Society and the Highest Stages of Moral Development

Gerhard Sonnert Michael L. Commons
Department of Physics Department of Psychiatry
Harvard University Harvard Medical School

Sonnert, G., & Commons, M. L. (1994). Society and the highest stages of moral development. *Politics and the Individual*, *4*(1), 31-55.

In this paper, a discussion of Kohlberg's stage theory of moral development results in a re-definition of Moral Stage 6, Kohlberg's highest stage. Our task here is to construct Moral Stage 6 in a non-arbitrary way that satisfies General Stage Model (Commons & Richards, 1984a) criteria for what constitutes a stage. Moral Stage 6 is found to manifest itself only inter-individually. At this stage, morality is no longer a property of individuals, as it is at earlier stages, but a property of the social enterprise of discourse. This leads to a novel politization of morality and, conversely, to a moralization of politics.

Challenging the prevalent relativism in the study of political ideology, a number of researchers have recently applied theories of moral development in this field (Emler, Renwick & Malone, 1983; Habermas, 1979; Rosenberg, 1988a, 1988b; Rosenberg, Ward & Chilton, 1988; Wagner, 1986, 1990; Weinreich-Haste, 1986). Whereas the relativist approach posits that a non-arbitrary, acultural evaluation of political ideologies is impossible, a theory of moral development provides the means for just such an evaluation. It can discover structural commonalities among seemingly unique and diverse political ideologies; it can order political reasoning in a non-relativist sequence; and it can point to possible avenues for the development of political thinking.

Modern societies have partially differentiated the political from the moral domain, developing areas of superficially moral-free political pragmatism. This is reflected in the popular distinction between "political" and "moral" questions: pragmatic compromising is considered acceptable in the former but unacceptable in the latter. A key mechanism for creating these relatively moral-free areas of politics is framing the issues in terms of continuous variables, first of all money. Continuous variables invite compromise and take the moral sting out of the issues. The **moral** difference between alternative tax-hike proposals of 2% and 4% is not obvious, and a compromise is generally welcomed. However, the absence of conspicuous morals in the political sphere of power-play and haggling is possible only on the basis of a widely-shared and unquestioned background belief that such politicking is justified within the given parameters.

Politics is not identical with morality, but political and moral reasoning are connected insofar as they both deal with issues of justice. Political legitimacy is the focal point where politics and morality converge. Individuals and social groups tend to grant legitimacy to, or withdraw it from, political institutions, programs, and decisions on the basis of whether or not they view them as just. Thus, political reasoning and behavior builds on moral judgments. As part of the wider effort of examining political reasoning and action from a developmental perspective, we introduce in this paper new descriptions of the highest stages of moral development and discuss their repercussions in the political domain.

Kohlberg has produced the most influential theory of moral development both within individuals (Kohlberg, 1969, 1984) and within society (Power, Higgins & Kohlberg, 1989). As a cognitive developmentalist, Kohlberg considered mathematical and logical intellectual operations, such as those used to discern amount, balance, and causality, to underlie moral operations, i.e., moral reasoning. The hierarchy of moral stages was thus grounded in cognitive functioning. Kohlberg showed that in the course of moral development, reasoning about justice increases in complexity and becomes more inclusive.

Cross-cultural data from over 30 societies indicate that people's development follows an invariant sequence, but that the end points of development vary (Colby & Kohlberg, 1987; Snarey, 1985). Movement through the developmental stages is unidirectional. Barring mental illness or loss of mental capacity, people's reasoning moves only up through the stages, never down, within test accuracy limits.

In what follows, we give our short interpretation of Kohlberg's stage theory. Kohlberg posits three periods of development in the moral domain: the preconventional, the conventional and the postconventional. Each of these three periods is subdivided into two stages so that Kohlberg's model comprises six stages of moral development (see Table 1).

Table 1
Correspondence of Stage Models

General Stage Model		Moral Stage (revised Kohlberg stages)	Approximate age
Aconventional			
0a	sensory & motor actions	-1	0-0.5
0b	circular sensory-motor actions	-1/0	0.5-1
1a	sensory-motor actions	0	1-2
1b	nominal actions	0/1	2-3
Preconventional			
2a	sentential actions	1	3-4
2b	preoperational actions	½	4-6
3a	primary operations	2	6-8
3b	concrete operations	2/3	8-10
Conventional			
4a	abstract operations	3	10-12
4b	formal operations	3/4	12-17+
5a	systematic operations	4	18+
Postconventional			
5b	metasystematic	5	20+
6a	paradigmatic	6	?
6b	cross-paradigmatic	(7)*	?

* No corresponding Moral Stage

The preconventional period (Moral Stages 1 and 2) begins in early childhood and extends through elementary school. At Moral Stages 1 and 2, people justify actions in terms of avoiding punishment and obtaining rewards. At these stages, people are particular persons ("myself" or "my parents") who do particular things ("If I don't keep my promises, then my parents won't trust me anymore"). Adults generally consider the moral reasoning of this period inadequate.

The conventional period (Moral Stages 3 and 4) begins at the onset of post-elementary school education and extends across the life-span of all but a small portion of the population. This period generates the conventional norms of adulthood. Reasoning at each stage of this period contains enough logic that it can find its most elaborate expression in some current adult philosophy.

At Moral Stage 3, the *Group stage*, action is justified in terms of the reputation and characterization of the groups or individuals that are involved. Groups and people can, for instance, be good or bad, nice or nasty. Action is often judged on the basis of groups' or individuals' underlying sentiments or motives. Role and person may be confused.

At Moral Stage 3/4, the *Bureaucratic stage*, the reasons given for labeling an action as fair and good are logical and abstract. Bureaucratic norms, laws, rules, and regulations guide behavior and are seen as "given"; they cannot be changed to fit particular situations or individual needs. Role and person are no longer confused as they were at the previous stage.

At Moral Stage 4, the *Institutional stage*, the yardstick for evaluating the morality of an action is the preservation (or destruction) of a system--or a society. Norms, laws, rules and regulations now form a logically coherent system. Societal law is seen as a meaningful way of regulating rights or duties of individuals or groups. People at this stage reason in terms of how an action would affect their individual role and status within the system, or how it would affect the system's capability to function. Hence, there are tensions between societal and personal rights, and between societal and personal duties. Individuals experience such tensions, for instance, in the conflict between independence from and dependence on both others and the system. Most persons at this stage accept their subordination to the societal system. "What would happen to society if everyone...?" is a question characteristic of this stage. A minority of individuals rejects the system outright and totally and emphasizes individual independence.

The postconventional period (Moral Stages 5 and 6) begins sometime after adolescence; however, fully postconventional thinking and action appear only after early adulthood (Colby & Kohlberg, 1987). Some contemporary philosophies use postconventional arguments. In fact, Armon (1984b, 1989) has reinterpreted philosophical debates in terms of conventional versus postconventional arguments. In any known society, only a small portion of members achieves postconventional stages of reasoning.

At Moral Stage 5, the *Universal stage*, universal abstract principles form the basis of moral and political action. They emerge from notions of universal human rights and dignity. These principles have been articulated by members of a number of modern societies (Reiser et al., 1987). They have found their fullest expression in the works of philosophic, political, and religious thinkers. For example, a principle in Rawls' (1971) theory of justice posits that actions should not worsen the situation of the least advantaged.

Moral Stage 5 principles are universal and general in their application, irrespective of the particular persons affected. At this stage, society is seen first as a creation of individuals and second as the environment in which people develop. Emphasis on the interdependence between individuals and society reconciles the dependence stances and the corresponding independence stances of the previous stage. Moral Stage 5 principles not only serve the interest of individuals but also that of society. From a developmental-stage perspective, the principles coordinate the rights and the duties of the individual with those of society. The methods of decision-making are due processes ranging from lotteries to voting, with a preference for joint decision-making. Unconventional decisions may be sanctioned as long as they appear reasonable in the light of higher principles.

Kohlberg (1969, 1981; Kohlberg, Levine & Hewer 1983) also posited a Moral Stage 6. This proved to be the most problematic stage in his model for conceptual and empirical reasons. Kohlberg had problems in articulating the logical and conceptual difference between Moral Stage 6 and Moral Stage 5 (Diessner & Commons, in preparation). And empirically, he found it difficult to identify subjects reasoning at Moral Stage 6 (Kohlberg, Levine & Hewer, 1983, p. 60).

Plan of the Argument

This paper advances a new proposal for Moral Stage 6 within a modified Kohlbergian framework. By addressing the conceptual problems with Moral Stage 6, we hope that the empirical problems will be solved as well. We assert that a meaningful Moral Stage 6 cannot be defined within individual moral development; it has to be understood as the property of a social enterprise--of discourse. Thus, Kohlberg was looking for Moral Stage 6 in the wrong place and, as a consequence, came up empty-handed.

We will begin by reviewing the evolution of Kohlberg's own theory of developmental stages. Next, the General Stage Model (Commons & Richards, 1984a, 1984b; Commons & Rodriguez, 1990) will serve as the basis for discussing certain shortcomings of Moral Stage 5 reasoning and for critiquing Kohlberg's notion of Moral Stage 6 (Kohlberg, 1986; Kohlberg, Boyd & Levine, 1990). Then, we will develop a discursive notion of Moral Stage 6 and explore its individual, methodological, and societal implications.

History of Kohlberg's Stage Theory

When discussing Kohlberg's theory of moral development, one should not overlook that the theory itself, and the notion of Moral Stage 6 in particular, developed over a number of years. Colby and Kohlberg (1987) distinguish a number of historical periods of Kohlberg's theory. The "old period," in our terminology, was characterized by global scoring of the content of subjects' responses (Kohlberg, 1969; Kohlberg & Kramer, 1969). In this period, Moral Stage 6

was defined as "Conscience or principle orientation. Orientation not only to actually ordained social rules but to principles of choice involving appeal to logical universality and consistency. Orientation to conscience as a directing agent and to mutual respect and trust." (Kohlberg, 1967, p. 171; 1984, p. 42).

In a transitional period (Kohlberg, 1973), Moral Stage 4/5 and other half-stages were added to the model to fix problems emerging with global scoring, but the difficulties in content-scoring proved intractable. The introduction of structural scoring (Kohlberg, 1976), which solved most of those problems, marked the end of the transitional and the beginning of the "middle period." In this middle period, details of structural scoring were developed, and that solution finally appeared in Colby and Kohlberg's (1987) "Scoring Manual." Stage definitions were systematized (Kohlberg, 1976; Kohlberg, Levine & Hewer, 1983). They were now based largely on Selman's social perspective-taking considerations (Selman, 1976, 1980; Selman & Byrne, 1974). Kohlberg had become increasingly uncertain about Moral Stage 6 and, over a period of time, transferred material from Moral Stage 6 to Moral Stage 5.

In the last period, or "new period," Kohlberg's (1986, p. 486; Kohlberg, Boyd & Levine, 1990) thinking was greatly influenced by Habermas and brought a "new" Moral Stage 6 that we discuss here. This new stage put increased emphasis on discourse and addressed the shortcomings of a more complete Moral Stage 5 (now augmented by material from the old Moral Stage 6).

Discerning these developments in Kohlberg's theory is difficult because the formal definitions of Moral Stage 6 found in tables in his publications did not really track the substantive changes in his theory but remained remarkably constant. The above definition of Moral Stage 6 from the old period (Kohlberg, 1967, p. 171) re-appeared unchanged in the middle period (Kohlberg, 1984, p. 42). Likewise, the following definition of Moral Stage 6 can be documented from the middle (1976) through the new period (1987):

> *Following self-chosen ethical principles.*
> Particular laws or social agreements are usually valid because they rest on such principles. When laws violate these principles one acts in accordance with the principle. Principles are universal principles of justice: equality of human rights and respect for the dignity of human beings as individual persons (Kohlberg, 1976, p. 35; 1984, p. 176; 1986, p. 489; Colby & Kohlberg, 1987, p. 19).

Nonetheless, in the new period, much of the material described as Moral Stage 6 in the above-quoted passage appeared in Kohlberg's Moral Stage 5. For example, Kohlberg asserted that, at Moral Stage 5, a "rational moral agent [is] aware of universalizable values and rights" (1986, p. 495).

General Stage Model

To counter the possible objection of arbitrariness in the definition of moral stages, we use the stage criteria of the General Stage Model (GSM) (Commons & Richards, 1984a, b). In this model, the hierarchical complexity of tasks was the basis for construing the notion of stage. In turn, the hierarchical complexity of tasks was grounded in mathematical models (Coombs, Dawes & Tversky, 1970) and information science (Lindsay & Norman, 1977).

Commons and Richards (1984a, b) emphasized that developmental theory addresses two conceptually different issues: (1) the hierarchical complexity of tasks to be solved; and (2) the psychology, sociology and anthropology of how task performance develops. The description of stages in the General Stage Model belongs to issue 1 because it is strictly analytical, delineating a discrete hierarchy of task complexity. The General Stage Model consists of a set of axioms that define a stage sequence, and it describes the necessary analytical properties of stages. It does not posit any detailed empirical forms of stages or any empirical processes that cause stage change. The present paper focuses on the analytic properties that Moral Stage 6 should have if it exists. It does not claim to be an empirical description of an existing phenomenon.

At this point, we will briefly review the main elements of the General Stage Model. At the heart of the model is the notion of hierarchical (vertical) complexity--as opposed to non-hierarchical (horizontal) complexity. *Hierarchial complexity* is the order of logic required to complete a task (Commons & Richards, 1984a; Commons & Rodriguez, 1990). *Non-hierarchical complexity* is the number of bits of information needed to complete a task, as described in information theory. By definition, for every yes-no question embedded in a task, there is an answer containing one bit of information. Here, we are concerned with hierarchical complexity only. The successful completion of a task requires an action of some order of hierarchical complexity. At each order of hierarchical complexity, actions, including reasoning, are defined in terms of the actions of the previous order. Actions at a higher order of hierarchical complexity organize and transform lower-order actions. We say that higher-order actions coordinate the actions of the next lower

order. This organization of lower-order actions is new and unique and cannot be accomplished by those lower-order actions alone.

The General Stage Model not only meets the four Hard Stage Conditions of Kohlberg and Armon (1984) but also a fifth condition, namely that the "Logic of each stage is explicit" (Trudeau & Commons, 1991). Therefore, the sequence of stages can be explained analytically, and a logical analysis of tasks can be performed to determine their order of hierarchical complexity (Commons & Grotzer, 1990; Commons & Richards, 1984a, b; Commons & Rodriguez, 1990). For example, multiplying 3 x (9 + 2) requires a distributive action at the concrete order of hierarchical complexity. The distributive action is as follows:

3 x (9 + 2) = (3 x 9) + (3 x 2) = 27 + 6 = 33.

This action coordinates the primary-order actions of adding and multiplying by uniquely organizing their sequence. The distributive action is therefore one order more complex than the acts of adding and multiplying alone. This action is required in both long multiplication and long division. As an equally complex alternative to the distributive action, one could follow a priority rule--do operations in parentheses first. In either case, doing both addition and multiplication in a coordinated manner is an advance over just doing addition and multiplication.

What follows is a capsule description of the analytic properties of the stages in the General Stage Model in descending order. When moral stages are viewed from a General Stage Model perspective, it is useful to translate the Moral Stage numbers into the General Stage Model (GSM) stage numbers. Table 1 shows their correspondence.
GSM Stage 6b (Cross-paradigmatic, Moral Stage 7) requires actions that coordinate and integrate GSM Stage 6a (Paradigmatic, Moral Stage 6) fields. A Stage 6a field consists of coordinated sets of GSM Stage 5b (Metasystematic, Moral Stage 5) metasystems. A Stage 5b metasystem (or "supersystem") coordinates GSM Stage 5a (Systematic, Moral Stage 4) systems. A Stage 5a system is a coordination of GSM Stage 4b (Formal, Moral Stage 3/4) formal operational relationships--formal operations mark the top stage of Piaget's (Inhelder & Piaget, 1958) system of stages as he designated them. A Stage 4b formal operational relation coordinates GSM Stage 4a (Abstract, Moral Stage 3) variables. A Stage 4a variable coordinates GSM Stage 3b (Concrete, Moral Stage 2/3) specific instances. One such variable is an abstract subset of "rights" formed from the specific rights found at the concrete stage (GSM Stage 3b). Another variable is an abstract subset of "duties."

The compatibility of the General Stage Model with, and its usefulness for, more empirical theories of development have been demonstrated. For instance, Commons and Grotzer (1990) showed that the General Stage Model can adequately account for Kohlberg's stages of moral development. Armon (1984a) has used both Kohlberg's theory and the General Stage Model to create a general theory of the development of ethical reasoning. Sonnert (in press) considers the empirical ramifications and limitations of Moral Stage 6. Commons and his collaborators are currently examining possible Stage 6 social perspective-taking in actual institutions.

In this paper, we will apply the General Stage Model to the discussion of Moral Stage 5 and Moral Stage 6. We will state the order of hierarchical complexity at Moral Stage 5 by noting what lower-stage actions Moral Stage 5 actions coordinate. Then we will show which coordinations cannot be performed at that stage. These coordinations are characteristic of the next higher stage, Moral Stage 6. We thus focus on how to coordinate metasystems of the metasystematic stage (Moral Stage 5), but the general method of applying the General Stage Model to the moral domain is the same for every stage.

Accomplishments and Shortcomings of Moral Stage 5

Recall that a Moral Stage 4 (GSM Systematic Stage, 5a) system coordinates formal operational relations (GSM Formal Stage, 4b) among variables (GSM Abstract Stage, 4a), i.e., abstract propositions about rights or about duties. Moral Stage 5 (GSM Stage 5b) metasystems coordinate Moral Stage 4 systems of rights and systems of duties by applying moral principles. Moral Stage 5 reasoning claims universal validity of its moral principles. This universality claim is not derived from some external authority or privileged perspective, as at earlier stages, but from the nature of the metasystematic Moral Stage 5 coordination itself.

This Moral Stage 5 universality claim fails because of one fundamental problem: there are many possible Moral Stage 5 (metasystematic) coordinations. Yet within a Moral Stage 5 framework, such Moral Stage 5 coordinations cannot be in turn coordinated in a way that would sustain the Moral Stage 5 universality claim. Such a claim must be based either on the completeness of the metasystematic coordination or on a non-arbitrary selection rule, but, as we will

argue, any such coordination would be either inconsistent if complete or arbitrary if incomplete. Thus, Moral Stage 6 is necessitated by the logical failure of Moral Stage 5.

Systems are complete (Gödel, 1977) when there are no valid statements outside of them. In order to be complete, a coordination of Moral Stage 5 (GSM Stage 5b) metasystematic coordinations would have to contain **all** Moral Stage 5 (GSM Stage 5b) metasystematic coordinations. But if one attempted to combine all metasystematic coordinations, the result would be internally inconsistent. For example, if two people had slightly different moral supersystems, a melding of the two in their entirety would be inconsistent. At some point, adding a statement from one person's supersystem to the other's supersystem would make the enlarged supersystem inconsistent. From elementary logic (Luchins & Luchins, 1965), we know that a contradiction in an inconsistent system would allow for any proposition to be proved as true or false.

In the moral domain, the price for avoiding inconsistency is incompleteness. In that case, a rule is needed for choosing a particular incomplete supersystem over others. Without such a rule, the selection of a consistent but incomplete supersystem would be undetermined and arbitrary. A possible rule for selecting an incomplete supersystem that is consistent would, for instance, be taking the common core of the supersystems in question. Alternatively, one might take the Moral Stage 5 supersystem preferred by most people. Another plausible rule would be to take the supersystem that embodies the most cherished principles, even though the supersystem as a whole might not be the most popular. All of these rules, however, achieve only a pseudo-coordination of Moral Stage 5 supersystems. Their very task--selecting an incomplete but consistent coordination of Moral Stage 5 supersystems--immediately re-appears in a new guise--selecting an incomplete but consistent coordination of selection rules--and so on ad infinitum. Such selection rules do not really coordinate different sets of moral stage 5 principles, but they are a particular kind of "universal principles" themselves. In short, from a metasystematic perspective (GSM 5b, Moral Stage 5) one would not know which incomplete but consistent set of metasystematic coordinations to choose. Hence any such incomplete set would be arbitrary.

The failure of selection rules implies that individual metasystematic coordinators are unable to perform a satisfactory coordination of Moral Stage 5 supersystems in monological isolation. The problem of arbitrariness can only be solved in actual cooperation among the individual coordinators involved. Thus, it becomes apparent that Moral Stage 5 underrates the role of concrete individuals. The typical strategy of Moral Stage 5 coordinations is to abstract from individuals' particular biases in order to attain a universalist moral solution, such as in Rawls' (1971) "veil of ignorance." The content of Moral Stage 5 universalist principles is of course very humanitarian, but they are adapted over the head of the actual people involved. The unique features of the situation and of the potentially different metasystematic coordinations of the other people are lost. Thus, one of the shortcomings of Moral Stage 5 is its reductionism--disregarding the particularity and uniqueness of human beings (Puka, 1987, 1990). Moral Stage 5 resorts to parentalism to support the rights and duties of the less advantaged through some imaginary mechanism of social contract.

Concomitant with the abstraction of the individual, Moral Stage 5 has an abstract concept of society and does not fully understand its structural dynamics and limitations, such as illustrated by Arrow's (1951) welfare theorem, Plott's (1967, 1971; Plott & Levine, 1978) voting dilemmas, and Hardin's (1977) tragedy of the commons. As regards the voting dilemmas, for example, a simple majority rule system for three or more candidates may be hampered by nontransitive preference and also produce a candidate the majority do not prefer. Nontransitivity occurs when Candidate A is preferred over Candidate B, Candidate B is preferred over Candidate C, but Candidate C is preferred over Candidate A. In this situation, one cannot decide who the most preferred candidate is. Or, the voters may prefer the Conservative Candidate A (33% of the vote) over the Conservative Candidate B (32% of the vote). In a three-way election, the Liberal Candidate C (with 35% of the vote) wins, although 65% of the voters preferred a conservative candidate.

Investigation of such dilemmas in voting systems has produced efforts of creating a voting supersystem that mends the flaws of the simple voting systems. However, there are many different proposed supersystems. The complete combination of two Moral Stage 5 (GSM Stage 5b) voting supersystems that each include different sets of the desirable properties found in Moral Stage 4 (GSM Stage 5a) voting systems will be inconsistent. And no rational political scientist can non-arbitrarily determine that a particular (internally consistent) voting supersystem is best. Thus, the voting dilemma, as well as Arrow's and Hardin's dilemmas, illustrates that "monologic" attempts at solving it are either complete but inconsistent or consistent but incomplete and arbitrarily selective.

Critique of Kohlberg's Moral Stage 6

Before discussing the shortcomings of Kohlberg's notion of Moral Stage 6, we introduce a model of transition steps between stages. This model serves two immediate purposes. It sheds light on Kohlberg's problems with his Moral Stage 6 by locating it precisely in the middle of a transition, and it creates empirical hypotheses about structural "pathologies" of discourse that will be noted later on. Additionally, it eliminates some confusion about Kohlberg's Moral Stage 4/5. Our model makes apparent that Kohlberg's Moral Stage 4/5 is not really a stage but only the first two transitional steps toward Moral Stage 5.

The transition-step model uses signal detection theory (Commons & Richards, 1984b) and Kuhn and Brannock's (1977) work to modify Piaget's four-step, probabilistic model (Flavell, 1963) of equilibration. The model describes the steps through which a "new equilibrium," i.e., a new stage, comes about (Commons & Hallinan, 1989). Deconstruction of the previous stage actions takes place in Steps 0, 1 and 2. The old stage actions are rejected, varied and temporized, all of which increases the number of alternative actions. Construction of the new stage action takes place in Steps 3 and 4. The model extrapolates Piaget's four-step model by extending it to a new domain (moral development) and to stages beyond formal operations.

The end point of the four steps is a temporary equilibrium point--a new stage. The intermediate steps, however, are not stages in the sense of the General Stage Model; they are not analytical constructs possessing the necessary properties of stages. Rather, the steps belong to the realm of empirical science and describe stage acquisition in an empirically testable manner. We give a detailed example of the transition steps beyond Moral Stage 4 in the Appendix.

Table 2
4-Step Model of Equilibration in the Moral Domain
General pattern

Step 0	A = A* with B*	Temporary equilibrium point (thesis) fails
Step 1.	B	Negation or complementation (antithesis)
Step 2.	A or B	Relativism (alternation of thesis and antithesis)
Step 3.	A and B	Smash or Fusion (attempts at synthesis).
Substep 1		Elements from A and B are included in a non-systematic, non-coordinated manner. Incorporates various subsets of all the possible elements
Substep 2		Incorporates subsets that produce hits at stage n. They also produce false alarms.
Substep 3		Incorporates subsets that produce correct rejections at stage n. They also produce misses.
Step 4	A' = A with B	New temporary equilibrium (synthesis and new thesis)

We now delineate the transition steps between the stages (see Table 2). At Step 1 of the transition-step model, subjects negate the coordination held at the previous stage and espouse an alternative coordination. We call the previously held coordination the "thesis" (A) and the newly adopted coordination the "antithesis" (B) to highlight the dialectical flavor of our transition model.

At Step 2 of a transition, people alternate between a thesis and its antithesis. They are aware of various possibilities and use relativistic and contextualizing arguments as to which to adapt, but do not integrate these possibilities.

Step 3 is probably the most interesting steps from a psychological perspective. Construction now replaces the deconstruction of Steps 1 and 2. At Substep 1 of Step 3, subjects recognize that the alternatives have to be fit together. Yet because they have not previously coordinated such systems, they tend, figuratively, to smash the thesis and antithesis together, picking different subsets of the parts to make the new whole function. This is a pattern-recognition step because a new template that represents a sufficient coordination has to be created. The new coordinations so generated begin to be successful at Substep 2. Overgeneralizations are reduced at Substep 3 at the expense of successful coordinations.

Finally, at Step 4, the coordination, or "synthesis," is completed and a temporary equilibration is established.

The General Stage Model expects a real Moral Stage 6 in Kohlberg's justice reasoning model to be paradigmatic (GSM Stage 6a)--the third postformal stage. In contrast, Kohlberg's (1986) description of Moral Stage 6 from the new period can be easily identified in General Stage Model terms as a full coordination of systems of rights and systems of duties. "These [universal Moral Stage 6] principles may be expressed ... in terms of the language of human rights (and reciprocal duties)...." (p. 497). Such a coordination of reciprocal systems of rights and duties into a metasystem requires only metasystematic thought (GSM Stage 5b), because systems of rights and systems of duties belong to Moral Stage 4 (GSM Systematic Stage 5a). Thus, Kohlberg's above description of Moral Stage 6 belongs to the GSM metasystematic Stage 5b (Moral Stage 5) according to the General Stage Model (Diessner & Commons, in preparation).

Even from the intrinsic perspective of Kohlberg's own theory, his (old and new) notions of Moral Stage 6 should really have been Moral Stage 5. Diessner and Commons (in preparation) show that most of what Kohlberg called Moral Stage 6 can be understood in terms of his own Moral Stage 5. They furthermore point to inconsistencies in Kohlberg's treatment of his new Moral Stage 6. In our opinion, these inconsistencies result from a melange of Moral Stage 5 (GSM Stage 5b) and Moral Stage 6 (GSM Stage 6a) elements. Kohlberg's new Moral Stage 6 is not a full-blown Stage 6 in General Stage Model terms, but hovers around the transitional steps between Moral Stage 5 and Moral Stage 6.

Kohlberg had increasingly become aware of the importance of actual discourse and added a discursive element to his definition of the new Moral Stage 6 (Kohlberg, 1986, pp. 497, 506). But he did not take the radical step of severing Moral Stage 6 moral reasoning from the individual entirely and attributing it to the discourse itself, as Habermas (1979, 1990) did. Instead, he proposed three mechanisms of individual moral reasoning that could take the place of actual discourse at Moral Stage 6, i.e., moral "musical chairs," blindness as to one's own position in the hierarchy, and hypothetical discourse. But these three mechanisms belong to Moral Stage 5 (GSM Stage 5b); they are no different from Moral Stage 5 principles. Kohlberg could not sustain his glimpse of the discursive nature of Moral Stage 6 without contradiction, because the traditional individualistic focus of his theory prevented the formulation of a consistent Moral Stage 6. Our model of transitional steps locates Kohlberg's new Moral Stage 6 at the third transitional step (GSM 6a-3, Substep 1) because it acknowledges the need for coordinating different Moral Stage 5 perspectives in discourse, but is unable to achieve a sensible coordination and reverts to Moral Stage 5 solutions.

Definition of Moral Stage 6

Within the General Stage Model, we define Moral Stage 6 using the hierarchical coordination criterion. Moral Stage 6 (GSM Stage 6a) coordinates Moral Stage 5 (GSM Stage 5b) supersystems. The crucial point is that a Moral Stage 6 paradigmatic coordination cannot be monologically determined, because it depends on the agreement of the participants of the coordination. It changes and evolves as participants exchange arguments. No *a priori* consideration can predict the direction or extent of the change. In other words, actual interaction among the metasystematic reasoners is needed in order to arrive at a Moral Stage 6 paradigmatic coordination. Adapting a valuable Habermasian concept, we call this social process of paradigmatic coordination discourse and its outcome (the agreement of all participants) consensus (Habermas, 1990).

Given the discursive properties of Moral Stage 6, there will be no single instantiation of Moral Stage 6. Moral Stage 6 respects the uniqueness of all individuals but integrates their idiosyncratic perspectives in a discursive solution. By contrast, at the end of Moral Stage 4, people see these individual differences as either arbitrary or contextual, without understanding how to transcend them. And at Moral Stage 5, people evaluate these differences according to universal principles.

Moral Stage 6 consensus is historically unique and limited. The ideal has to be saturated with the actual and divested of its context-free universalist validity claim that is typical of Moral Stage 5. Moral Stage 6 replaces hypothetical discourse with actual discourse, hypothetical universalization with real consensus. This in no way implies cultural relativism or the like, because the coordinations at Moral Stage 6 are highly constrained by hierarchical complexity requirements. Although there are many Moral Stage 6 solutions, the creation of each is highly unlikely.

Moral Stage 6 and the Individual

The central element of individuals' "reasoning at Moral Stage 6" is their realization that individual moral reasoning is limited and that the attainment of Moral Stage 6 is beyond their individual efforts. Individuals must have the potential to reason at Moral Stage 5, but, in addition, they must understand the dilemma of Moral Stage 5 reasoning and surrender part of their moral autonomy to the discourse where Moral Stage 6 reasoning takes place. At Moral Stage 6, there is no

issue of individual versus societal rights (Moral Stage 4), nor of their coordination (Moral Stage 5). Individual rights are constituted in the societal process, which, in turn, depends on the input of all individuals.

Individuals must be critically aware of the limitations of their own viewpoint of political, societal and moral issues. As Habermas points out, individuals have to replace strategic with non-strategic communication before real discourse can begin. Strategic communication is goal-oriented. By fixing an unconscious as well as conscious agenda of goals to be achieved in interaction, strategic communication makes it impossible to hear others accurately or to co-construct a perspective of reality. One has to have the strength to stop defending and asserting oneself while communicating. This may require metasystematic ego development and beneficence.

Kohlberg's more traditional Moral Stage 6 notions of unconditional respect for, and solidarity with, others (Kohlberg, 1986; Kohlberg, Boyd & Levine, 1990) can easily be accommodated within the new notion of the social nature of Moral Stage 6. Moral Stage 6 is the most radical expression of unconditional respect for others, because a Moral Stage 6 solution depends on the actual approval of all the other discourse participants. Each and every participant has "vetopower" against the possibility of a Moral Stage 6 solution. At this stage, people have to speak for themselves; there is no way in which all possible contributions can be simulated. Whereas local and particular views represent the effects of each person's biography, genes, and culture, people may arrive at shared, unifying principles that transcend these particularities. The concept of principles takes on a new meaning, however. At Moral Stage 5, moral principles are universal at the price of limiting respect for actual others; at Moral Stage 6, principles are subjected to discourse in order to universalize respect for others.

Whereas Moral Stage 5 solidarity rests on abstract principles that are considered valid for all humans, Moral Stage 6 solidarity extends to all humans precisely because there are no abstract universal principles. Everyone is needed as a potential partner in discourse and someone's failure to participate makes a Moral Stage 6 solution impossible. If, for some reason, people are incapable of articulating their point of view, a Moral Stage 5 solution can still be found for them by taking their "best interests," according to universal principles, into account. A Moral Stage 6 solution, by contrast, cannot proceed in such a vicarious way.

Moral Stage 6 "reasoners" know that perfect understanding between one another is impossible. Ultimately, all individuals are alone because they cannot completely take another's perspective. But at the same time, all individuals are viewed as interdependent, because everybody may have to rely on everybody else in discourse. Intermediate particularistic social groups, such as those based on gender, race, religious community, or nation, lose significance. There is a Hegelian dichotomization between the universal community of humankind on the one hand and the autonomous and emancipated individual on the other hand (also see Armon, 1989). The Moral Stage 6 collective identity is the cosmopolitan community of citizens of the world.

Moral Stage 6 and Methodology

The social construction of Moral Stage 6 requires methodological adaptations. Kohlberg's standard method of probing an individual's responses to a moral dilemma cannot detect Moral Stage 6 reasoning, because it is impossible to reason at Moral Stage 6 as an individual in isolation. Attempts to locate Moral Stage 6 in an individual are doomed to fail, because they look in the wrong place. Any method for studying Moral Stage 6 must concentrate on the social process of arriving at a consensus. This may require the construction of new dilemmas in which groups of Moral Stage 5 reasoning people are asked to work together in developing a solution.

At an individual level, the traditional technique of probing interviews accomplishes only the following two tasks. First, the individual's proficiency at Moral Stage 5 reasoning can be established. Moral Stage 5 reasoning is a precondition of Moral Stage 6 discourse. Second, the individual's understanding of the limitations of Moral Stage 5 and of the social quality of Moral Stage 6 can be examined. Thus, an individual's "Moral Stage 6 responses" are characterized by insight into the impossibility of individual Moral Stage 6 reasoning rather than by Moral Stage 6 reasoning.

Moral Stage 6 and Society

By locating Moral Stage 6 in the social enterprise of discourse, we acknowledge that an individual is unable to carry out a Moral Stage 6 reasoning process, without assigning religious, mystical, or transcendental properties to Moral Stage 6. Various scholars posited a transcendental stage at the top of their developmental hierarchy. Baldwin's (1906) version of postformal reasoning was transcendental. Gilligan and Murphy (1979) and Gilligan, Murphy and Tappan

(1990) suggested the need for a transcendental stage. Some of the people Kohlberg chose as exemplars for Moral Stage 6 had a transcendental apocalyptic view of the world, which supposedly liberated them from reasoning at Moral Stage 5; furthermore, Kohlberg briefly described a Moral Stage 7 with transcendental properties (Kohlberg, 1984; Kohlberg & Power, 1981; Kohlberg & Ryncarz, 1990). Some researchers (Alexander, Drucker & Langer, 1990; Koplowitz, personal communication, June 21-23, 1985) have divided the transcendental from the traditional developmental stages. For a discussion of transcendental solutions in terms of hard and soft stages see Kohlberg (1990).

The social construction of Moral Stage 6 entails a new relationship between moral reasoning and society. At all stages of development, moral reasoning is, of course, related to society. On the one hand, the political legitimation of a society must make sense to its members' moral reasoning. On the other hand, the societal environment influences the moral development of individuals. But at stages lower than Moral Stage 6, the connection between society and moral reasoning is relatively loose, whereas at Moral Stage 6, morality itself becomes societal. At this stage, individuals realize that they cannot be moral in isolation and that discursive processes are necessary to achieve satisfactory moral solutions. There can be no Moral Stage 6 sages in immoral times. Thus, moral concerns are firmly directed toward politics. Morality manifests itself in political processes of organizing discourses and achieving consensual solutions. We term this the politization of morality.

The politization of morality puts the political system under pressure to accommodate a discursive style of politics. As we noted above, modern societies have partly differentiated the political and moral domains and thereby created areas of relatively moral-free politics. This is clearly not the type of politics in which people with a politicized morality are interested. They might radically question the moral-free "politics as usual," demand a thorough discursive review of existing institutions, and advocate the strengthening of elements of direct democracy--a political system with enhanced egalitarian and democratic features, reduced hierarchical and authoritarian components, and a de-differentiation of morals and politics. We term such a program the moralization of politics. Its strength depends on how numerous and influential the people with a politicized morality are.

The discursive moralization of politics has its difficulties. Discourse, in the Habermasian sense, demands universal membership across space and time and is patently counterfactual. It is an ideal, utopian yardstick by which to evaluate real-life processes. Within our framework, a discourse does not necessarily need universal membership. The structural characteristic of discourse--the coordination of Moral Stage 5 coordinations--is the same whether five or five hundred people participate. Still, a real discourse requires a large amount of time and societal resources. Because of the inefficiencies of discourse, large-scale discursification might precipitate a crisis within the political system. A fully-discursive society seems impossible. The complexities of discourse and its limitations as a blueprint for a real-life society are discussed in more detail by Sonnert (1994).

So far we noted some external limitations to the moralization of politics through discourse. We now turn to the internal problems of discourse. Not every agreement qualifies as a Moral Stage 6 consensus (coordination of Moral Stage 5 coordinations). In fact, Moral Stage 6 solutions are very rare and difficult to achieve. From our transition-step model, we can derive empirical hypotheses about two structural "pathologies" of discourse that occur as groups and individuals move towards Moral Stage 6 (see also Appendix). We expect two converse pathologies, each located in a distinct phase of the transition. In the first phase (GSM 6a-3, Substep 2), the novel idea of discourse creates a great deal of excitement and enthusiasm, and people wish to apply it widely. Often, reasonably good solutions are generated through discourse, but the problem lies in a certain over-eagerness and naïveté among the discourse participants. They tend to underestimate the difficulty of creating a successful Moral Stage 6 solution. Overlooking inconsistencies and overgeneralizations, they may accept coordinations even if they are not correct. They also may not realize that a particular consensus was really based on strategic communication by some participants.

In the second phase (GSM 6a-3, Substep 3), the mood swings to skepticism about discursive solutions. People become experts in finding flaws with any proposal generated in discourse, detecting inconsistencies and overgeneralizations, as well as strategic communication. The problem now is that people have become too critical of proposed Moral Stage 6 solutions and reject even those that are viable. In this phase, they appear unable to produce any workable solution in discourse.

Conclusion

The study of moral development has traditionally examined individual moral development. At Moral Stage 6, however, the focus on individual moral development becomes insufficient, because moral reasoning can no longer be understood in terms of the individual. At this stage, moral reasoning is socially constituted. As a consequence, morality becomes politicized. The political system is considered immediately responsible for the attainment of morality. This in

turn may lead to a moralization of politics--the discursification of the political system. In this process, danger might beckon from the Scylla of a functional crisis of the political system if discursification goes too far, and the Charybdis of a crisis of legitimacy if the political system is perceived as unresponsive to Moral Stage 6 demands of discursification.

Author Notes

Earlier versions of this paper were presented at the *Third Beyond Formal Operations Symposium held at Harvard: Positive Development During Adolescence and Adulthood*, June, 1987, at the *Fourteenth Annual Scientific Meeting of the International Society of Political Psychology*, Helsinki, Finland, July 1, 1991-July 6, 1991, and at the *Sixth Adult Development Symposium*, Suffolk University, July 12-July 14, 1991. Charles Levine commented on an earlier presentation version and suggested a proof for a theorem. Victoria Alexander, Archie Brodsky, Eric A. Goodheart, Thomas G. Gutheil, Patrice M. Miller, David Moshman, and William Puka made valuable editorial suggestions. Requests for reprints should be sent to Michael L. Commons, Ph.D., Lecturer and Research Associate, Department of Psychiatry, Harvard Medical School, Massachusetts Mental Health Center, 74 Fenwood Road, Boston, MA 02115-6196.

References

Alexander, C. N., Druker, S. M., & Langer, E. J. (1990). Introduction: Major issues in the exploration of adult growth. In C. N. Alexander, & E. J. Langer (Eds.), Higher stages of human development: Perspectives on adult growth, (pp. 3-32). New York: Oxford University Press.

Armon, C. (1984a). Ideals of the good life and moral judgment: Ethical reasoning across the life span. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development (pp. 357-380). New York: Praeger.

Armon, C. (1984b). Ideals of the good life: A cross-sectional/longitudinal study of evaluative reasoning in children and adults. Unpublished doctoral dissertation, Harvard Graduate School of Education, Cambridge, MA.

Armon, C. (1989). Autonomy. In M. L. Commons, J. D. Sinnott, F. A. Richards, & C. Armon (Eds.), Adult Development: Vol. 1, Comparisons and applications of adolescent and adult developmental models. (pp. 179-196). New York: Praeger. New York: Praeger.

Arrow, K. J. (1951). Social choice and individual values. New York: John Wiley.

Baldwin, J. M. (1906). Thoughts and things or genetic logic (3 Volumes). New York: Macmillan.

Colby, A., & Kohlberg, L. (1987). The measurement of moral judgement: Vol. 1. New York: Cambridge.

Commons, M. L., & Grotzer, T. A. (1990). The relationship between Piagetian and Kohlbergian stage: An examination of the "necessary but not sufficient relationship." In M. L. Commons, C. Armon, L. Kohlberg, F. A. Richards, T. A. Grotzer, & J. D. Sinnott (Eds.), Adult Development, Models and methods in the study of adolescent and adult thought, 2. (pp. 205-231). New York: Praeger.

Commons, M. L., & Hallinan, P. W. with Fong, W., & McCarthy, K. (1989). Intelligent pattern recognition: Hierarchical organization of concepts and hierarchies. In M. L. Commons, R. J. Herrnstein, S. M. Kosslyn, & D. B. Mumford (Eds.), Quantitative analyses of behavior: Vol. 9, Computational and clinical approaches to pattern recognition and concept formation. (pp. 128-153). Hillsdale, NJ: Lawrence Erlbaum Associates.

Commons, M. L., & Richards, F. A. (1984a). A general model of stage theory. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development (pp. 120-140). New York: Praeger.

Commons, M. L., & Richards, F. A. (1984b). Applying the General Stage Model. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development (pp. 141-157). New York: Praeger.

Commons, M. L., Richards, F. A., & Armon, C. (Eds.). (1984). Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development. New York: Praeger.

Commons, M. L., & Rodriguez, J. A. (1990). "Equal access" without "establishing" religion: The necessity for assessing social perspective-taking skills and institutional atmosphere. Developmental Review, 323-340.

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). Mathematical psychology: An elementary introduction. Englewood Cliffs, New Jersey: Prentice-Hall.

Diessner, R. & Commons, M. L. with Arabos, V. N. (in preparation). Differentiating postconventional stages of justice reasoning.

Emler, N., Renwick, S., & Malone, B. (1983). The relationship between moral reasoning and political orientation. Journal of Personality and Social Psychology, 45, 1073-1080.

Fischer, K. W., Hand, H. H., & Russell, S. (1984). The development of abstractions in adolescents and adulthood. In M. L. Commons, F. A. Richards, & C. Armon (Eds.). Beyond formal operations: Late adolescent and adult cognitive development (pp. 43-73). New York: Praeger.

Flavell, J. H. (1963). The developmental psychology of Jean Piaget. New York: D. Van Norstrand Co.

Gilligan, C., & Murphy, J. M. (1979). Development from adolescence to adulthood: The philosopher and the dilemma of the fact. In D. Kuhn (Ed.), Intellectual development beyond childhood (pp. 85-99). San Francisco: Jossey-Bass.

Gilligan, C., Murphy, J. M., & Tappan, M. B. (1990). Moral Development beyond adolescence. In C. N. Alexander, & E. J. Langer (Eds.), Higher stages of human development: Perspectives on adult growth. (pp. 208-225). New York: Oxford University Press.

Gödel, K. (1977). Some metamathematical results on completeness and consistency; On formal undecidable propositions of Principia Mathematica and related systems I; On completeness and consistency. In J. Heijehoort (Ed.), From Frege to Gödel: A source book in mathematical logic 1879-1931. Cambridge: Harvard University Press (originally published 1930, 1931, 1931, respectively.).

Habermas, J. (1979). Communication and the evolution of society. T. McCarthy (Trans.). Boston: Beacon Press.

Habermas, J. (1990). Moral Consciousness and Communicative Action. C. Lenhardt & S. W. Nicholsen (Trans.). Cambridge, MA: MIT Press.

Hardin, G. (Ed.). (1977). Managing the commons. San Francisco: W. H. Freeman.

Inhelder, B., & Piaget, J. (1958). The growth of logical thinking from childhood to adolescence: an essay on the development of formal operational structures. (A. Parsons & S. Milgram, Trans.). New York: Basic Books (originally published 1955).

Kohlberg, L. (1967). Moral and religious education and the public schools: A developmental view. In T. Sizer, (Ed.). Religion and Public Education. Boston: Houghton Mifflin.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), Handbook of socialization theory and research (pp. 347-480). Chicago: Rand McNally.

Kohlberg, L. (1973). The claim to moral adequacy of a highest stage of moral judgment. Journal of Philosophy, 70, 630-646.

Kohlberg, L. (1976). Moral stages and moralization. In T. Lickona (Ed.), Moral development and behavior. (pp. 31-53). New York: Holt, Rinehart and Winston.

Kohlberg, L. (1981). The philosophy of moral development: Moral stages and the idea of justice. San Francisco: Harper and Row.

Kohlberg, L. (1984). Essays on moral development: Vol. 2. The psychology of moral development: Moral stages, their nature and validity. San Francisco: Harper & Row.

Kohlberg, L. (1986). A current statement on some theoretical issues. In S. Modgil & C. Modgil, (Eds.), Lawrence Kohlberg: Consensus and controversy. (pp. 485-546). Philadelphia, PA: The Falmer Press.

Kohlberg, L. (1987). Child psychology and childhood education: A cognitive-developmental view. New York: Longman.

Kohlberg, L. (1990). Which postformal levels are stages? Adult Development: Vol. 2. Models and methods in the study of adolescent and adult thought. (pp. 263-268). New York: Praeger.

Kohlberg, L., & Armon, C. (1984). Three types of stage models used in the study of adult development. In M. L. Commons, F. A. Richards, & C. Armon (Eds.). Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development (pp. 383-394). New York: Praeger.

Kohlberg, L., Boyd, D., & Levine, C. (1990). The return of stage 6: Its principle and moral point of view. In T. E. Wren (Ed.). The moral domain: Essays in the ongoing discussion between philosophy and the social sciences (pp. 151-181). Cambridge, MA: MIT Press (originally published 1986).

Kohlberg, L., & Kramer, R. (1969). Continuities and discontinuities in childhood and adult moral development. Human Development, 12, 93-120.

Kohlberg, L., Levine, C., & Hewer, A. (1983). Moral stages: A current formulation and a response to critics. Contributions to Human Development 10. Basel: S. Karger.

Kohlberg, L., & Power, C. (1981). Moral development, religious thinking and the question of a seventh stage. In L. Kohlberg (Ed.). Essays on moral development: Vol. 1. The philosophy of moral development (311-372). San Francisco: Harper & Row.

Kohlberg, L., & Ryncarz, R. A. (1990). Beyond justice reasoning: Moral development and considerations of a seventh stage. In C. N. Alexander, & E. J. Langer (Eds.), Higher stages of human development: Perspectives on adult growth (pp. 191-207). New York: Oxford University Press.

Kramer, G. H. (1977). A dynamical model of political equilibrium. Journal of Economic Theory, 16, 310-334.

Kuhn, D., & Brannock, J. (1977). Development of the isolation of variables scheme in experimental and `natural experiment' contexts. Developmental Psychology, 13, 9-14.

Kuhn, T. S. (1962). The structure of scientific revolutions. Chicago: University of Chicago Press.

Lindsay, P. H., & Norman, D. A. (1977). Human Information Processing: An introduction to psychology, (2nd ed), New York: Academic Press.

Luchins, A. S., & Luchins, E. H. (1965). Logical foundations of mathematics for the behavioral sciences. New York: Holt Rinehart and Winston.

Plott, C. R. (1967). A notion of equilibrium and its possibility under majority rules. American Economic Review, 57, 787-806.

Plott, C. R. (1971). Recent results in the theory of voting. In M. D. Intrilligator (Ed.). Frontiers of quantitative economics. (pp. 109-127). Amsterdam: North-Holland.
Plott, C. R., & Levine, M. E. (1978). A model of agenda influence on committee decisions. American Economic Review, 68, 146-160.
Power, F. C., Higgins, A., & Kohlberg, L. (1989). Lawrence Kohlberg's approach to moral education: A study of three democratic high schools. New York: Columbia University Press.
Puka, W. (1987). Moral Development without the philosophical captivation. Moral Education Forum, 12(4).
Puka, W. (1990). The majesty and mystery of Kohlberg's stage 6. In T. E. Wren (Ed.). The moral domain: Essays in the ongoing discussion between philosophy and the social sciences. (pp. 182-223). Cambridge, MA: MIT Press.
Rawls, J. (1971). A theory of justice. Cambridge, MA: Harvard University Press.
Reiser, S. J., Bursztajn, H. J., Gutheil, T. G., & Appelbaum, P. S. (1987). Divided staffs, divided selves: A case approach to mental health ethics. Cambridge, England: Cambridge University Press.
Rosenberg, S. W. (1988a). The structure of political thinking. American Journal of Political Science, 32, 539-566.
Rosenberg, S. W. (1988b). Reason, ideology and politics. Princeton, NJ: Princeton University Press.
Rosenberg, S. W., Ward, D., & Chilton, S. (1988). Political reasoning and cognition. Durham, NC: Duke University Press.
Selman, R. L. (1976). Social-cognitive understanding: A guide to educational and clinical practice. In T. Lickona (Ed.), Moral development and behavior. (pp. 299-316). New York: Holt, Rinehart and Winston.
Selman, R. L. (1980). The growth of interpersonal understanding. New York: Academic Press.
Selman, R., & Byrne, D. (1974). A structural developmental analysis of levels of role taking in middle childhood. Child Development, 45, 803-5.
Snarey, J. R. (1985) Cross-cultural universality of social-moral development: A critical review of Kohlbergian research. Psychological Bulletin, 97, 202-232.
Sonnert, G. (in press). Limits of morality. A sociological approach to higher moral stages. Journal of Adult Development, 1(2).
Trudeau, E. J., & Commons, M. L. (1991). Mathematically demonstrated hierarchical complexity of tasks and behavior development theory. Presented at the 17th Annual Convention for the Association of Behavior Analysts, May 1991.
Wagner, J. (1986). A critique of past and present theories of political tolerance: Discovering rational, democratic citizens through the application of developmental stages. Political Behavior, 8, 45-80.
Wagner, J. (1990). Rational constraint in mass belief systems: The role of developmental moral stages in the structure of political beliefs. Political Psychology, 11(1), 147-171.
Weinreich-Haste, H. (1986). Kohlberg's contribution to Political Psychology: A positive view. In S. Modgil & C. Modgil, (Eds.), Lawrence Kohlberg: Consensus and controversy. (pp. 337-361). Philadelphia, PA: The Falmer Press.

Appendix
Stage Transition Steps: A Voting Example

The following hypothetical example illustrates the transition steps from the Systematic Stage 5a (Moral Stage 4) to the Paradigmatic Stage 6a (Moral Stage 6). Choosing the familiar issue of voting systems for this example should make our transition model more intelligible. Voting systems are key determinants of what makes governments "just" in the citizens' eyes. They are thus among the core issues of political legitimacy, located at the intersection of moral and political reasoning.

In the case of voting systems, the issues and many of the possible solutions are well known. Hence, to some, the task of choosing of a particular voting system may appear relatively easy. Usually in research on moral reasoning, individuals are asked to discover solutions anew, without the benefit of historical knowledge. Under such circumstances, high-stage solutions would be much rarer and much harder to understand.

In this example, hypothetical subjects face the task of designing a just voting system, as they progress through the developmental transition steps. In the following, we examine only a few properties of voting systems to highlight the transition steps; we do not attempt complete descriptions of voting systems at the various steps. And, of course, the mentioned voting systems are not the only ones possible at each step. They represent examples drawn from a multitude of possibilities.

Fully equilibrated Moral Stage 4

(GSM Stage 5a-4) *A = A* with B**. The subjects perceive voting systems either as systems of rights or as systems of duties. The subjects are thus divided into two groups: One group takes the perspective of voters' duties. This group

emphasizes the voting system's purpose of generating a stable government. Therefore, it prefers a winner-takes-all constituency system (system A) because it encourages the existence of only two large parties and tends to lead to landslide victories. The other group takes the perspective of individual voters' rights. It advocates a system of proportional representation (system B) because, in its opinion, this will better represent individuals' views.

Transition Steps from the Systematic Stage (Moral Stage 4) to the Metasystematic Stage (Moral Stage 5)

(GSM Stage 5b-0) *A.* Subjects discover their Systematic Stage 5a solutions to be failures for a variety of reasons. For instance, because constituency boundaries are crucial for the outcome of elections under system A, there are heated and protracted battles about how to draw these boundaries. On the other hand, the proportional system (system B) brings about a proliferation of political parties and groups that permanently destabilizes government. "Horse-trading" in coalition negotiations leads to results that do not represent the will of the voters. In particular, splinter groups may hold the balance of power and thereby become more influential than their numerical strength would warrant.

(GSM Stage 5b-1) *B.* Subjects shift their attention from one voting system to another. Under the weight of the criticism leveled against their favorite voting system, those who supported the constituency system now become supporters of proportional representation, and vice versa. For instance, subjects who previously stressed the voters' duty to support the societal system (system A) now emphasize their own rights as voters (system B). Kohlberg considered this shift to being self-oriented as the first part of Moral Stage 4/5. At this transition step, however, the choice of an alternative single voting system fares no better than the previous system. Soon, the weak points of the alternative system are also discovered, but the subjects do not know how to escape this dilemma.

(GSM Stage 5b-2) *A or B.* Subjects begin to alternate between one voting system and another. "Cultural relativism" and "dualism" prevail. Subjects now point out that voting systems are embedded in Moral Stage 4 cultural, political, and historical frames of reference. Kohlberg referred to this relativism as the second part of Moral Stage 4/5. Subjects say that one voting system may be good in one country or area of politics, the other in others. This step brings a concurrent application of apparently contradictory systems. Contradiction is avoided by contextualization. Yet there is no true integration of systems.

(GSM Stage 5b-3) *A and B.* At this step, subjects would like to combine the different voting systems, but do not do so in a systematic and sensible manner. Systems of rights and systems of duties are considered together, but there is no full reversibility (e.g., subjects do not clearly understand that voters' rights correspond to voters' duties).

At **Substep 1**, elements or subsets of elements from voting system A and voting system B are combined in a non-systematic manner. Subjects attempt to construct a supersystem that orders the contextualization of the voting systems. For instance, to guarantee individual rights, subjects propose instituting proportional representation (system B) in national elections. To enforce the stability of the government (voters' duties), they apply the constituency system (system A) in local elections. But it becomes apparent to the subjects that this is not a good solution. The problems of each voting system continue to manifest themselves in their respective arenas.

At **Substep 2**, elements or subsets of elements from voting system A and voting system B are chosen so as to produce a supersystem that has most of the desirable features of system A and system B. The resulting synthesized supersystem generates useful metasystematic solutions to voting problems for the most part. For instance, subjects propose a modified constituency system with multi-representative constituencies. Within each constituency candidates run "at large," and the three top vote getters are elected. Subjects argue that this will increase the representativeness of the voting system without risking fragmentation, a problem inherent in systems based on proportional representation. They do not realize that their proposal still might be problematic.

At **Substep 3**, subjects discover and reject overgeneralizations, inconsistencies, and other shortcomings in voting supersystems, yet no adequate solutions are proposed. For instance, subjects realize that, in the voting supersystem from substep 2, "gerrymandering" is still a problem and that the multi-representative constituencies invite countless forms of strategic voting. Problems with coalitions and government instability appear to persist. But it remains unclear to the subjects how to accommodate both representativity and stability in the voting supersystem; there is no successful coordination.

Fully equilibrated Moral Stage 5

(GSM Stage 5b-4) *A' = A with B*. Subjects now discover new relationships and interrelationships between the voting systems that allow them to coordinate the voting systems successfully. For the most part, the subjects produce useful solutions and avoid most overgeneralizations. Diessner and Commons (in preparation) show that this fully metasystematic culmination (5b-4) is in fact Kohlberg's old Moral Stage 6. Subjects have learned that rights and duties are in a reciprocal relationship to one another--the double golden rule: One person's rights are others' duties, and vice versa (Kohlberg, 1986).

A consistent supersystem (A) is formed. For instance, subjects base this supersystem on the utilitarian principle of the greatest good for the greatest number and design it to protect the interests of the abstract "average voter." The principle of the "average voter" may translate into an emphasis on majority rule: the majority viewpoint should be expressed in the government. Subjects realize that the voters have not only the right to see their preferences accurately represented, but also the duty to support stability and governability of the political system. This duty arises from the right to stable and efficient government. The government, conversely, has not only the right to have stable majorities in the elective body--a right which arises from the duty to provide stable and efficient government--but also the duty to represent the population's preferences.

Subjects now propose a voting supersystem which is based on proportional representation but which excludes small parties with less than 10% of the popular vote from representation. Subjects argue that this supersystem adequately expresses voters' preferences while preventing political fragmentation and instability (Kramer, 1977).

Transition Steps from the Metasystematic Stage (Moral Stage 5) to the Paradigmatic Stage (Moral Stage 6)

In the following section of our example, we introduce an alternative to the voting supersystem grounded in the concept of the "average voter" (supersystem A). The alternative supersystem (B) is based on the concept of the "least advantaged voter" derived from Rawls' (1971) philosophy.

(**GSM Stage 6a-0**). *A'*. As in the last step, the subjects favor voting supersystem A (based on the notion of "the average voter"). Supersystem A contains a consistent subset of possible conditions for a voting system. But however admirable, this supersystem is ultimately doomed to fail when a new situation arises which it cannot adequately address. Subjects come to recognize the inherent weakness of the supersystem under which they have been operating. For instance, small minorities who comprise less than 10% of the population complain that the 10% threshold completely wipes out their representation.

(GSM Stage 6a-1) *B'*. Subjects come to reject voting supersystem A and consider an alternative supersystem. Supersystem B consists of a different subset of possible conditions for a fair voting system. It is based on a rival principle, such as the concern for "the least advantaged voter." Subjects oppose the majority rule approach of supersystem A because it does not sufficiently protect disadvantaged minorities. They consider it essential that, in supersystem B, representation be provided to disadvantaged minorities, even if these minorities are numerically small.

Subjects now propose a legislative body consisting of representatives of societal groups. Each group gets the same number of representatives, regardless of numerical size. The groups specifically designated as disadvantaged groups cannot be overruled. Subjects soon recognize a number of problems with this alternative supersystem. For example, this supersystem puts the government at the mercy of the representatives of the disadvantaged groups. This arrangement threatens governability. Moreover, it sharply discriminates against members of larger, and not disadvantaged, societal groups in terms of representing their views in government.

(GSM Stage 6a-2) *A' or B'*. Subjects consider a number of proposals and attempt to contextualize them. For example, the fair voting proposals of supersystem A and of supersystem B are alternatively supported and or rejected depending on the context. In viewing a homogenous society, subjects consider supersystem A, based on "the average voter," fair because the citizens in this society are seen to have very similar interests and an almost identical outlook on life so that the "average voter" would, in effect, represent everybody and thus converge with the "least advantaged voter." In viewing a heterogeneous society with conspicuous castes or underclasses, these same subjects consider supersystem B, based on "the least advantaged voter," fair. Supersystem B, subjects say, gives a greater voice to minority interests by including them in coalition-style politics. Subjects think that as long as the minority factions are not too numerous or too discontent with the politics of the majority, supersystem B would function in a manner

conducive to the welfare of everybody. Such a metasystematic contextualization appears to make the implementation of rival supersystems possible without jettisoning universal principles altogether.

Yet the dichotomous notion of two types of societies is overgeneralized and remains blind to critical problems within societies. Even the supposedly homogeneous societies contain "invisible" minorities. And empowering disadvantaged minorities in heterogeneous societies can lead to the reversal of oppression, rather than to its elimination. Contextualization does not achieve a true integration of supersystems.

(GSM Stage 6a-3) *A' and B'*. Subjects attempt an integration of the voting supersystems. Now, they begin to understand that there is no abstract "ideal" voting supersystem and generally recognize the necessity of actual discourse in order to coordinate competing supersystems.

At **Substep 1**, subjects consider actual dialogue the preferable way of combining contradictory voting proposals of the metasystematic stage. But because they still think they can coordinate the supersystems monologically if necessary, they tend to revert to monological reasoning when dialogue appears to run into difficulties. In our example, subjects just combine the two supersystems into a single the form of government: namely a bicameral legislature. House A is elected according to supersystem A and house B according to supersystem B. As to coordinating the relationship between the two houses, subjects propose that both houses must concur on each legislative action. The necessity of concord between the two houses results in a complete stalemate when it becomes an absolute condition for the exercise of political power.

At **Substep 2**, the subjects' discourse begins to generate a paradigm that coordinates voting supersystem A, based on the principle of the "average voter," and supersystem B, based on the principle of the "least advantaged voter." Now subjects unreservedly place their faith in the discursive process itself, rather than in some set of abstract principles (metasystematic, Moral Stage 5) or laws (systematic, Moral Stage 4). They understand the necessity of co-constructing solutions in discourse. Subjects realize that the discursive process must be non-strategic. No one should conceal their goals either from the other parties or from themselves. The preferences of the average voters as well as those of the least advantaged voters must coalesce around a solution to the problem at hand. Both "average voters" and "least advantaged voters" need to feel satisfied by their participation in reaching that solution. A key issue is how to define who is disadvantaged. Subjects understand that such a definition should consider the views of disadvantaged voters speaking for themselves (supersystem B), as well as the average voter's view of what constitutes disadvantage (supersystem A).

Subjects see that their own discourse cannot substitute real discourse among all citizens; therefore they consider their own solutions tentative. Among the subjects, there is great enthusiasm about the newly-established idea of co-constructing solutions through discourse, and a tendency to apply it to as many problems as possible.

Subjects decide to give house B priority over house A, because they think that such an arrangement would best provide redress to the plight of the disadvantaged after a long history of neglect and oppression. This coordination solves more problems of interest to the subjects, i.e., it provides for governability, but it still produces overgeneralizations. Problems can now occur in two realms because of the emphasis on the discursive process. Not only might solutions be unworkable; they might also have been generated in a non-discursive, i.e., strategic way. But at this step, subjects are not well-versed in detecting problems in either direction.

At **Substep 3**, subjects develop a proclivity to reject false attempts at coordinating supersystems. The attitude of optimism about discursive solutions from the previous substep turns to skepticism or pessimism. Subjects realize that previous attempts to coordinate Moral Stage 5 supersystems are either incomplete (do not cover all important issues) or inconsistent (include conflicting rules).

In our example, subjects find that the solution of substep 2 basically reduces house A to window-dressing because it is being constantly overruled by house B. Moreover, subjects discover that the previous discursive solution was based on strategic misrepresentations of disadvantage. Some subjects who had successfully argued for the inclusion of certain societal groups among the "disadvantaged groups" in house B had had the strategic agenda to establish a power base for those groups' "leaders." These group representatives now frequently pursue their own power interests rather than representing the truly disadvantaged. The problem is that, in their general disenchantment with discourse, subjects tend to reject not only bad solutions, but also good ones.

Fully equilibrated Moral Stage 6

(GSM Stage 6a-4) *A" = A' with B'*. At the fully paradigmatic Moral Stage 6, non-strategic discourse co-constructs a paradigm for addressing the dilemma posed by various conflicting voting supersystems. In our example, a paradigmatic voting system emerges from the subjects' discourse. Subjects emphasize that real discourse among the voters would be needed to make the subjects' proposal a true Moral Stage 6 voting system. They also note that an alternative to a paradigmatic voting system would be to decide as many questions as possible in real discourse among the citizens, rather than delegating these decisions through a voting system.

Subjects now outline three areas in which the relationship between the two houses differs: in the first area, house B has priority; in the second area, house A has priority; and in the third area, the houses have to agree (e.g., basic constitutional changes).

At this stage, non-strategic and feasible coordinations of metasystematic supersystems are generally accepted, whereas strategic and non-feasible coordinations are rejected. Subjects stress that ultimately the people themselves have to determine to which societal group they belong and how strongly they identify with it. Emphasis is put on democratic structures within societal groups. Strategic claims by group leaders that their group is disadvantaged are unacceptable.

Subjects realize that a paradigmatic solution depends on the circumstances of the society in question. The degree of societal homogeneity or fragmentation and the manner in which groups constitute their collective identities and allegiances are decisive in determining an appropriate paradigmatic voting system. The subjects realize that people cannot define an "ideal" voting system once and for all. Rather, the voting system has to be open to discursive review as societal conditions change.