The Emergence of Norms

Russell Hardin


Stable URL:
http://links.jstor.org/sici?sici=0014-1704%28198007%2990%3A4%3C575%3AETON%3E2.0.CO%3B2-F

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*Ethics* is published by The University of Chicago Press. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ucpress.html.

*Ethics*
©1980 The University of Chicago Press

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2002 JSTOR
This is not to deny that many improvements have been won in the condition of the working class and women since the achievement of universal suffrage in the liberal state. However, like the suffrage itself, many of these achievements were won with the help of extraelectoral and parliamentary political action and they have done little to challenge the structure of class and sex inequality. (The limitations of liberal democratic voting have also been revealed in the 1970s by the fates of the elected reform governments in Chile and Australia.) DuBois's ambiguous verdict on womanhood suffrage reflects the tension between the two dimensions of universal suffrage in the liberal state. "Democratic participation" may be the key to women's emancipation, but periodic exercise of the franchise to choose national and local representatives at a time, on issues, and for candidates about which the elector has no choice is an exceedingly weak and minimal form of democratic participation compared with that in, say, the suffrage movement itself. The historic importance of womanhood suffrage lies far less in the reforms and the enlargement of individual capacities that its most ardent advocates predicted, than in the fact that it has thrown into sharp relief a basic contradiction of liberalism. The institutionalization of (more or less) formal political equality between men and women has revealed the continuing social subordination and inequality of women (the separate spheres) as a political problem. The problem is part, and a fundamental part, of the wider question of the democratization of the liberal capitalist state. Radical critics and political theorists have devoted much attention to class inequality and democracy; they have been much less willing to face up to the implications of sexual inequality for the claims of the liberal state to be democratic. The old battle between nature and convention of the contract theorists and the patriarchalists is far from concluded. If it is to be decisively joined again then the lessons of the suffrage movement must be learned and feminist scholarship extended.

The Emergence of Norms*

Russell Hardin

In *The Emergence of Norms* Edna Ullmann-Margalit sets herself the impressive task of presenting a general explanation of the generation of


*Ethics* 90 (July 1980): 575–587
© 1980 by The University of Chicago. 0014-1704/80/9004-0008$01.00
norms. Despite the great variety of contexts in which norms arise, she supposes that game theoretic representation of the essential features of these contexts can reduce the vast class of diverse norms to a far smaller number of general types. She argues that distinctive norms emerge to regulate individual behavior in three types of social situations, which she represents as prisoner’s dilemma (PD) games, coordination games, and inequality games. Her “hunch” is that these three types of cases “form the core cases in this field of inquiry” (p. viii). Along the way she supports her argument by showing how it resolves various issues which she thinks have been poorly understood by others—in particular, she proposes that her PD norms, which arise in PD situations, are precisely the field of application for Marcus Singer’s generalization argument.

Ullmann-Margalit’s method is that of Carnap’s “rational reconstruction,” which involves telling “a story of how something could happen—and, when human actions are concerned, of what is the rationale of its happening that way—not of what actually did take place” (p. 1). What she actually does is even less than this. In speaking of “the emergence” of norms, she has in mind what Hume means when he speaks of “the origin” of civil government.¹ They mean, not how norms or government came to be, but rather what interest or purpose they serve. Ullmann-Margalit is finished with her explanation of each type of norm once she has shown that a norm to regulate individual behavior in the relevant type of situation would be collectively beneficial to some group or society. Her thesis (and its limits) would be clearer if presented in a few paragraphs, whereas it gains in impressiveness by being filled out to book length, because her numerous, often clever discussions of collateral issues obscure what is central to her argument even while shoring up that argument.

PRISONER’S DILEMMA NORMS

The first of Ullmann-Margalit’s types of norms, PD norms, is the easiest to define. A particular PD norm is a norm to cooperate rather than defect in a particular situation which has the objective form of a PD game. If such a norm were followed by all or almost all of those collectively involved in such a situation, all would be better off. Hence, it is in the collective interest that there be such a norm. These last three statements constitute an explanation, in Ullmann-Margalit’s sense, of the emergence of PD norms in PD situations. She baldly asserts “that in the type of situations concerned it is likely that there will as a matter of fact be” PD norms (p. 60), and she leaves the field to take on collateral issues. Perhaps she thinks the rest is sociology. Sociological analysis aside, however, there is still much that can be argued, even by rational reconstruction. For example, one might wish to discover just how important PD norms might be. They may be potentially valuable for cleaning up our lives, for making them a bit nicer or even quieter; and

we can see how they might emerge to regulate minor interactions in our daily lives. But it is less clear how—or even that—they similarly emerge in more crucial contexts.

As is perhaps too common in ethical debates, in arguing her case Ullmann-Margalit selects examples which are generally not very important, whereas her conclusion has the ring of great importance. It is plausibly a PD norm at work when, as in one of her examples, a neighborhood public walks the long way around a corner rather than takes a shortcut that would etch a path through the grass. Is it a PD norm at work when a metropolitan public reports thefts to the authorities? One's first inclination might be to answer yes. The general public can plausibly expect to be at least slightly better off every time a theft is punished, but virtually no theft can be punished unless it is first reported. Hence, we stand to benefit collectively from a norm that requires us individually to report thefts, or, in short, a norm to punish theft. Is it a PD norm? Note that in this case the conditional for the relevant PD norm constitutes a staggeringly complex counterfactual. What if no one reported thefts? No thieves would be punished. What kind of world would it be in which no thief were ever punished, and how would it compare with our own world in which thieves are commonly punished?

Tristram Shandy's Uncle Toby reprimanded his manservant, Trim, for supposing that the King of Bohemia was upset that in all Bohemia there was no port: "How the deuce should there—Trim?" "It might, said Trim, if it had pleased God—"

—I believe not, replied my uncle Toby, after some pause—for being inland, as I said, and having Silesia and Moravia to the east; Lusatia and Upper Saxony to the north; and Bavaria to the south: Bohemia could not have been propelled to the sea, without ceasing to be Bohemia—nor could the sea, on the other hand, have come up to Bohemia, without overflowing a great part of Germany, and destroying millions of unfortunate inhabitants who could make no defence against it—Scandalous! cried Trim—Which would be speak, added my uncle Toby, mildly, such a want of compassion in him who is the father of it—that, I think, Trim—the thing could have happened no way.²

If Uncle Toby were with us now he would, in his kindly manner, dismiss any argument that turned on whether we objected to a world in which no thief were ever punished, as, alas, he would also dismiss the present observation. Yet the norm to punish theft can be construed as having emerged from a PD situation only by treading heavily on Toby's good sense. It is absurd to compare the generalized case in which all report observed thefts with that in which no one reports them. When we report thefts, is this then not following an ethical norm which produces net benefit? Surely it often is. Nevertheless, the institution of punishing theft

cannot easily be seen as a collective good in the sense that, if there were no such institution, then it would be collectively beneficial for us to share in the cost of creating one. Yet, given that we already have such an institution and that it works much of the time, an additional increment of punishing theft might well be a collective good in the analogous sense that to share in the cost of providing that increment would seem collectively beneficial.

Although we cannot extrapolate backward from the increment to the whole institution, it is plausible that, given a long history of institutional sanctions against theft, the problem of making an incremental contribution to the ongoing practice of punishment has some of the character of a prisoner's dilemma. There is some slight cost to reporting a theft, especially if one might eventually have to testify, which might seem to outweigh one's own expected gain from any punishment which followed the reporting. Moreover, given that we now have an ongoing institution to punish theft and, correlatively, given that we have built our lives and our society on the expectation that one's property will remain one's own within limits until it is voluntarily given up, almost all of us might strongly prefer to continue the punishing of theft. Hence, adherence to a norm of reporting thefts is mutually desirable. But it can hardly be argued that the norm which sustains that institution has emerged as a PD norm from a state of affairs in which there was no such institution. The fact that there is such an institution is virtual proof that there is such a norm. More generally it will often follow that, if a relevant group has had a high degree of cooperative success, then it has a norm of cooperation, and there is, therefore, reason for the norm's existence — if not, then not on both counts.

One may take on many of the most important institutions in society in a similar manner to argue against the notion that the existence versus the nonexistence of any of them constituted a PD situation which has been resolved by the emergence of the relevant PD norm. For many modest problems, however, such as the neighborhood shortcut across a beautiful lawn, one might plausibly argue for that notion, as Ullmann-Margalit does.

COORDINATION NORMS

Ullmann-Margalit's second type of norm is less readily defined because it regulates behavior in coordination games, which are less a part of the vernacular of contemporary social philosophy than are PD games. In the literature of game theory, coordination games and their "solutions" were apparently first discussed at length by Thomas Schelling,3 who proposed that successfully solving such a game is likely to depend on extragame considerations, such as the psychology, past history of mutual experiences, and so forth, of the players in the game. In a variant of one of Schelling's examples, consider two old friends who have not seen each

other in a while. One calls the other to say she will briefly be in New York Monday and suggests they meet. But their communication is incomplete, and they fail to reach any agreement on where or when to meet. Is their prospect hopeless? Schelling thinks not, because there are “prominent” choices they can make in order to coordinate their actions. He proposes that there is even a most prominent time and place for the friends’ meeting. In support of his proposal he cites numerous data collected casually at parties and in classrooms showing that many people readily spot a particular prominent solution. (Those who enjoy games should pause before referring to the footnoted time and place.)

David Lewis convincingly argues that, if such a coordination problem recurs often enough for a given population, one among the plausibly many prominent outcomes may well begin to be “selected” almost invariably, so that it becomes a “convention” always to behave as though that outcome will be selected by all parties, which is to say that one should likewise select it in order to benefit from the successful coordination. Lewis further argues that a convention in his sense is a norm, and Ullmann-Margalit asserts that a “co-ordination norm” will arise in coordination situations to add to the motivation of all concerned to follow the convention.

Despite extended discussion (pp. 85–89) of why a convention should turn into a norm, Ullmann-Margalit is not convincing on “why it is needful that regularities which solve recurrent co-ordination problems be norms” (p. 89). Her answers are essentially two: (1) A norm serves a collective purpose in a coordination problem so complex that only an articulated and understood rule is likely to be sufficient to guide behavior toward coordination. This seems to mean merely that a well-understood convention works better than one which is ill understood. (2) Having a convention be supported by normative social pressure enhances the salience of the particular actions it involves.

It is useful for a moment to break Ullmann-Margalit’s coordination norm into two components: the behavior and the norm. The behavior is simply Lewis’s convention. Whereas for the PD norm in many contexts it is unclear how the norm or the associated behavior emerges, for the coordination norm it is clear how the behavior emerges but not compelling that there be a norm that governs such behavior. There is sufficient self-interest motivation to adhere to an established convention that there need be no further norm-based motivation. Perhaps the norm could be useful to regulate the behavior of slow learners—but only if the norm could be learned more easily than the self-interest could be comprehended. Is it a norm not to drive on the left in most countries, or merely good sense? Now it may be that there is a tendency to derive “ought” from “is,” so that those who adhere to a particular dress convention or who speak a particular vernacular often pass moral judgments on any of their peers who deviate from these conventions.

4. Ibid., p. 56.
6. Ibid., pp. 97–100.
But that does not appear to be Ullmann-Margalit's point in speaking of the emergence of coordination norms, which she thinks serve some collective purpose. Yet her coordination norm can arise only after there has been successful coordination, that is, when there is no longer a problem, so that the norm resolves no problem. Morally, it is at best a very weak cousin of her PD norm, or so it seems.

In the end, the most compelling claim Ullmann-Margalit makes is a rather backhanded one: If I act in my own interest on the assumption that a recurrent pattern of behavior will continue, then, if my assumption proves false, I will tend not to fault my prediction from past behavior but to blame those who have now altered their behavior. This is an ought-from-is derivation which Ullmann-Margalit thinks is justified if the recurrent pattern of behavior was a resolution of a coordination problem (pp. 88–89). This may tell us how coordination norms arise, but we may still ask why it is “needful” that conventions be norms. The answer may lie in H. L. A. Hart’s original characterization of what Ullmann-Margalit calls norms of obligation: (1) They imply social sanctions against deviation. (2) They are thought important “because they are believed to be necessary to the maintenance of social life or some highly prized feature of it.” (3) “It is generally thought that the conduct required by these [norms] may, while benefiting others, conflict with what the person who owes the duty may wish to do.”? Ullmann-Margalit notes that coordination norms do not qualify as norms of obligation (pp. 12–13). It is clear from her account that they fail only on Hart’s third criterion. But do they? I think the answer, at least for certain important coordination norms, is that they do not fail Hart’s criterion, as can be seen in Hume’s account of the peculiar differences between benevolence and justice. To set up Hume’s distinction, first consider more fully the recurrent problem of meet me Monday in New York. The resolution of this problem is a convention, as justice is a convention in Hume’s analysis.

Note that to be alerted to look for a prominent solution in a coordination problem, as in the Monday meeting problem, greatly enhances one’s likelihood of finding one. In one of my classroom exercises, a student sensibly said she would simply have given up on her old friend for Monday after having her communication cut off. If this problem arose frequently, and if many people would simply give up, the convention which might arise from the situation might well be to conserve one’s Monday efforts for more profitable entertainments than forlornly seeking one’s old friend in New York. If almost everyone coordinated on not trying to meet, this too would yield a prominent solution to the game and hence a convention. (We might benefit from having a norm to help us coordinate sooner or on a better one of the possible conventions. But Ullman-Margalit’s coordination norm enters, if at all, only ex post facto when there is no longer a problem, to

reinforce the convention we have already established.) Which convention arises in a coordination game is historically as well as analytically determined. What could be a convention is analytically determined, but which of the many possible conventions arises is historically determined by accident, psychology, past history, etc. Whether an individual should act as though a better outcome were the convention, when in fact a poorer outcome is, depends on whether the entire group (or most of it) can easily enough be induced to switch as well. An individual requires no moral motivation to follow the poorer convention, but one might be morally moved to face loss to oneself for a while in order to initiate a shift by the whole group to a better outcome. (Again, this is not the kind of moral behavior which is covered by Ullmann-Margalit’s coordination norms.)

Now consider Hume’s distinction between benevolence and justice as virtues. Hume notes that, “as the good resulting from [an act of benevolence] is in itself complete and entire, it also excites the moral sentiment of approbation without any reflection on further consequences, and without any more enlarged views of the concurrence or imitation of the other members of society.” The good resulting from acts of justice is distinctively different; it “is not the consequence of every individual single act, but arises from the whole scheme or system concurred in by the whole or the greater part of the society.”

In a classic paper, Jonathan Harrison is at pains to understand what Hume could mean by his distinction and seems to find him inconsistent. In part Harrison is misled by his comparison of justice to the action (a convention) of two people rowing a boat. If either person does not row, the boat will not progress. Hence, the good which consists in the progress of the boat requires infallible participation by all concerned. But this conclusion is a consequence of the logic of the number two in this case, and it need not be generalized to the convention of justice in a society of considerably more than two people (or to a trireme with hundreds of rowers). Justice is not so fragile as to fall with a single miscarriage, but it can fall with frequent enough miscarriages. An occasional driver in the left lane does not destroy the convention which makes driving reasonably safe.

Harrison suspects Hume of having thought that “even if good consequences would be brought about by the general performance of a certain type of action, I have no duty to perform it, so long as I have good reason to believe that actions of that type will not, in fact, be generally performed.” But he thought this only “some of the time,” because Hume, “though he thought that we had no duty to be just in a

10. Ibid., p. 109. The example is Hume’s (Inquiry, p. 123), and it seems not to have misled him.
state of nature, thought that our duty to be benevolent was still incumbent upon us, for our duty to be benevolent, unlike our duty to be just, is in no way dependent upon the performance of benevolent actions by other people.” Against Harrison’s perplexity here, Hume was consistent if he meant to say that one is not obligated to attend to a potential convention in a case when an alternative but deficient convention is well established and is followed by most others, whereas one is obligated to be benevolent when one’s own benevolence alone is sufficient to cause good. (One may be obligated to attend to a superior but unestablished convention if one’s action is likely to bring about coordination on the superior convention.)

After distinguishing the forms of benevolence and justice, Hume immediately put them into conflict: “Cyrus, young and inexperienced, considered only the individual case before him and reflected on a limited fitness and convenience when he assigned the long coat to the tall boy, and the short coat to the other of smaller size. His governor instructed him better, while he pointed out more enlarged views and consequences, and informed his pupil of the general, inflexible rules necessary to support general peace and order in society.” In this context, Hart’s third criterion for a norm of obligation has a role. Without commitment to the norm of justice, an attitude of benevolence or considerations of short-term utility might often override the convention of justice. One would not sensibly refuse to conform to the convention of justice in general, but one might occasionally deviate for good reasons (out of benevolence or self-interest). That being so, justice in the long run might fall prey to deviant good reasons in the short run, and a norm that adds to the commitment to the convention may be, in Ullmann-Margalit’s term, “needful” to support the convention of justice. Our interests may best be served if we come to say that crime should be punished because it is wrong rather than because it is in our interests to punish it. To see the needfulness of coordination norms in this way, however, is to see such norms as very closely related to PD norms whose purpose is to overcome the intensity of short-run narrower concerns in order to benefit long-run greater interests.

NORMS OF PARTIALITY

Ullmann-Margalit’s third type of norm should not require extended discussion. Norms of this type, norms of partiality, arise in situations like that depicted in Game 1 below. In this game, the upper left and lower right cells are equilibria in the sense that, if one of these outcomes occurs, neither player can benefit from unilaterally changing strategy. Ullmann-Margalit defines such a game also to be “strategically unstable,” by which she means that one of the players could be benefited if both players changed strategies. Norms of partiality arise in such situations, she says, to motivate the players to stick with an original status quo of

inequality—such as the outcome (2,1)—even though the worse-off player might hope that a unilateral change in strategy could eventually lead the other player also to change to the first player’s absolute benefit and relative advantage. For example, slaves might adopt a norm in support of their inferior position, even though by rebelling they could improve their lot both absolutely and relatively. Again, such norms might emerge from a feebleminded ought-from-is derivation, but presumably Ullmann-Margalit wants to make a greater claim as well as to argue that the norms serve some collective interest.

<table>
<thead>
<tr>
<th>Column</th>
<th>Row</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,1</td>
<td>0,0</td>
</tr>
<tr>
<td>0,0</td>
<td>1,2</td>
</tr>
</tbody>
</table>

Game 1

In this game, the upper left and lower right cells are equilibria in the sense that, if one of these outcomes occurs, neither player can benefit from unilaterally changing strategy. Ullmann-Margalit defines such a game also to be “strategically unstable,” by which she means that one of the players could be benefited if both players changed strategies. Norms of partiality arise in such situations, she says, to motivate the players to stick with an original status quo of inequality—such as the outcome (2,1)—even though the worse-off player might hope that a unilateral change in strategy could eventually lead the other player also to change to the first player’s absolute benefit and relative advantage. For example, slaves might adopt a norm in support of their inferior position, even though by rebelling they could improve their lot both absolutely and relatively. Again, such norms might emerge from a feebleminded ought-from-is derivation, but presumably Ullmann-Margalit wants to make a greater claim as well as to argue that the norms serve some collective interest.

As with her coordination norms, one can easily see how the behavior associated with partiality norms might emerge while doubting that any further purpose would be served by the emergence of the norms. The behavior is indeed the same as that in coordination games (Game 1, which Ullmann-Margalit presents on p. 168, is a coordination game in Schelling’s and Lewis’s sense).¹³ The worse-off party is generally a group (all slaves) who, in order to concert their strategy to change the situation, must overcome an established convention. If only one or a few members of the group change, there are likely to be coercive institutions ready to punish them. Those institutions can be disabled only by massively concerted action. A norm might be marginally useful (even to slaves) to restrain the odd slave who would charge off alone to attack the plantation owner. Similarly, one can comprehend a mutiny or a riot as the rare instance of successful coordination by some group on an

¹³. See, e.g., Schelling’s example 1, p. 60.
alternative behavior that produces a superior outcome for the mutineers or rioters. How the coordination happens is also comprehensible—recall the triggering events of the American ghetto riots of the sixties.

On this account, the coordination game for the suppressed is a subgame within the larger coordination game for the whole society. Of course, it makes sense for the suppressed to coordinate in disrupting the outcome of the larger game only if they can thereby affect, to their own advantage, the strategic choice of the suppressors. But suppose that, in the event of failed coordination on rebellion, mutiny, or whatever, there is selective retribution, but that it is not perfectly selective. The net may be cast a little too wide, and innocents may be punished along with activists, as in Nazi-ruled Czechoslovakia, Videla’s Argentina, or Pinochet’s Chile. Those not inclined to risk rebellion might naturally strive, in their own self-interest, to discourage others as well. And they might sensibly pass severe moral judgment on activists, so that there might be a widespread norm among the suppressed to keep one’s peers from fighting their suppression. (In addition, one might rationalize one’s cowardly failure to rise against suppressors by imputing positive normative value to submission.) Hence, the collective interest of the suppressed might be served by Ullmann-Margalit’s norm of partiality. But in general, her inequity situations seem better understood as simply defective conventions, that is, coordination at inferior outcomes on the part of the disadvantaged group, which is to say that there is a coordination subgame within the larger game. And it is, as usual in recurrent coordination games, rational of the suppressed not to rebel unless there is good reason to think they can coordinate, so that a norm of partiality will have little additional effect on choices.

THE GENERALIZATION ARGUMENT

In discussing her PD norms, Ullmann-Margalit argues that the generalization argument (GA) of Singer and others is not as general as they claim. Singer asserts that his weak GA applies to all and only those actions which have moral overtones. Ullmann-Margalit asserts that any valid GA applies only to PD situations (pp. 53–58), which on her account are only a part of the moral realm. Neither presents more than a partial argument. Ullmann-Margalit shows only that a strong version of the GA does validly apply to PD situations, as seems self-evident. Her argument against the validity of broader applications is a collection of situations which a straw man GA obviously fumbles. Oddly, however, Ullmann-Margalit has the material with which to do in the GA in its strong version on ground which it should govern. The strong GA should apply to coordination situations in which a defective convention has been established, but the principle of individual action derived from the GA for those situations would commonly be wrong.

Singer’s GA is: If the consequences of no one’s doing X would be undesirable, while the consequences of everyone’s doing X would be desirable, then everyone ought to do X. An alternative, which is stronger in that it applies far more generally, is a comparative version (which lends itself readily to appropriation by utilitarians): If the consequences of everyone’s doing a certain sort of thing would be worse than those of some alternative, then it would be wrong for anyone to do such a thing.

Suppose we are involved in a recurrent coordination game in which we are collectively and individually better off if every one of us does X, or everyone does Y, etc., than if we all (or nearly all) fail to coordinate on any one of many possible actions. Also suppose we have established a convention always to do Y, although a convention always to do X would be clearly superior for everyone. The strong (comparative) GA, but not Singer’s weak GA, recommends individual action. Consider what it seems to recommend in the case of a particular deficient convention: the contemporary typewriter keyboard. The location of letters on the standard keyboard is clearly not the best possible arrangement. But the standard arrangement is a convention which seems incorrigible. The strong GA seems to suggest, therefore, that, say, Olivetti should arrange the keys differently, because if all typewriters were more sensibly arranged, all typists would be better off. The suggestion is probably wrong in this case because the cost (at least to the present generation of typists) of changing the convention might outweigh the benefits to be gained. But even if that were not so, the suggestion would still seem wrong—because Olivetti’s action likely would not be expected to lead to a changed convention, so that the net result of the action would be detrimental rather than beneficial. Hence, one cannot derive a principle of individual action from the strong GA in this context. Similarly, the drivers of Sweden may long ago have wished their driving-on-the-left convention were the driving-on-the-right convention which prevails in most of the world. But it would have been wrong and foolish of an individual Swedish driver to switch to driving on the right. Any formulation of the GA which recommends such a switch is self-evidently silly.

To straighten out such kinks, Neil Dorman proposes his modified GA (MGA): If the consequences of not having a rule against Y would be undesirable, then there ought to be a rule against Y. Ullmann-Margalit says “that PD norms in fact constitute an explication of these rules of Dorman’s” (p. 59). She is not strictly right, because Dorman’s rules may apply as well to coordination as to PD situations. But such rules are less interesting than either Singer’s GA or Ullmann-Margalit’s norms in that

15. Ibid., pp. 72–73.
they are less applicable to individual choice than to lawmakers' choices. For example, in the Swedish driving convention, the MGA could recommend a changed convention only to someone whose action would actually bring about a simultaneous switch by every driver in Sweden, as Swedish government action eventually did. Cyrus might be instructed by Dorman's MGA; his subjects could not be.

Ullmann-Margalit says that the comparative judgment is essential for a GA. Singer is at great pains to argue the contrary, and he insists on his weak version of the GA that obligates us to action only to prevent harm and not to cause good. The distinction is often blurred, so that his GA is ragged at the edges, but we are not always incompetent to make the distinction. (In this respect, Singer precedes recent trends in ethical theory toward changing the terms of discussion from what is right to what are one's rights. This trend is perhaps a belated reaction to the demanding ethics of duty, a positive ethics, fully hortatory and minatory, of the utilitarians Sidgwick and Moore and of their interwar critics, such as Ross. Charles Fried and others argue for a restricted ethics of rights, a negative ethics, almost exclusively minatory. Singer includes some duties but protects us with rights against other claims to duty.) A utilitarian adherent of the GA would agree with Ullmann-Margalit that it applies to all PD situations. For Singer, it would presumably apply only to those recurrent prisoner's dilemmas in which one has been cooperating (so that one's defection would harm others), not to those in which one has never cooperated (unless perhaps one has got others to cooperate through some sort of agreement). For a utilitarian it would apply as well to a failed prisoner's dilemma if one's own cooperative action would yield net benefit, but not to defective coordinations unless one's own action would plausibly effect enough reaction on the part of others to produce a superior convention. Hence, to equate the GA with the motivation of Ullmann-Margalit's PD norms is nearly to equate it with a utilitarian's GA. Ullmann-Margalit's intuition here seems compelling: Neither the weak nor the strong GA yields as credible a moral stance as the utilitarian's GA. These three positions, the fully hortatory strong comparative GA, the minatory weak GA, and the utilitarian's GA, do not exhaust the realm of plausible moral positions in such contexts. Hart and others have argued for cooperating in prisoner's dilemmas when there is an institution or practice of cooperating. For Hart the norm of "obligation is due to the co-operating members of the society as such and not because they are human beings on whom it would be wrong to inflict suffering." Insofar as Ullmann-Margalit's account of norms is positive or predictive, her PD norms when they arise may well fit Hart's stance rather than that of any of the GAs.

In sum, The Emergence of Norms is often concerned with emergence and generally with norms, but it seldom successfully argues for the

emergence of norms. Nevertheless, Ullmann-Margalit's mode of inquiry—looking for underlying game theoretic structures of social interactions in order to discover patterns of norms which might govern them—is powerful, original, and likely to be influential. Indeed, that mode recommends against one or her more audacious and important conjectures: that her three classes might encompass the whole realm of social norms. Given the rich diversity of possible games, it would be odd if only prisoner's dilemmas, coordination, and the inequality type of coordination games had moral overtones.

Reviving the Virtues*

Arthur Flemming

A bibliography of modern philosophical works on the virtues would be very short. It is as though philosophers had taken Bentham too seriously when, with his customary candor, he judged the virtues poor conceptual soldiers: "There is no marshalling them; they are susceptible of no arrangement; they are a disorderly body, whose members are frequently in hostility with one another. . . . Most of them are characterized by that vagueness which is a convenient instrument for the poetical but dangerous or useless to the practical moralist."1 Prevailing opinion has it that the virtues can only be organized by a principle of right, as various dispositions to do the right thing. A theory of the virtues would then be at most an epiphenomenon of the theory of right. Perhaps this explains why, with few exceptions,2 recent work on the virtues has rarely gone beyond commentary on historical sources, a sure sign that a subject has been safely buried. In separate volumes with an unpromising title in


Ethics 90 (July 1980): 587–595
© 1980 by The University of Chicago. 0014-1704/80/9004-0009$01.00