Degrees of Freedom in the Social World:  
Towards a Systems Analysis of Decision*

MARIAM THALOS  
Philosophy, State University of New York, Buffalo

I. INTRODUCTION  

THOMAS SCHELLING taught us that in ordinary human affairs, conflict and common interest are ubiquitously intertwined.¹ For when it comes to variety, the occasion of pure conflict (known to some of its friends as the zero-sum game) is as under-represented in human affairs as the occasion of undiluted common interest (known as the pure coordination game). The undiluted extremes are the exceptions, when it comes to counting kinds, while the mixed-motive kind of occasion is the rule. Things look a bit different, however, when one looks at sheer numbers of true-life occasions, as I will explain.

Schelling also taught us that in the diverse space of mixed-motive affairs, intermediate between pure conflict and pure coordination, there is more true-life collaborative behavior—more trustings and promise-keepings—than our prescriptive decision theories can accommodate, never mind explain. This collaborative behavior, as long ago another Thomas—Thomas Hobbes—was painfully aware, requires some explaining precisely because of the presence of ineliminable conflict of interest. The problem confronting those who aspire to explain such collaboration is to identify weighty motivations in its favor, in the face of the weighty, but countervailing motivations.

What shall concern me here is not the collaboration that transpires in the face of conflict of interest, but the success which meets us more than halfway in those limiting cases of pure coordination, so as (for example) not to collide in roadways and corridors. For in sheer quantity, the occasions for pure coordination outnumber the occasions for anything else perhaps a hundredfold. We

¹I owe a special debt of gratitude to my colleague Barry Smith for, among many other things, giving me the title of this paper and directing me to write the rest. Therefore I can take responsibility only for material from the subtitle onwards. Special thanks also to Robert Goodin, who generously and in many ways supported the larger project of which this is a small part, and advised publication of this small part in the present venue. I would like to thank the members of the University of Manchester Department of Government and several referees of this journal for insightful and evocative remarks. Thanks also to Alison Wylie, Newton Garver and those in attendance at the Marvin Farber Conference on Applied Ontology, State University of New York at Buffalo, April 1998, as well as those present during presentation of materials from this essay at the biennial meeting of the Philosophy of Science Association, Kansas City 1998. Research for this piece was financially supported by the Center for Cognitive Science and the Baldy Center for Law and Public Policy, State University of New York at Buffalo.


©Blackwell Publishers, 1999, 108 Cowley Road, Oxford OX4 1JF, UK and 350 Main Street, Malden, MA 02148, USA.
successfully pass hundreds, if not thousands, of strangers in passageways every
day, distribute indistinguishable parcels and papers, and organize ourselves
around symmetrical tables, all without appointing experts in indecision
resolution, or even stopping to exchange words about how to proceed. And
the numbers in which we do this successfully, without missing a beat, outrun our
failures in magnitudes so remarkable, we hardly notice our efforts as efforts,
much less as successful ones.

Success at coordination is the most perplexing of all social phenomena,
because it is enjoyed by creatures that can appreciate, and share as common
knowledge, the fact that there are (inconveniently) many ways whereby to
achieve the common goal. And ever since Schelling we have been suspecting that
traditional game theory, even with the most up-to-date of Bayesian refinements,
will never explain this success. I propose here to confirm these suspicions, and
thus put an end to doubt on this point. I shall argue that traditional game theory
owes the resounding failure to its Kantian account of agency.

My strategy will be first to propose an anti-Kantian account of agency that—
for a very important reason—takes as its point of departure precisely the kinds of
pure coordinations that Schelling demonstrated are so difficult for traditional
game theory to explain. With this account of agency I shall develop a proposal
for handling pure coordination problems, and in the process forge an alliance,
not with the mathematicians, as the Bayesians do, but with a certain school of
cognitive and social psychologists propounding a simulation-based account of
sociality. The fruits of this alliance will be resources for a new science of society.

Then, once the explanatory goalposts vis-a-vis coordination come into clearer
view, because we will have seen how some account actually reaches the goal, I
shall argue that traditional game theory, founded as it is on the Kantian notion of
agency, cannot match the success of the proposal I am promoting. Finally, I will
show that my proposal is in a position to contribute to the resources for
explaining collaboration even when there is indeed conflict of interest: it shall
provide a missing factor which, together with existing motivation, can direct an
agent in one direction rather than another.

II. DECISION ANALYSIS AND ITS BAYESIAN MANIFESTATIONS

A sizable portion of the choices we make in ordinary life are social: they rest on
considerations of other people’s wishes and needs, as well as on considerations of
the effects of our decision on others. This is not to say that these choices flow
from some kind of altruism, but instead that they involve the anticipation of
others. For we make most of these social choices partly by looking over other
people’s shoulders to see, if we can, what they shall be choosing. And we view
other people as also looking over our own shoulders. The discipline that studies
such matters is known (somewhat to its misfortune) as game theory, and the
decision makers there studied are called *players*. The outcomes of “games” are outcomes for which all the players are partly responsible.²

Bayesian analyses of decision theory aim at treating cooperative games, defined as ones in which exists a potential for enforceable agreements amongst the players,³ as they treat non-cooperative games, which admit of no such agreements. And so Bayesians aim at representing collective decision making as a species of individual decision making, not as something different in kind. The goal is manifestly reductive, guided by the idea that cooperative games do not deserve a separate category of their own. It is to assimilate what may be called *collective rationality*, the process of deliberating collectively to achieve agreement for the sake of coordinating action, to what may be called *individual rationality*, the process of achieving decision as an (undistributed) individual. The general colonizing move is to handle the overt process of deliberations as a series of strategic bargaining steps, in a competitive game played out amongst the members of the coalition, within the boundaries of the larger game. The bargaining game is itself viewed as governed by independent rules of interaction amongst multiple players, and therefore clearly *not* as something which someone can undergo purely as a single, unified decision body. Under the Bayesian proposal, it thus becomes *impossible* to view the process of deliberation as a means of bringing into being a single, multi-member decision-making body aiming at collective goals. Instead, deliberation with others comes to be viewed exclusively as a means for each participating individual to reach an individual end, within a purely competitive framework. This is, for example, a goal of Harsanyi and Selten’s monumental work on equilibrium selection although, as they themselves acknowledge, the goal is never reached.⁴

The distinctively Bayesian procedure, in service of the assimilative goal, is to define a decision problem very abstractly, by “essentials” only. The defining characteristics of decision problems are allowed to include preference structures, rules of play, and knowledge that these things are themselves common knowledge to all participating. And participants, incidentally, are assumed to be consummately rational beings, in possession of common knowledge of this

---

²I shall use the term “decision theory” in such a way that game theory—which specializes in problems of strategy faced by two or more players each of whom is rewarded according to what all players choose—is a species of decision theory.

³The original classical definition defined them also in terms of the permissibility of communication; this has come to be seen as both unnecessary and counterproductive.

⁴A *General Theory of Equilibrium Selection in Games* (Cambridge, Mass.: MIT Press, 1988). In the Postscript they write: “This means that our theory uses two independent, and ostensibly very different, criteria of rationality. One of them, risk dominance, is based on *individual* rationality: it is an extension of Bayesian rationality from one-person decisions to n-person games involving strategic interaction among n players, each of them guided by Bayesian rationality. . . . In contrast, payoff dominance is based on *collective* rationality: it is based on the assumption that in the absence of special reasons to the contrary, rational players will choose an equilibrium point yielding all of them higher payoffs, rather than one yielding them lower payoffs. That is to say, it is based on the assumption that rational individuals will cooperate in pursuing their common interests if the conditions permit them to do so” (p. 356).
mutual consummate rationality, as well as unbounded deductive powers. The participants are then assumed to pass from knowledge of essentials about their situation, through putting to one side (if necessary) such things as cultural norms and all other common knowledge not specifically relevant to the decision problem, via principles of practical reasoning, to a solution that compels action. As I will explain, the Bayesian approach cannot avoid expecting the decision maker to be a consummate logician, and in that regard exaggerates the human endowment. It shall be my thesis that the Bayesian approach at the same time understates the agent’s nonlogical resources, by denying her the capacity for forming true alliances. And thereby Bayesians deny themselves the ability to explain true-life feats of coordination. We will then explore how we must mend our view of the human endowment, so as to make explaining coordination, among other things, possible.

III. NASH EQUILIBRIUM

The Bayesian approach to coordination forms one stream of the so-called Nash Program, which commends solutions to dilemmas in the form of a Nash equilibrium. To define a Nash equilibrium we shall need to introduce the notion of best reply. A certain option for a given player is a best reply to other players’ choices, if it best advances that player’s aims (in Bayesian terms, maximizes his utility), in light of the other players’ current choices. Then a Nash equilibrium is a combination of plays, one from each player, such that each one is a best reply on its player’s part, to the other players’ choices. A Nash equilibrium point is attractive to game theorists because it is self-enforcing, in the following sense: if each player believes that the others are doing their parts in the equilibrium stratagem, then each can be certain he can do no better for his aims by unilaterally deviating from equilibrium, and so will follow through, without there being a penalty imposed for deviation. For each player’s part in the equilibrium is a best reply for each to what the others are doing. Now, an option which is a best reply to other players’ choices, no matter what the others choose, is called dominant. A dominant strategy is consequently an unconditional best reply.

5I will not survey the variety of solution concepts here, as it will take us too far afield. Suffice it to say simply that all solution concepts so far proposed, with the possible exception of correlated equilibrium, have had the minimum property of being Nash equilibria. A sophisticated treatment of this issue is Michael Bacharach, “A theory of rational decision in games,” Erkenntnis, 27 (1987), 17–55.

The sticking point—the essence of the coordination problem—is that coordination games possess multiple, non-interchangeable Nash equilibria, no one of which is also dominant. For example, suppose you and I each are given only one opportunity to name either Heads or Tails, without communicating; we win a prize so long as our selections match.\(^7\) The outcome of the equilibrium point Heads/Heads is just as agreeable to us as the outcome of the equilibrium Tails/Tails. So the trick is to reach one equilibrium point, rather than the other, without benefit of communication.\(^8\) An unembellished Bayesian approach, which simply directs its adherents to maximize individual utility, without giving further instruction in the matter, does not provide a mechanism for deliberating among competing equilibria, so does not prevent miscoordinations.

Now in Schelling’s own informal empirical studies, participants in the Heads-Tails game were able to converge upon Heads at a rate of 6 to 1.\(^9\) (Heads would be your choice too, I will wager.) Why? Schelling draws our attention to the fact that the choices (Heads and Tails) are labeled differently, and also to the supporting fact that Heads enjoys a psychologically more prominent or salient position in the Heads-Tails scheme. We are invited, with Schelling, to admire players for selecting Heads, having grasped some—any—mutually graspable asymmetry in the matter. But our admiration cannot substitute for an account of why the prominent should win out over the underemphasized. After all, the fact of asymmetry does not itself explain why the underemphasized should not be elected.\(^10\)

There is now a very large literature devoted to refining the Nash equilibrium concept, so that the refinement will both (1) eliminate intuitively incorrect equilibria, and (2) leave only one equilibrium left standing as an option. For once upon a time it was thought that, if there is one unique Nash equilibrium point, then every theory, no matter what else it might say, can embrace it as the solution to the game, since it has no competitors.\(^11\) This is a quite problematic proposal, for it is tantamount to simply and unceremoniously postulating that every coordination problem possesses (for example) a Bayesian solution—that, in other words, every coordination problem possesses a solution that can be supported purely on individual utility-maximizing grounds. But there are at the same time very serious reasons for embracing the contrary of the existence postulate.\(^12\)

\(^7\)This, and many other similarly compelling examples, first made appearance in Schelling’s pathbreaking *The Strategy of Conflict*.

\(^8\)In an important sense, communication does not really help. For if I am allowed to tell you (but not show you) that I am selecting Heads, how exactly does this give you a reason to choose Heads yourself?


\(^10\)In fact, there are circumstances where the under-emphasized deserves to win, because being under-emphasized is important under the prevailing conditions.

\(^11\)The monumental Harsanyi and Selten, *A General Theory of Equilibrium Selection* (Cambridge, Mass.: MIT Press, 1988), for example, rests on this idea, as do many other refinements of Nash equilibrium.

Michael Bacharach, who is perhaps the only person to axiomatize what may plausibly be called the theory of noncooperative games in Bayesian style as practiced in the 1980s, has proved that such a theory, which is without principles of deliberation beyond “Maximize!”, has no solutions to games with unique Nash equilibria.\(^{13}\) The ground of this proof is quite simple, and is captured in an argument, put forward independently by Margaret Gilbert, against use of salience as potentially identifying an outcome as a (unique) solution. The argument, in effect, is that even in the presence of a salient strategy combination, there is still nonetheless no reason for one player to believe that his co-players will or are likely to perform their parts in that combination.\(^{14}\) So if there is no reason for me to think you will choose your part in that prominent combination, I have no compelling reason, not even on grounds of maximizing utility, to choose my part in it. And so prominent considerations are not rationally compelling in the way they need to be in order to qualify as Bayesian solutions.

One promising Bayesian approach, appearing since Bacharach’s work, is the idea that dilemmas ought to be individuated also in terms of initial beliefs, because beliefs—or at least initial beliefs vis-a-vis potential solutions—fall into the category of essentials of a decision situation. Brian Skyrms has perhaps the most prominent proposal, as well as the best worked out, along these lines, but Cristina Bicchieri has another.\(^{15}\) The basic idea is simple. They say that in order to achieve a specific enough definition of a dilemma to reach a compelling solution, we must model also the knowledge of players as to how they assess the others are initially likely to play. And also to model the means by which they update systematically their opinions of each others’ opinions, iterating the updating process until there is convergence.

The results of this approach are in fact interesting, but, as I now argue, they are also of limited use. The approach requires that an initial assessment of the opponent or partner’s play is ascribable to each player by the others. These assessments are called prior probabilities or simply priors (as known to friends and enemies alike). Allowing priors to individuate between games may be very much in keeping with the Bayesian program—which, while embracing much diversity among its adherents, is nonetheless unified by the idea that it is exclusively antecedent attitudes (towards, among other things, propositions) which dictate subsequent turns of mind and will. But the priors proposal runs into the following, fundamental difficulty. If players fail, for whatever reasons, to form priors, or simply fail to form them in a way which is commonly known to be common knowledge, then such factors as would otherwise seem relevant (such

\(^{13}\) Bacharach, “A theory of rational decision in games.”


as, for example, that some equilibrium point is in point of fact most salient, despite its not being common knowledge that it is so) cannot be used to discriminate one dilemma from another. And so the Skyrms proposal will have an application only in those cases in which factors that otherwise seem relevant to individuating one dilemma from another also show up as making an impact on priors. (So the new proposal will succeed in identifying solutions where, by contrast, the unembellished Bayesian approach fails, only in those instances where players have at least comparative opinions of each proposition to the effect that the opponent will do such-and-such, that these opinions are common knowledge, and that it is also common knowledge that players update this opinion systematically, according to some definite rule.) But this is a tall order, even when a salient solution is known to be salient. For example, in our Heads-Tails dilemma, each player might know that heads has a psychological first place for the normal citizen, without having any reason to think that the opponent is more likely to choose heads (Gilbert’s point again).

Now Bayesians generally promote the manufacture of (logically consistent) priors if, on introspection, we do not find them on the shelf (as it were). One wishes to object here, for rather delicate reasons. For it is not that one is against the manufacture of priors, as such. How could one be? But rather one wishes to object to the demand for priors on the grounds that issuing a mandate for their manufacture is out of place, as an edict of strategic rationality. On at least two counts. First, if I make no estimates at all, implicit or explicit, vague or definite, as to the chances that you will choose Heads, how is it the business of a theory of choice to insist that I do so? For how can this insistence be justified in terms of how my interests will be served? The business of a theory of choice is to tell me how to advance my interests; it is not to tell me how to advance my opinion or knowledge as such. For, if I can produce priors at will, why should I not produce them so that their magnitude is to my advantage, rather than apportioned according to the evidence? The point is simply that epistemological imperatives cannot simply be a subcategory of practical imperative.

The second ground for objecting to the priors mandate is simply that holding any prior whatever is liable to introduce logical inconsistencies into my corpus of belief, in the presence of a common knowledge assumption (an assumption that everyone knows what everyone else knows about the decision problem as such). We might bring this point out as follows, vis-a-vis the Heads-Tails dilemma. Suppose I believe you are more likely to choose Heads than Tails, possibly on grounds that Heads is psychologically more prominent for you, and also that this is common knowledge between us. This belief, in the presence of a common knowledge assumption, presupposes that I have considered the impact of my own

---

mind on the likelihood of your playing Heads. But this, in turn, can be true only if I know what I will do, or have some estimate of it, and this cannot be true unless I have already completed deliberations, or at least begun them, which has got to be false if I am forming a prior only. Skyrms defenders are certain to object that I have confused priors with posteriors—thus confusing the process of deliberation with its end products—and that anyway inconsistencies tend to disappear as rational deliberation advances in real time. But even if they are right (and I am not yet prepared to concede the point), my point will simply be, why should we begin deliberations using priors in the first place, if they are the sorts of things that require subsequent refinement or correction? Particularly if something more ready-to-hand is available.

The Bayesian might reply that a directive to rectify belief, so as to achieve greater likeness to the truth, is a decidedly desirable feature of their prescriptive enterprise. For they propose to explain human-originated events by drawing attention to what individuals ought to do in the interests of self, via belief if necessary.

I agree that we can no longer explain social behavior without speaking about what (objectively) promotes the interests of individuals, in evolutionary terms, since what we aspire to explain is not individuals’ actual rationales for what they actually do, but rather what contributes to the success of the actual behavior. Even so, it does not follow that we will be allowed to prescribe X to an individual, when something lesser, Y, will do equally well. This principle of economy too is among the first principles of evolutionary thinking. I think something less costly (again, in evolutionary terms) than full-fledged priors will do for purposes of achieving the observed coordinations, and therefore that Bayesian tolls are extortionate.

Now the Bayesians might simply insist that I always do have a degree of belief in the propositions they are drawing attention to, whether I know it or not, if only I will reflect on the matter. I have only to Ramsify or Savagize (think about choosing between certain gambles) and these exquisite items will present

---

17This resurrects an objection to priors that Skyrms believes he has answered to tolerable satisfaction, to the effect that it is illegitimate to make assessments of one’s own likelihoods, to which Skyrms answers we need not. My point in the text is that a likelihood assessment of others, in the presence of common knowledge, presupposes a likelihood assessment of self.

18Consider Schelling’s Holmes-Moriarity example. They are on separate trains, unable to communicate, having to choose the station at which to get off, and both will be rewarded if they succeed in getting off at the same station. We can agree that each of Holmes and Moriarity will get off at station A if he expects the other to do the same. And that he will expect the other to do the same only if he expects himself to expect the other to do likewise, and so on. The problem of inconsistency, as I see it, is that those who use priors use them to argue also that Moriarity (for example) infers that he expects himself (Moriarity) to get off at the next station from the fact that he expects Holmes to expect him to get off at the next station. They have got inference engines working in both directions at the same time. At the very least, it is hard to keep track of where one is in the deliberation process.

19See Adam Morton, “Folk psychology is not a predictive device,” Mind, 105 (1996), 119–38, for related points about entangled anticipations.
themselves for inspection, materializing out of thin air if necessary. So since they are already more or less on hand, I might as well put them to use. But again: why is it the business of a theory of choice to insist that I bring these particular priors forward (assuming they are on the shelf, and just require dusting off), and subsequently correct them? Why are beliefs about what the other person is initially likely to do relevant to my current coordination problem when beliefs about their hair color, as such, or some other salient property of our situation, as such, are not? Surely a theory of choice must have a prescription for me, even if I do not or cannot form a full complement of priors, and should direct me to put aside all such beliefs as would foreseeably lead to an undesirable end.

In summary, I view the demand for priors as a case of making a virtue out of a necessity—a necessity that in any case is simply an artifact of the Kantian foundations of Bayesian theory, as I will now endeavor to explain.

IV. DEPENDENT COORDINATIONS: AN ANTI-KANTIAN PROPOSAL

True, the regularities in human social behavior rest on the demands of strategic rationality, or prudence as I prefer to call it. What are those demands? To my mind, decision theory as we know it today, because it has been shaped by Bayesian dogma, has made serious mistakes in articulating the nature of prudence, primarily by embracing a Kantian view of agency. The trouble with the Bayesian proposal is that the very metaphysics of agency is made to disappear, and is replaced by a logic of decision.

Kant, as is well known, put forward the proposal that autonomy is a property of the rational will. Now it is this idea, that a rational individual is not one who is moved externally, on which rests the Bayesian presumption that an agent requires positive reasons in favor of one course of action over another, and so requires some kind of deduction, as the solution to a dilemma. In the Kantian style of analysis that Bayesians adopt, prudence consists purely of adhering to principles or propositions. Since everything that matters for the purpose of evaluating principles also falls in the category of propositions, at which internal mental states like belief and desire can be directed, the analysis of decision can subsequently be conducted in the (logical or formal) space of propositions and principles. Because the propositional content is all that matters, the metaphysics of agency itself is forced off center, while logical analysis takes center stage. Decision analysis becomes the preserve of the logician-mathematician. This is the dogma I wish to challenge.

When two friends lose each other in a crowd, how do they proceed to find each other again? As we may say, it depends. For the answer to their problem cannot be reached \textit{a priori}, but depends on the nature of their venture up to the point of getting lost. Each might find herself heading towards the point of origin, if that is sufficiently close by. Or towards the destination point of their journey, if that is closer. Or some intermediate point, if it exerts more pressure on each of their sensibilities than does either of the first two options. But if separated from my toddler, or my dog, on the other hand, I do not head towards a departure or destination point, nor do I direct my path towards a prominent intermediate point, but instead seek out wandering places likely to attract the attention of a toddler or dog, close to the point of being separated.

I will draw two morals from these contrasting examples. The first is this: whether they succeed in coordinating or not, the adults choose as they do partly—but only partly—because the option each chooses enjoys a psychologically more prominent position on their spectrum of perceived options. But it is the \textit{rest} of the story as to why they choose as they do which is most interesting, philosophically speaking. I propose to tell it as follows: under circumstances when coordination with others is required and there is an appropriately prominent strategy combination, individuals who perceive the prominent combination \textit{as} prominent, typically find themselves in possession of an \textit{impulse} to act so as to realize it as the solution. (They could experience also secondary impulses to resist acting on the original impulse. And these secondary impulses may be stronger, as impulses, than the original. But notice that this fact does not diminish my point—to the effect that impulses come into play and require taking account of.) The impulse towards the prominent is a species of compulsion rather than a species of reason, which nonetheless admits of being either overridden or bypassed under appropriate conditions, as well as it admits of being complied with.

I propose, in other words, to understand certain features of circumstance as directly exerting influence on agents, in such a way as gives rise to conditional impulses to act. The properties of circumstance can give me, on occasion, not a \textit{reason} for acting, but simply an \textit{impulse} towards a certain action—an impulse whose force can be weighed against that of reasons, as well also as against that of other impulses.\textsuperscript{21} Thus when our efforts to coordinate succeed, this success is explained by factors both external and internal to our minds. And the external factors are features of our decision-making body or \textit{system}. For if it did not belong to the system, \textit{as} a system, our behavior would not be \textit{coordinated}; it would be merely correlated. This difference explains the fact that we succeed even when circumstances are perturbed slightly in this or that direction. Merely

\textsuperscript{21}This solves Gilbert’s problem by rejecting the proposition that only a \textit{reason} for action can rationally justify a piece of behavior on grounds of prudence, and embracing instead the proposition that compliance with certain impulses is permissible so long as no compelling reason is raised \textit{against}. 
correlated behavior is not stable or robust under a wider range of variations, whereas coordinated behavior is.

Thus, when separated from you, I am moved in the first instance to choose a prominent point not because I have reason to believe you will choose it, or a reason to believe you will think I will choose it, or anything of the kind. (Although after the fact I will be entitled to say that I believed you would do likewise because you had the same perspective on our situation.) I choose in the first instance simply because the point of reference exerts a certain pressure directly on my will, which I can but do not elect to resist. I receive from my circumstances an impulse which deserves compliance because I have no reasons to override it. And having no reason is not itself another reason: it is, just as the term has it, a lack or insufficiency of reasons. (More on this matter and other metaphysical matters in sections VI–IX.)

The second moral I wish to draw from my examples concerns the differences when it comes to coordinating with a toddler or pet. I give in to the impulse to seek a peer somewhere that (for possibly unrelated reasons) draws my own attention, because I have no reason to believe that if I took my peer’s perspective I would be choosing differently. (More on this in sections VI–IX.) But the story is quite different when it comes to being separated from my toddler. I must bypass the impulse I receive, as myself, in that case. And I must take his perspective. I try to imagine the situation from his point of view, confident that he is positively not perspective-taking like me. I model him neither as a maximizer of utility, just like myself, nor as a deliberate duplicator of my reasoning processes—although he is both, I am quite certain, at least on many occasions. Perspective-taking is not the same thing as modeling the other party after self-as-rational, although in the case of the two friends, the two are indistinguishable in their results.

Perspective-taking is normally an exercise in practical rather than theoretical phenomenology. It is a simulation of the other’s experience, as an experience, from the inside, normally so as to achieve a solution to a choice problem. And in any case, perspective-taking is never, due to its nature, conducted from a perspective-free style. It is simulation; and simulation of you is done from your eyes’ point of view. Simulation makes your perspective, as a perspective, present to me, and in a position therefore to supplant another perspective (if I so let it) in a controlled fashion—running in a Window, so to speak. It consequently renders your perspective capable of being something from which I, just like you, can take action. My simulation of you makes it possible for your perspective to be the ground of our joint, successful coordinative effort—an effort that can leave my own, abandoned perspective in the dust (doing absolutely no work from a systems point of view). The normal product of an exercise in simulation is a piece of behavior. Thus a simulation is, as such, a piece of practical reasoning. Moreover, it lends itself to a systems analysis of our practical reasoning, as I will soon discuss. The systems analysis will allow us to view factors external to
ourselves (for example, factors which generate impulses) as acting on the decision-making unit, as a unit.

By contrast, the modeling of the other after self is an exercise in deduction of the other’s behavior from certain modeling principles. As such, it is an exercise in the scientific enterprise. If successful, its product is an explanation of a piece of behavior, not another piece of behavior. Modeling is an exercise in theoretical rather than practical reasoning. (To suppose otherwise is to collapse practical reasoning into theoretical reasoning, where—ironically—the Bayesian aspires to achieve collapse in the reverse direction.) Modeling, as such, is the enterprise of representing the other’s internal state for scientific or theoretical purposes, not the enterprise of duplicating it in a controlled fashion for practical ends. And one gains no advantage from looking at a pair of model-framing individuals through the lens of a systems analysis: the system they compose, as a system of action takers, is a purely limiting case of systems analysis—a case in which systems analysis gives no distinctive result.

V. EMPIRICAL COGNITIVE PSYCHOLOGY

The term perspective-taking originated in the 1930s, in the work of Jean Piaget and G. H. Mead. Together with the concept of simulation, that of perspective-taking now plays a central role in the contemporary scientific writings of a school of cognitive and social psychologists, eminently represented by Paul Harris and Robert Selman. Rediscovery of the idea of perspective-taking has kindled criticism of the research agenda in cognitive psychology which has come to be known as the theory theory. Theory theory is to psychology what Bayesianism is to the economic sciences. According to the theory theory, all predictions of another’s behavior (and particular attention is drawn to children) are due to none other than a proto-scientific, representational theory of mind, that is acquired by the prediction-maker in the normal course of human development.
theory opponents hold that the representational ("theory") mode of presentation is not the only mode in which humans cope with each other. Thus the debate I am hoping to kindle within the normative discipline of decision and action theory has an empirical counterpart in the empirical realm.

The empirical evidence, as to whether children in particular simulate predominantly or represent, is inconclusive. The theory theorists claim that simulation theory cannot account for all the evidence, and the simulation theorists proceed to show just how a simulation theory can do precisely what is claimed it cannot. But the evidence concerning adults has been around for some time: adults, without question, exercise a capability for simulation. And surely there is no good reason to deny adults both capacities.

That humans exercise a power for simulating other perspectives, and do so on some if not all occasions, strongly weighs against treating individual agents as decision "atoms." Such a treatment will be, at best, suboptimal. For the decision to treat individuals as atoms goes hand in hand with a decision to view individuals as, fundamentally, independent systems. Then spotty correlations amongst the atoms can be handled as a matter of incidental, "external" interactions between them. But in fact this simplifying model of the fundamental character of human affairs is misleading at best. (At worst, it is grossly oversimple, as often happens with reductive proposals.) Humans in community form systems. In fact humans form systems of nested systems. Thus human social behavior is best handled within a systems approach. The reductive alternative makes the social, as social, disappear, as social psychologists have been complaining for many years.

Harris, for example, criticizes theory theory for oversimplifying the human cognitive toolkit, and thereby obscuring the very pronounced, qualitative differences between the child (and the adult in ordinary life), on the one hand, and the scientist, on the other. These qualitative differences, he feels, have to do with the role of the imagination in the contrasting enterprises. And Selman writes that "the greater the emphasis we put on explanations that are based within the individual, the more diffuse will appear factors of a more sociological nature." Thus theory theorists tend to obliterate the qualitative difference between the exercise of powers that are purely intellectual, and the exercise of powers with a distinctively social component or dimension. For theory theory does not recognize the characteristically social character of social life: it aims to reduce the social to the purely intellectual, just as Bayesian decision analysis aims to eliminate corporate decision making as a special branch of decision making, distinct from individual decision making. According to Selman, putting oneself in another's position, and subsequently considering one's own actions from that

26 P. Harris, "The work of the imagination," *Natural Theories of Mind*, pp. 299–302.
alternative position, is the operation, uniquely human, that allows for the existence of a social self.  

Ultimately, my criticism of Bayesianism’s belief-desire style of analysis will be that it is in principle unable to explain the overwhelming success we enjoy when we attempt to coordinate in cases where there is absolutely no conflict of interest. This criticism is best displayed after we have seen an alternative to the Bayesian proposal, which (as I shall show) explains what needs explaining. I shall afterwards advise the coordination of perspectives, in the name of success, where my Bayesian counterparts counsel modeling others after self-as-rational. Abandoning my perspective, after an exercise of perspective-taking, is what is called for when I seek to coordinate with my toddler; indeed I should be faulted for modeling him after myself-as-modeler. So why suppose something qualitatively different is required when I seek to coordinate with my friend? 

By advocating perspective-taking I shall also explain the significance of salience considerations for coordination, whilst agreeing that they cannot bear the entire burden of explaining why people choose in accordance with them. On my account, two peers who aspire to coordinate for the sake of being reunited, should not choose purely on grounds of prominence, or even on the assumption that the other (like self) sees a certain option prominently; I agree with Gilbert, that prominence is not enough as a reason. But it is enough to generate an impulse. And that can be enough, full stop. Hence I am recommending the prominent solution precisely in those instances in which the prominent solution gives rise to the same impulse to act in all would-be coordinators, an impulse that is not overridden but instead reinforced through the social act of perspective-taking. Thus the grounds on which I say coordination is achieved are precisely those on which it is permitted by prudence: coordination rests, not purely on reasons, but on a combination of reasons and impulses.

Schelling, I believe, was groping towards this anti-Kantian view of agency, but was nonetheless reluctant to make a clean break with the requirement of deduction. He wrote, for example, that certain signaling properties of circumstance are the sorts of things that “seem to ‘point toward’ [one equilibrium point over against another]. They provide either a reason or an excuse for believing or pretending that [one equilibrium point is better than another]; since we need an excuse, if not a reason, for pretending, if not believing, that one of the equilibrium pairs is better, or more distinguished, or more prominent, or more eligible, than the other.”

Now a critic will no doubt suggest that perspective-taking, if required at all by prudence, is required only instrumentally, for the sake of enabling a prediction of

\[\text{Ibid.}\]
\[\text{Adam Morton, “Game theory and knowledge by simulation,” } \textit{Ratio}, 7 (1994), 14–25, offers similar counsel.}\]
\[\text{The Strategy of Conflict, pp. 297–8.}\]
my toddler’s behavior, or my pet’s, and so assists me in establishing an opinion that will serve as a prior. This idea, to my mind, over-intellectualizes the process I undergo when I coordinate with my little one, so over-intellectualizes what prudence requires. (The prediction requirement is precisely the point at which the Bayesian demands are extortionate, exaggerating the demands of prudence.) For I do not take my toddler’s perspective in order to predict his behavior. I do it for less intellectual reasons. I take his perspective in order to put myself in a position to be directly in contact with those same psychological factors as give rise to his impulses. It should come as no surprise—if indeed it is true—that perspective-taking is a good means to predicting his behavior. Indeed, if what I am saying about the toddler is right, then perspective-taking is arguably a necessary step in achieving a correct prediction of his behavior. But predicting behavior is not always a requirement of prudence; what prudence requires is (at least in our examples) coordinated action. So, when asked to make a likelihood assessment of my toddler’s being found where I am looking—when in the grip of that electrifying fear for his safety which compels me to take his perspective—I, for one, should certainly give the Bayesian no satisfaction; I will be unable to produce any such assessment, except after the fact. Can I be faulted, on grounds of prudence, for having made no such likelihood assessment? If the answer is—as I believe it is—a negative one, then the Bayesian proposal requires too much as a theory of prudence: it makes a virtue out of a necessity that only a Kantian would accept as a necessity. And by the same token its model of the rational agent is overly intellectual.

VI. SYSTEMS ANALYSIS OF THE SOCIAL

We can’t understand human life merely in terms of individual subjects, who frame representations about and respond to others, because a great deal of human action only happens insofar as the agent understands and constitutes himself as integrally part of a we.32

Ordinarily, when each of N persons has a decision to make, which impacts also what each of the other N–1 persons will receive as a result of these N decisions, we suppose that there are N degrees of freedom in the situation, one for each decision maker. (I will speak as if degrees of freedom come in packages of one per decision maker. But as will become clear in Section VII, this is too simple.) But if parties to such a decision problem seek to coordinate their behaviors perfectly, they will seek to choose as if they were a single decision-making body, so that the number of degrees of freedom is reduced from N to one (and thereby also the potential for mis-coordination). This, at least, is what happens in the ideal cases, when the attempt to reduce degrees of freedom actually succeeds. The reduction

to one degree of freedom is not always the aim, nor is it even the rule in human affairs. When a group of N voters each has to cast a vote for one of two candidates, and there are two political parties each sponsoring a candidate, and there is widespread party loyalty, then efforts will be expended to reduce the number of degrees of freedom from N to 2, not one.

Reducing degrees of freedom is a piece of social engineering. As such it typically takes hard work on the parts of those who wish it done. The attempt can go wrong in many ways, many more ways in fact than it can go right. But it can go right also in a number of metaphysically distinct ways, as now I wish to illustrate. It will be important for our purposes to mark the differences between the two ways I will now describe.

When our two separated adults seek to coordinate their actions for the sake of being reunited, they seek to act collaboratively. When they succeed, they reduce the number of degrees of freedom in their decision problem from 2 to 1, by sharing (as I will call it) the remaining degree of freedom; they seek to act as if the two of them were each one half of a single decision-making unit. They seek for their choice to be conditioned upon the choices others make; they seek to make their selection dependently, or nonautonomously.

The idea is this. When two of us favor a coordination, and each knows of the other’s favorable view of coordination, and each approves of the other’s favorable view, and each knows of the other’s approval, and so on, then each of us is in the best possible position for forming, on an impulse, a collaborative intention, as such—an intention to act in concert rather than to act unilaterally, an intention to act as part of a larger action-taking body.33 And only subsequently, and derivatively, does this intention proceed to give way to individual intentions to do one’s part in the collaborative enterprise. This sort of thing can occur only in what Selman calls Level 3 perspective taking, in which what he calls mutual perspective taking begins to emerge, and coordination of multiple perspectives becomes possible. Level 3 is where perspectives first become nested in other perspectives. This capability for coordinating perspectives in a controlled fashion is present in ten-year-olds.34 From the collaborative intention formed during a Level 3 activity can flow an individual reason for acting in one particular way rather than another (for example, a reason for choosing an equilibrium point which awards each of us more than any of the other equilibrium points) or, if conditions favor it, an impulse towards a certain action. Provided there is no countervailing reason (or impulse), the individual intention in that instance will carry the day. And this proposal, on my view, is the only one

---

33Bicchieri, *Rationality and Coordination*, views this iteration of knowledge and approval as an idealization, which may not hold good when we move away from a single decision maker. I view it as perfectly reasonable and nonidealizing, since each individual is undertaking it, at this stage, from the point of view of a single decision-making body. But I agree it is quite an idealization if the individuals undertaking it are Bayesian atoms.

that can account for the commonplace coordinations that a Bayesian account cannot aspire to explain. Here is why.

What a collaborative intention does, subsequent to its formation, is to correlate the behaviors of collaborators. This is how it functions to serve the interests of coordination. Thus, metaphysically speaking, what a collaborative intention does is reduce the number of degrees of freedom in, for example, a two-by-two decision problem, for a system of two decision makers, from two degrees to one.

My criticism of the Bayesian approach, now, is that it will never be able to duplicate the successes of a simulation account, purely because it cannot guarantee reductions in degrees of freedom in the way that a simulation account can. In other words, the Bayesian approach cannot accommodate a systems result, except artificially. Whereas a theory of agency built upon a simulation theory of sociality quite handily treats reductions in degrees of freedom within the compass of what the account takes itself to be responsible for explaining. When I simulate you, the number of factors that determine the outcome of the system composed of the two of us is reduced by at least one. For if I am successfully simulating you, then, in Harris’ terms, one of the default settings, in my simulation, matches a factor that determines the result of your practical reasoning. Therefore we are more likely to coordinate successfully. This sort of reduction of factors cannot occur in a Bayesian account, except artificially, by the theoretician herself setting the defaults to the same thing, by hand. This is in contrast with the convergence of settings being the result of a process described within the account. Thus, on the simulation approach, reductions in degrees of freedom lie within the boundaries of what the account takes itself to be responsible for explaining. This is not true of the Bayesian account.

Reductions in degrees of freedom, when they occur in a Bayesian theory like, for example, Robert Aumann’s account of correlated equilibrium, do not occur as part of what we the agents do (for example, by taking each other’s perspective), and what the decision analysis on offer describes us as doing. Thus the Bayesian account does not explain coordination, but leaves it (unexplained) as a matter of chance. Aumann’s account achieves accommodation of the phenomenon of coordination, but does not achieve its explanation.

Now, explaining coordination with a toddler or pet will naturally be a bit different. And the proposal I am now making will handle the differences in a very natural way. This is now the other sort of way in which things can go right. When I seek to be reunited with my lost toddler, who himself seeks to be found, we too seek to act collaboratively. And when we succeed in our effort, we reduce the number of degrees of freedom in our decision problem from two to one. However, in our case we do it not by sharing the remaining degree of freedom equally, as friends equal to each other in strategic maturity do. Instead I, as the parent, relinquish any claim to a portion of the degree of freedom we aspire to share to my child, who inadvertently, by default, claims it all for himself. I let his
perspective on things be my sole guide, because I know he cannot coordinate perspectives as competently as I do: he cannot act as if one half of a single decision-making unit, due to his extreme youth. His claim to the whole of the single degree of freedom, though not unchallenged by my claim to it, nonetheless overrides mine. His deficiencies give him first title to it.

Whose perspective will win is not a matter of justice or principle; it is a psychological matter, a contingent matter. Thus there has got to be theoretical space for handling it within the account. And this is why we need an account of the metaphysics of agency. But this is precisely what the Bayesians have no space for in their exclusively propositional framework. Within the Bayesian world, the contingent occurs purely as a matter of chance.

VII. MORAL METAPHYSICS, AS CONTRASTED WITH MORAL MATHEMATICS

How, precisely, does reduction in degrees of freedom occur? The matter has, up to this point, been sketched in only the lightest hand. I propose in this section to fill in a bit of detail. The full story, however, requires the space of a book.

The Bayesian views decision making as purely a matter of maximizing expected utility. The operation of maximizing is regarded as an autonomous operation on two ostensibly independent inputs: (1) beliefs (as to probabilities of states of affairs external to the agent), and (2) preferences. The autonomous operation of maximizing is then construed as a formal or mathematical operation.

By contrast, the anti-Bayesian proposal I am now sponsoring views decision making as involving also the (non-formal, non-mathematical) operation of impulses, which act upon an agent either through the normal functioning of perceptual organs, or the functioning of those organs in a more agent-directed mode: that of simulation or imagination. So decision making, as the function of practical reason, here relies upon beliefs, preferences, plus the non-propositional, non-formal nudges that arise naturally as a result of perception and quasi-perception. How does reduction in degrees of freedom occur? Through non-accidental correlations amongst operations of perceptions and quasi-perceptions, as follows.

Let us say that each decision maker, at any given point in time, is in possession of three sets of inputs: beliefs (B), preferences (P) and impulses (I). Let us stipulate, merely for the sake of the illustration, that these three sets of inputs are at least initially independent. So a two-by-two decision problem, in which for example two separated companions seek to be reunited, is a problem with six

35This too is an idea that Schelling, The Strategy of Conflict, discussed brilliantly. He taught us that formidable capability does not always translate into advantage, and weakness can oftentimes be better utilized to advantage than strength. For example, a nation state unable to stabilize its economy is in a position to achieve economic assistance it might otherwise be refused, and a weak intelligence can provide protection against threats.
degrees of freedom—six sets of independent factors that determine the outcome. At least, this is so prior to simulation. Take, now, the case of myself and my toddler. I simulate his perspective in hopes that I will be prompted to move in his direction. If I conduct this simulation correctly (but of course it goes without saying that I am, sadly, liable to err), the two sets of initially independent impulses become correlated: my original impulses come to be (if I am successful) replaced by his. We still have two triplets of inputs: toddler beliefs ($B_t$), preferences ($P_t$) and impulses ($I_t$), plus parent beliefs ($B_p$), preferences ($P_p$) and impulses ($I_p$). But now ($I_t$) and ($I_p$) are not independent: if I am a skilled simulator of my toddler, ($I_t$) and ($I_p$) are identical. Voila! Only five degrees of freedom. In terminology borrowed from Harris, some of the default settings, in my set of inputs, have been made to match my toddler’s.

Moreover, simulation theory explains how this matching comes about: it occurs thanks to the process of simulation. A theory (such as Aumann’s) which merely postulates that ($B_p$) and ($B_t$) simply do contain such elements as would facilitate coordination does not give us the how of it. Or if it so happens to do so, because ($B_p$) and ($B_t$) are suitably related, by accident, the how is independent or external to the account of action itself, which is simply the formula ($B$)+($P$)→action ($A$).

Now a certain problem arises for my proposal. Not in connection with the toddler-parent case, which is probably the simplest sort of instance of how simulation can achieve a solution to a coordination problem. But in connection with what I have claimed is the case of coordination most numerous in its instances: coordination amongst equals. For if two such individuals endeavor to simulate each other, and do so successfully, then each will succeed in doing what the other ostensibly would have done. It looks, then, as if the two have swapped roles. So if they were not heading in the same direction before the swap, they are not doing so afterwards either. Hence simulation in that instance, if done skillfully, does not hinder. But neither does it help.36 We start with six degrees of freedom, and end with six, which have been somewhat reshuffled. Someone might even suggest that for a simulation to be complete, the simulator must simulate the simulations that it simulates, if simulations are expected on both sides. So there will be an infinite regress of simulations, and hence nothing but a blank screen, awaiting the result of a non-terminating process.

Both these difficulties focus attention on the thorny and potentially controversial issue of the nature of the process of simulation. This is not the place to launch into a treatment of that matter. But we can say enough to alleviate the difficulties just raised, as follows. When I undertake to simulate you simulating me, I do not get a completely blank screen. For there is much more to your perspective, which I am aiming to simulate, than your mere efforts to

---

36I am grateful to anonymous referees for bringing my attention to both this difficulty and that described in the following paragraph.
simulate our mutual simulations. It is this more, in addition to the mutual simulations, which does the work that needs doing—the work of reducing degrees of freedom. In the present case, for example, I exert efforts also to simulate the scene of our predicament—our immediate locus of being separated—as you view it; and this does not depend upon my being able to simulate you simulating me. We can call this simulation of your scene, a first-order simulation. When we share a degree of freedom, say for the purposes of pure coordination, what we do is compare elements in a first-order simulation with their counterparts in our own, unsimulated scenes. Only upon identifying such a match, say in prominent elements, will either of us undertake to accede to an impulse. When we seek to coordinate with equals, we are, essentially, choosing to act only on impulses which either derive from or survive such a filtration process. We defer acting on impulses that arrive through our own perspective alone, and act only (say) when we identify matches. (By contrast, when I relinquish the degree of freedom to my toddler, in hopes of duplicating his moves, I abandon my own perspective in preference to his.) Thus, in the case of the adults seeking to be reunited, reduction in degrees of freedom occurs when both of their impulse sets are replaced with something that deserves calling the collective impulse. Rather than two independent sets of impulses, \( I_1 \) and \( I_2 \), we get two copies of a single set of impulses \( I_{\text{collective}} \).37

The secret to coordinating, then, is deferring action until an impulse with the right credentials (for example, one which matches the unsimulated impulse) is produced by the simulation process. In this way, then, we avoid merely swapping. And we avoid too the delay until the simulation process will be complete: it is enough that we identify a unique match between the original and the first-order simulation. We do not need to check that the match appears also in second-order simulations.

Now someone might suggest that the process of simulation can be utilized to enrich the bodies of beliefs \( B_t \) and \( B_p \), in such a way as to ensure that \( B_t + B_p + P_t + P_p \) alone suffice for a prediction of the (same) actions. I do not deny this. For surely so powerful an account as can call upon an unlimited range of beliefs, and which is such as to impose no time constraints whatever upon the extraction of belief from a simulation process, will be potent enough for the job. My point has been simply that enrichment of \( B_t \) and \( B_p \) requires a much more complex process than that of simulation alone—as well as one that requires more computation time. So the claim is not that simulation theory permits us to explain more behavior, full stop. Rather, the claim is that simulation theory allows us to explain with less in the way of computations that must be attributed to action takers. The cunning of simulation theory is that it can do as much with

---

37The “how” of it is perhaps according to a process Robert Gordon calls “total projection” in “The simulation theory: objections and misconceptions,” *Folk Psychology*, ed. Davies and Stone, pp. 100–22.
so little. More with less is more economical. Hence a theory of prudence should not advise less with more, but more with less.

And the fact that simulation is less demanding of its practitioners is of fundamental importance. For, that which it does not insist upon is of singular biological significance. With the capability for simulation, an organism can do with the much simpler \((P) + (I)\) what the Bayesian can do only with the more sophisticated \((B) + (P)\). To put this point another way, simulation’s achievement is that it shows how predictions of behavior can be achieved once the would-be predictor grasps merely the goals of those whose behavior she aspires to predict. (This is a point that Harris is at great pains to emphasize.\(^{38}\)) So simulation can help us explain the predictions of certain individuals—namely, children and pets—whose capacities for attributing beliefs and cogitations to others is limited. Simulation is available to anyone with the ability to grasp goal-directedness, even if that same someone is unable to unlock the machinations of those minds whose goals they can quite readily apprehend. Simulation, then, is potentially a tool for unlocking the mystery of how animals coordinate behavior—so as, for example, to orchestrate the flock’s departure south for the winter. Attribution of sophisticated mutually attributed beliefs to geese is—by all accounts—something of a stretch.

This last point explains why a richer metaphysics than the Bayesians’ \((B) + (P)\) is absolutely necessary. While a theory that calls upon \((B) + (P) + (I)\) as the building blocks of action is metaphysically less moderate, and certainly less reductionist than one which calls only upon \((B) + (P)\), it nonetheless affords an important empirical advantage: the license to attribute less in the way of intellectual or cognitive sophistication to entities that manifest the undeniable capability for anticipating goal-directed behavior.

VIII. THE METAPHYSICS OF AGENCY

My proposal so far might seem afflicted with contradiction. There appears to be a confusion of freedom with its very absence. When I say that parent and toddler share a degree of freedom, I give it to toddler not parent. I do it on the grounds that the toddler suffers a certain handicap, due to tender age, and therefore is owed title to the degree of freedom, while the parent, by contrast, can exercise a power the toddler does not possess. And this seems to get things rather backwards. To answer this criticism, I shall have to discuss the metaphysics of freedom in my sense of the term—which is a sense borrowed partly from physics. It is a systems notion, thus appropriate for handling the social. It is decidedly not appropriate for handling moral responsibility, as now I will explain.

Freedom, in my sense, is not simply the capacity for putting some plan into action. If it were, there could be no reductions at all in degrees of freedom among

\(^{38}\)“Understanding pretence.”
full-fledged agents. Freedom in my sense, and as the physicists use the term, is the absence of constraint. Formally:

**Def:** A factor is a *degree of freedom in a decision system* $S$, if and only if it is numbered among those factors that shape the state of system $S$.

The notion of shaping in this definition is used primitively, and is governed by the following axiom:

**Ax:** If a factor $X$ shapes the state of a system $S$, or if it shapes another factor of system $S$, then it is false that $X$ is given shape to by any other factor.

It is clear that, according to my account, there is no such thing as a degree of freedom, in the social world, absolutely: degrees of freedom are what they are *only in relation to a system; they are what they are only in relation to the whole taken as a system*. By the same token we do not require any entity such as a common mind or a supra-individual to define the notion of degree of freedom.

Now, lack of discipline (which is the toddler’s handicap in our example, come by honestly through simple youth) gives someone a certain title to a degree of freedom. Whereas the ability to comply with an imperative, or to advance an aim, is no credential whatever, and is instead a counter-credential. For such is precisely what I have in mind by constraint on action. Now it does not disqualify something’s being my action, something for which I am to be held accountable, that I brought it about through following an imperative or seeking a particular aim. Nor does it count against its being under my control. But it does count against my title to a full degree of freedom in it. Social constraint in my sense thus *presupposes, rather than denies*, that individuals are exerting their wills in a rational fashion—however one does this. When we count the number of degrees of freedom in a social system, we are not looking at the capacities for action on the parts of social players. We are looking, instead, at the numbers of *uncontrolled or unconstrained dimensions* of decision making that can swing the social system towards one set of behaviors rather than another. Contrariwise, the degree of control or constraint in a social system is directly proportional to the level of social functioning within that system, which depends on the capacities for decision making being exercised.

Reductions in degrees of freedom occur among individuals with the capacity to see things from the point of view of others within the group. (However the exercise of such a capacity is neither necessary nor sufficient for a reduction in degrees of freedom: there are no guarantees of this sort in nature.) And now we can say, exactly, why the ability to see things from the point of view of others can help in reducing degrees of freedom: it helps with reducing the number of uncontrolled factors in the social system. When I look at things from the point of view of my toddler, I short-circuit the uncontrolled factors which arise from my own impulses, and allow his impulses to propel the social system towards the (in this case common) goal. Similarly when two political parties are vying for votes,
the ability to see things from the point of view of others, either those in one’s party or those in the other, enables one to disseminate the sort of information that will create the sorts of voting impulses one wishes created.

IX. METAPHYSICS AGAINST THE BAYESIAN: THE SCIENCE OF SOCIETY

The Bayesian agent is a consummate intellectual. His actions are grounded in reasons that have a foundation entirely in his internal corpus of propositional attitudes. Agency, on the Bayesian account, consists of principles of choice applied to this corpus of attitudes; thus the whole business can be transacted propositionally, abstracted from the mortal flesh and its waywardness. (Consequently a Bayesian model cannot admit of the phenomenon of weakness of the will, but must explain how we, mistakenly, think the alleged phenomenon exists.) It is the intellectualism, the metaphysics-free approach to explaining and prescribing behavior, rather than the principles selected for application within it, which I am suggesting is mistaken. I am saying that we cannot expect a purely intellectual system to perform as humans do, on the grounds that such systems are, by design, fundamentally independent of each other, while flesh and blood can form fundamental, nonintellectual alliances, over and above the sorts of alliance that stand-alone computers as well can enjoy. (In metaphorical language, humans can elect to “network” themselves, without wires.) I am saying that the flesh is subject to bonds which allow separate agencies to coalesce as their aims coincide. This feature of my proposal therefore serves the foundational needs of classical economics, which is concerned with economic classes and their responsibilities for disposing of economic surplus, far better than the Bayesian foundations suggested by John Roemer.39

The ability to view matters from the point of view of someone else is not a purely intellectual exercise, although it has its intellectual elements. It requires the action of psychological wet-ware. No human knows what it is like to be a bat, and this reality is not due to purely intellectual deficits on our part: we simply have not got the right equipment to exercise, well or badly. For the knowledge in question is not purely intellectual. To know what it is like to be a bat, one must be in a position to experience, among other things, the force and feel of a bat impulse. For this is a part (albeit a small one) of grasping the agency structure of bats.

The Bayesian proposal to model human others, no less than oneself, purely through the media of propositional attitudes (namely, beliefs and desires), therefore misrepresents the metaphysics of agency, simply by taking the metaphysics, as such, off stage. By doing so, the Bayesian and her ilk commend decision theory as a branch of mathematics. Against this suggestion

I propose that the metaphysics of agency, as metaphysics, is central to understanding the successes of coordination, and so is central to understanding the nature of decision. Thus I commend decision theory—which is the fundamental science of society—as a branch of natural science.

Our considerations have shown that belief-desire analysis does not capture enough of what goes on purely externally to individual agents. It does not capture enough of what does not get reflected in the internal mirror of nature. Thus, our considerations show that the Bayesian framework of looking to internal factors exclusively is at fault. For the external pressures on an agent, which advance rather than retard coordination, cannot always be reflected in that internal mirror, in the mediating propositional attitudes.

Now the Bayesian might wish to develop the following response. Surely the concept of subjective probability—of priors—can handle any external factors. The Ramsification procedure is just the sort of thing one needs to convert external factors impacting an agent into material that can be handled in the form of propositional attitudes. In fact, the Bayesian might wish to claim that priors were designed to serve just this very purpose. Thus the Bayesian might wish to hijack my argument in the service of imposing the priors requirement. This, however, does not solve the problem of explaining coordination. For it is that which comes before priors (if priors ever come on the scene) which explains coordination. That which is prior to priors, and it alone, can serve the function of reducing degrees of freedom. As we have already noticed, priors come on the scene (if they do) too late in the game to bear the explanatory burden of coordination. The “conversion” of the external into priors is precisely where we need to begin. It is the very process of coordinating.

We have found that explaining coordination requires being able to measure the internal against the external in a systematic way. The positive metaphysical proposal for doing this is now as follows. Reasons and actions are mediated by materials that are not themselves under the control of agents. Reasons are gatekeepers across the channel along which impulses pass, either checked or unchecked. I shall not say what impulses are, psychologically speaking, as I do not have to but can leave it up to others. (And anyway we are never told, on the other side, what beliefs or desires amount to; so we are doing no worse on this score than the competition, and it is not immediately clear one can do better in any case.) Suffice it to say that impulses are not, as such, the things for which agents can take responsibility. Reasons, on the other hand, are the sorts of things for which we can hold agents accountable. They function either to obstruct the passage of an impulse, or let it pass unmolested, subsequently to produce action. As much as anything else, the process of maturation is the process of learning to harness impulse, of which there is no short supply, for the sake of

40 Sometimes, though I dare say rarely, a reason can give rise to an impulse. But I shall have to leave this subject for another occasion.
aims, of which there is at the very beginnings of life considerable shortage. The balance of impulses to reasons shifts over time, if all goes well.

Thus, on my analysis, we have two classes of explainers: first, the gatekeepers (something like the reasons with which we are already acquainted) and second, the gated. Their interaction is at least partly the job of psychology to illuminate. The Bayesian neither seeks such illumination, nor believes that any is required in the normative theory of action.

X. THE COLLECTIVE

Only individuals choose and act. Collectivities, as such, neither choose nor act and analysis that proceeds as if they do is not within the accepted scientific canon.41 It is no small minority that will remind us of the unremarkable truth that groups cannot act in any way but through their members. The critic objects to the distributed self. I reply that my account does not require that collectivities be possessed of preferences at all, much less collective wills.42 I do not require that it is a self which is distributed, only that decision-making bodies can be. I seek the mechanism by which the many come to act as one, not an author of the achievement. For better or worse, morally speaking, this mechanism (which involves nesting of perspectives) gives rise to the potential for social engineering. And the idea that social engineering can reduce the number of uncontrolled social factors, which determine the dimensions of action in a social situation, is not at all problematic—at least it is not metaphysically problematic. But it can be a problem for moral theory.

---

42M. Gilbert, On Social Facts (London: Routledge, 1989) is prepared to give an account of collective desires and beliefs. I am not.