False Rationality and The Tragedy of The Commons: Toward a More Pertinent Approach to Social Dilemmas

Leandro F. F. Meyer
Instituto Socioambiental e dos Recursos Hídricos
Universidade Federal Rural da Amazônia
Belém, PA – Brazil

Marcelo José Braga
Departamento de Economia Rural
Universidade Federal de Viçosa
Viçosa, MG - Brazil

Prepared for delivery at the Workshop on the Ostrom Workshop (WOW5) conference, Indiana University Bloomington, June 18-21, 2014.

Corresponding author: Leandro F. F. Meyer, Instituto Socioambiental e dos Recursos Hídricos, Universidade Federal Rural da Amazônia, Belém-PA, Brazil.
E-mail: leandro.meyer@ufra.edu.br

ABSTRACT

In Seven Complex Lessons in Education for the Future, Edgar Morin points to *disjunction*, *false rationality*, *reductionism* and *closed specialization* as essential problems that challenge our ability to generate pertinent knowledge in general. The standard theory of social dilemmas offers one of the most striking examples of how pertinent knowledge can be lost amid the rationalizations generated by disciplinary specialization. The latest developments in the study of social dilemmas devote an increasing amount of attention to cognition, belief systems, valuations, and language. However, the developments in this field operate almost entirely under epistemological assumptions that recognize only the instrumental form of rationality and deny that "value judgments" or "moral questions" have cognitive content. This standpoint erodes the moral feature of the choice situation and prevents the acknowledgment of the links connecting cognition, inner growth, and moral reasoning. It also deemphasizes the significance of these links to achieving cooperative solutions to many social dilemmas. Concurrently, this standpoint renders mysterious the role of communication and mutual understanding in promoting cooperation in those situations. The presentation brings the epistemological issue to the fore in order to introduce a proposal that enlarges the Institutional Analysis and Development framework by integrating moral cognitivism and Action Logic into it to describe orders of development as discrete meaning making stages. The presentation advances an empirical strategy to test the power of alternative models of human valuation to predict the mixed choices of the participants in social dilemmas experiments under similar institutional conditions, including different uses of communication.

INTRODUCTION

In *Seven Complex Lessons in Education for the Future*, Edgar Morin (1999) points to *disjunction*, *false rationality*, *reductionism,* and *closed specialization* as essential problems that challenge our ability to generate pertinent knowledge in general. Pertinent knowledge, in turn, is construed as resulting from learning processes that organize and articulate a wide range of scattered information about the world and that enable us to grasp key problems within their relevant contexts, multidimensional totalities and systemic relations.

The mainstream approach to social dilemmas within the field of economics offers one of the most striking examples of how pertinent knowledge can be lost amid the disjunctions, reductions, and rationalizations generated by disciplinary specialization. This statement follows from the recognition that, in conducting

institutional analysis and design to cope with morally relevant conflicts of action, the bearing of knowledge has more to do with how we account for the heterogeneity of action logics and conditional choice patterns than with our ability to support a "representative agent" and produce wide generalizations.

Along with the exertion to move institutional analysis beyond the classical state-market dichotomy, the Ostroms' studies call attention to the limitations ensuing from a generalized application of the standard model of "economic man" and its associated self-interest assumption. As Aligica (2013) puts it, Elinor Ostrom has repeatedly drawn attention to the problem of actor and social heterogeneity and its implication for both institutional order and institutional theory. Actually, "[t]he homogenization, supra-simplification, and formulaic conceptualization is our theories of institutions is in many respects a function of a parallel homogenization of human agents that we practice at micro-level" (Aligica, 2013, p. 5)

Yet, as institutional economists seek to incorporate the heterogeneity of concrete individuals in the study of institutions to copy with social dilemmas, the lack of a theory of human valuation specifying the determinants of individuals' utility judgments renders the prediction of both expectations and behaviors in social dilemmas virtually impossible. Heterogeneous agents may respond quite differently, and often inconsistently, to similar incentive structures. This fact jeopardizes, in turn, the relevance of our proverbial understanding that human agents are not homogenous as it hinders our ability to set up arrangements of real help to cope with social dilemmas rather than rules that may make things worse.

We suggested elsewhere (Meyer; Braga, 2014) that scholars can overcome this difficulty by integrating models of sociocognitive and moral development into the framework of institutional analysis. The perspective of developmental psychology was

presented as one that can provide the principles to understand the heterogeneity of action logics, motivations and valuations, and improve the predictability of behaviors brought about by alternative incentives. In that study, we presented results from hypothesis testing linked to a cognitivist-developmental theory of human valuations. These results indicated that cooperative motivations and choices in public goods provision dilemmas (as well as in common-pool appropriation dilemmas) are associated with further *stages* of interior development of the participants. We then suggested that, in order to provoke the expected behavior or the intended choice, institutions should be tuned to the characteristics of each psychosocial centralization stage, as motivational needs, aims, and means differ between each stage, and that incentives could be designed to promote swifter movement of individuals along the path of interior growth.

We believe that the reception of these results and conclusions has been disadvantaged, nonetheless, due to the presence of unquestioned epistemological assumptions leading to the belief that the instrumental form of rationality is the only form of rationality and that "value judgments" or "moral questions" have no cognitive content. Even though the critical role played by heterogeneous valuations and motivations is recognized in modern institutional analysis, this epistemological standpoint prevents the acknowledgment of the links connecting cognition, inner growth, and moral reasoning and deemphasizes the significance of these links to achieving cooperative solutions to many social dilemmas. Concurrently, this standpoint renders mysterious the role of language and communication in promoting mutual understanding, shared expectations, and cooperation in those situations.

In this article we bring the epistemological issue to the fore in order to introduce a proposal that, as we see it, opens the Institutional Analysis and Development (IAD) framework to receive moral cognitivism and multiple action logics to describe orders of

development as discrete meaning making stages. We began by pointing to the epistemological matter as the main barrier preventing the recognition of the validity of the developmental perspective on human valuations. Next, we illustrate the bearing of our suggestion by addressing the disputed role of communication in promoting cooperation among individuals who strive to solve their social dilemmas by themselves. We argue that the Bloomington School approach to communication has been limited *not* so much by its alignment with a form of radical pluralism that sets it apart from the Habermasean account but rather due to its unchallenged adoption of the foundationalist epistemology. This perspective eventually leads the Bloomington School to stick with a unitary form of rationality—the instrumental—and thus becoming *less* pluralistic than Habermas's view, which comports two types of rationality—the instrumental and the communicative.

Following Habermas (1984, 1987), we argue that because foundationalism severs the internal connection between norms and justifying grounds, it conceals the noninstrumental feature of rationality and distorts the significance of communication in producing normative agreements in morally relevant conflicts of action. Working as a blinding paradigm, foundationalism discredits moral argumentation—over and above any form of spiritual knowledge—and redirects both research and teaching about social dilemmas toward an exclusively instrumental-utilitarian approach. Admitting that different methods are performative of different realities (Law; Urry, 2003; Esbjorn-Hargens, 2007), we close the article by highlighting the significance of moving beyond foundationalism in order to enhance the pertinence of our knowledge about social dilemmas and set up a research program that further the pluralistic bent of the Bloomington School of Intuitional Theory.

A number of factors have been preventing the recognition of the implications of many reputed developmental models to approach rational action in social sciences. A basic difficulty ensues from the common belief that psychosocial development refers to childhood and adolescence, that is, to the first twenty years of life. "Traditionally," as Marchand (2005), puts it; "experts in developmental psychology analyzed the growth of the child and of the adolescent, holding that development ends before adult life begins" (cf. e.g. Inhelder; Piaget 1955; Piaget, 1970/1972). Were that the case, this fact would evidently move the developmental framework out of serious consideration for addressing social dilemmas, since the actors involved in most relevant situations are typically adult humans. Currently, however, researches are coming to an increasing understanding that human subjective and intersubjective developments have indeed the potential for evolving all the way through the adult life (Graves, 1971; Riegel 1973; Arlin 1975; Basseches 1980, Kramer 1983; Pascual-Leone 1984; Commons, Richards and Armon 1984; Commons, Sinnott, Richards and Armon 1989; Sinnott 1984, 1989; among others).

A second difficulty is that each of the numerous facets or streams of consciousness comprising the overall *Self* appears to have its own internal drives or laws of transformation toward greater complexity and integration. When considered in its entirety, the overall Self of particular individuals does not show a sequential or stage-like development, but appears instead as a rather fluid and flowing affair, with much overlapping and interweaving (Wilber, 2000b, p. 34). The simple intuition of what seems to be an almost infinite number of multiple modalities of individual personalities stirs a natural sense of incommensurability supporting ordinary relativistic objections against the stage developmental framework in general. However, modern psychological

structuralism takes all that intertwining into account and entails careful methodological design for assessing particular streams of consciousness and specific self-related competencies, which are defined as capacities not only to solve but also to recognize the very existence of particular types of problems (e.g. empirical-analytic, moral-practical or interpersonal relationship). Along these lines, as Wilber (ibid) reports, the bulk of research has continued to find that each self-related developmental line itself tends to unfold in a stage-like, sequential, and nested hierarchical fashion, and that self's *center of gravity*, so to speak, tends to hover around one basic action logic at any time (p. 35). Furthermore, according to him, "One of the striking things about the present state of developmental studies is how similar, in broad outline, most of its models are" (p. 5). In fact, by comparing a sizeable number of developmental models and theories, also Richards and Commons (1990) indicate that "The stage sequence [in all of that theories and modes] can be aligned across a common developmental space," and that, "The harmony of alignment shown suggests a possible reconciliation of [these] theories…" (p. 160; see also Commons, 1981).

Yet, when it comes to the subject of morally relevant conflicts of action, as in social dilemmas, the acknowledgment of the developmental framework is hindered most of all by the common idea that "value judgments" or "moral questions" are rationally undecidable. As Heath (2001) indicates, a critical consequence of this view, which is often unstated, is that "most social theorists simply assume that any agent who acts on the basis of a moral principle, or social norm, is not rationally justified in doing so" (p. 2). According to him, "This is what underlies the widespread tendency among social theorists to assume that instrumental action is the only form of rational action, and that norm-governed action must have some kind of nonrational source, such as conditioning, socialization, or habit" (ibid). Heath further points to how the presumption

of non-rationality makes it tempting to abandon the action frame of reference and supply purely functionalist explanations for the coherence of norm systems and the adaptability of norm-governed action. This trend is noticeable in the current blast of interest in the sociobiological evolutionary framework for explaining human sociability and adherence to norms—visible also in connection with the Bloomington approach to the stability of normative agreements.

As suggested, the mentioned hindrance is epistemological in nature. In Heath's words, "The traditional reason for thinking that normative commitments are irrational, or unjustifiable, depends upon a rather specific conception of rationality and justifiability known as foundationalism" (ibid, p. 2). As Heath (2001) summarizes it, foundationalism is a theory of justification that is intended to provide an answer to the fundamental problem in epistemology: the problem of infinite regress. Foundationalism suggests that any attempt to justify a given statement inferentially generates an infinite regress of new arguments that can be introduced in support of that statement but that will contain premises that themselves require justification. The only way to break this cycle, Heath remarks (p. 197), is to use the conclusion as a premise (i.e., to use circular reasoning) or to simply break the chain of reasons (i.e., to make an undefended assumption).

Foundationalism represents an instance of the latter strategy, as it holds that there is a class of "basic" beliefs (also called foundational beliefs) that are intrinsically (i.e., non-inferentially) justified by virtue of their empirical content. Foundational beliefs are said to be self-justifying or self-evident to the extent that they are not justified by beliefs or constructs other than sensorial perception.

Because validity claims that make moral statements and normative commitments cannot be grounded in any direct experience with the physical world, the foundationalist epistemology implies that these judgments are essentially non-cognitive.

This epistemological underpinning has obvious implications for how we explain rational choice in morally relevant conflicts of action. In particular, this view influences the role of communication in producing normative commitments in these situations.

While the instrumental conception of rationality does not itself presuppose or depend upon any sort of moral noncognitivism (Heath, 2001), the foundationalist standpoint is probably what explains, for instance, Hardin's (1968) emphatic repudiation of moral argumentation for addressing commons dilemmas.

He regards any attempt to instill a sense of responsibility in others as "tempting to anyone who wishes to extend his control beyond the legal limits" (p. 1247). Hardin goes as far as to draw on Batson's "double bind" situation—hypothesized as being part of the genesis of schizophrenia—in order to denounce the "serious pathogenic effects" resulting from any appeal to the responsibility of individuals that is intended to regulate social interactions (ibid).

Batson's double bind theory of the origin of schizophrenia involves the same sort of contradiction between verbal and non-verbal communication that Hardin applies in *The Tragedy of the Commons*, implying a necessary skepticism about any plea to the "conscience" regarding the regulation of resource use in the absence of sanctions. However, Batson's double bind principle involves a strong emotional and unidirectional message from a mother toward her infant, who cannot dispel the anxiety brought about by the contradiction through further communication (Tarnas, 1999, p. 445, cf. Bateson, 1972; 1979). Clearly, linguistic communication among adults working to overcome

their collective action problems is different in this respect; what Hardin actually presumes is that moral claims and verbal promises are fundamentally non-binding.

BEYOND FOUNDATIONALIMS: TOWARDS STAGES OF MORAL REASONING

In the previous section, we attempted to show that because moral noncognitivism holds that moral judgments are purely relative, it conflicts with both the idea that social norms can be justified using arguments (rather than simply being imposed by coercion or force) and the idea that rational agents can obey norms on the grounds of a recognition of the norm's *validity* rather than exclusively considering an utilitarian calculus or some non-rational motive. These limitations motivated Habermas (1984/1987) to maintain that instrumental models do not provide a sufficient basis for a *general* theory of rational action.

Heath (2001) indicates that a common response to the relativist point of view on morals has been to accept the formal component of the foundationalist analysis, seeking only to deny the narrow empiricist list of belief-types that are claimed to be available for "objective validation." However, he also notes that these theories suffer from well-known difficulties; thus, the relativist position seems quite strong in this context. Conversely, rooted in the new paradigm of epistemology generated by the *linguistic turn*, Habermas's strategy in responding to the relativist stance on morality denies the force of the regress argument entirely and is governed by a non-foundationalist defense of the cognitivist conception of moral judgment.

Following Heath's (ibid) outline, Habermas's discourse-theoretic view has two basic components. First, Habermas claims that non-cognitivist concerns about the truth-aptness of moral judgments is important only if one assumes that the truth represents some kind of correspondence relationship between sentences and the state of affairs in

the world.[1] If one denies that this sort of "objectivity" plays any role in vindicating the truth-claims associated with beliefs, then our ability to justify beliefs has nothing to do with their reference to the physical world (i.e., with the knowledge of what can be enacted using individuals' ordinary senses and their extension). Similarly, when the relativist questions the ultimate justifiability of moral judgments, the argument is persuasive only if it presupposes a "monological" conception of rational justification (i.e., when justification is tacitly treated as a process that involves only the agent's cognitive states and the objects of representation). This assumption has the effect of reducing all public practices of justification such that they are either secondary or derivative. However, if one assumes, as Habermas does, that *justification is always dialogical*—a process involving an attempt to justify a claim to some other person, so that justification *to others* is taken as the primary phenomenon—then there is no a priori reason to think that moral questions are any less decidable than empirical or scientific ones.

In summary, Habermas suggests that one can defeat moral non-cognitivism by rejecting the traditional project of analytic epistemology, including both the received (correspondence) theory of truth and the received (foundationalist) view of justification. Heath also suggests that one reason some theorists have taken this more radical step is that "foundationalism does not offer a very persuasive justification for any kind of belief, including empirical ones" (2001, p. 198).

Despite the revolutionary tone of this epistemological turn, "the first thing to notice about Habermas's theory of communicative action" as Heath (2001, p. 13) observes, "is that it is a *typological* theory" (emphasis in the original). In presenting his theory of communicative action, Habermas does not reject the instrumental conception

---

[1] Habermas refers to the correspondence theory of truth.

of rationality. As Heath (ibid) explains, Habermas takes as his point of departure that agents have access to a set of different, often incommensurable, standards of choice. Communicative action is action governed by a particular standard—namely, that of achieving *understanding*—whereas instrumental action is action that is governed by a different standard: that of achieving *success*.

According to Habermas's typology, *instrumental action* and *speech acts* form two "elementary forms of action" (Figure 1). The introduction of a second agent generates social action, understood as a complex phenomenon constructed through interaction between the two elementary forms of action. According to this view, rational agents engaged in social action always face the problem of interdependent expectations, which can be resolved by drawing upon the resources of either instrumental action or speech. When the actors are primarily interested in the consequences, social action becomes strategic action in the standard game-theoretic sense. However, when speech is used to coordinate the regress of anticipations, it generates the form of action that Habermas characterizes as communicative action (cf. Habermas, 1990, p. 133).

This basic scheme is indicated by the straight lines in Figure 1. The upward oblique line indicates that communicative action is not the same as speech. Like strategic action, communicative action also presupposes the basic teleological structure
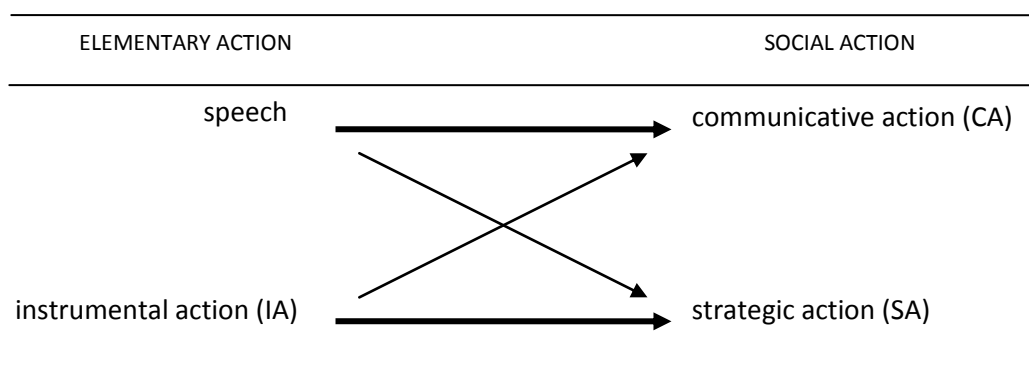

Figure 1: Elementary action types combine to produce social action types.
Source: Adapted from Heath (2001, p. 25)

of action inasmuch as the actors are assumed to continue to conduct their plans to attain a particular state of affairs. In Habermas's (ibid) words, the two social action types differ in that "for the model of strategic action, a structural description of the action directly oriented toward success is sufficient, whereas the model of action oriented toward reaching understanding must specify the propositions of an agreement to be reached communicatively that allows alter to link his action to ego's" (where the alter and ego are persons) (p. 134). In other words, when engaged in communicative action, actors are assumed to be "prepared to harmonize their plans of action through internal means, committing themselves to pursuing their goals only on the condition of an agreement—one that already exists or one to be negotiated—about definitions of the situation and prospective outcomes" (ibid).

When considering the empirical differences in the extent to which different groups or societies depend upon explicitly discursive procedures to secure social integration, Habermas offers a plain stage developmental account. In broad lines, his argument is aimed at showing that the stages that occur in his historical reconstruction of the development of communicative action, which takes the form of an interpretation of work by Emile Durkheim and George Herbert Mead,[2] are recapitulated in the ontogenesis of our capacity to speech and act, and are isomorphic to the stages described in Laurence Kohlberg's model of the development of sociocognitive and moral reasoning.[3] The connecting links are provided by Selman's account of sociocognitive development in relation to stages of social perspective taking, which Habermas reformulates in terms of structures of social interaction (see Table 1). "The point of this chain of argument is to connect structures of moral judgment to structures

---

[2] This phylogenetic account is in the fifth chapter of The Theory of Communicative Action (1984/1987, v. 2).

[3] The ontogenetic ground of Habermas's onto-phylogenetic parallel is developed in the fourth chapter of Moral Consciousness and Communicative Action (1990). The interested reader may want to consult in addition the third chapter of Justification and Application (Habermas, 1993).

of social interaction in such a way that their developmental-logical features stand out more clearly" (McCarthy, 1990, p. ix).

By defining discourse as reflective form of communicative action, Habermas situates the morally relevant presuppositions of practical argumentation as the tail end and point of reference in a constructivist *learning* process, in which complex forms of social action have given rise to competences resting on repeatedly reorganized sociocognitive inventories and perspective structures that have, in turn, permitted the emergence of more sophisticated forms of action. Viewed within the development of a complex structure of perspectives that culminates in a *decentered understanding of the world* displayed by subjects who act with an orientation toward reaching understanding, Habermas distinguishes *stages of interaction* in terms of different *achievements of coordination*, expressing a development that is *directed* and *cumulative*.

Skipping the details of Habermas's reconstruction, we single out only the reasons that, according to him, the ability to act from the perspective of a strict concept of morality (as an autonomous and rational sense of duty) can evolve only at the postconventional level, while the ability for acting strategically requires only an actualization of the structure of perspective applying to the preconventional level, without requiring any further reorganization of the sociocognitive inventory.

To show that this occurs, Habermas first redefines the preconventional types of action in terms of *forms of reciprocity* linked to different *structures of behavioral expectations* (not shown in Table 1). In this fashion, interaction controlled by authority is redefined in terms of an *asymmetrical form of reciprocity* which tends to obtain whenever the authority to control others' contributions to the interaction is unequal, as in the family. Conversely, the *symmetrical form of reciprocity* obtains when the participants exercise mutual control over their contributions to the interaction, as in

egalitarian friendship, for example (p. 147). These differentiations are correspondingly reflected in two different forms of action coordination: *authority-governed complementarity*, and *interest-governed reciprocity*, to which actors can resort in the face of both cooperative and competitive relationships.

Habermas suggests that authority-governed complementary and interest-governed symmetrical social relations define two different types of interaction that can embody the same perspective structure, namely: the reciprocity of action perspectives typical of Selman's level 2 of perspective taking (Table 1). According to Selman, the sociocognitive inventory of children at this level—i.e. analogously structured concepts of behavioral expectations, authority, motives for action, and the ability to act—enables them to control interactions by deception if necessary. An asymmetry between the developmental requisites for strategic action and action oriented toward reaching understanding starts to emerge as we recognize that in cooperative relationships, the participants renounce the use of deception, whereas in authority-governed relationships, the dependent partner cannot resort to deception, even in cases of conflict. "Hence, the option of influencing alter's behavior by means of deception exits only when ego construes the social relationship as symmetrical and interprets the action situation in terms of conflicting needs" (ibid, p. 148).

As shown in Table 1, Habermas correlates the justice concept based on the complementarity of order and obedience, which is built into Kohlberg's first stage of moral reasoning, with the considerations that will guide action when ones sees oneself as dependent, and tries to resolve the conflict between ego's needs and alter's demands by avoiding threatened sanctions. On the other hand, the concept of justice based on symmetry of compensation, set in Kohlberg's second moral stage, emerges only when one starts to see power as distributed equally, and may try to avail oneself of the

possibilities of deception that exist in symmetrical relations. Habermas then brings up results from Flavell's (1968) experiment in order to trace the reorganization of the preconventional stage of interaction and show how strategic action comes to be differentiated from competitive behavior.

In Flavell's experiment, two cups concealing different amounts of money are put side down on a table. Each cup bears a label in plain view indicating the payoff value supposedly hidden under the cup. The participants are shown that the relationship between the inscription and the actual amount hidden can be varied at will. Ego's task is to secretly distribute the payoffs in such a way that alter's will fail to guess where the greater amount is hidden. The point of the game is clear: alter will try to win as much as she can, and ego will try to prevent this by means of deception.

Habermas points out that if the participants in the experiment have the perspective structure of Selman's level 2 (see Table 1) they will choose what Flavell called strategy B. Following the strategy B, alter chooses the cup labeled "lower payoff," as she reasons that ego wants to fool her by not concealing the higher payoff under the cup labeled "higher payoff." On the other hand, participants who are able to engage in Selman's level 3 of perspective taking will choose Flavell's strategy C, which is a mixing strategy emerging from the recognition that alter sees through ego's strategy B. As this mutual (symmetrical) recognition establishes an infinite regress of anticipations, strategy C comes out from alter's realization that the chances of losing is as great as the chance of winning, no matter what she decides to do.

Habermas suggests that strategy C is characteristic of a type of action that is possible only at the conventional stage of interaction (Table 1), because it requires a coordination of observer and participant perspectives that is lacking at Selman's level 2, but necessary for the restructuring of preconventional competitive behavior into

strategic action. It is this shift that, according to Habermas, allows ego to attribute stability over time to alter's pattern of attributes and preferences, so that alter stops being perceived as someone whose actions are determined by shifting needs and interests and begins being viewed as a subject who intuitively follows rules of rational action. "Beyond this, however, no structural change in the sociocognitive inventory is required. In all other respects the preconventional inventory is adequate for the strategic actor" (1990, p. 150).

On the other hand, as Habermas puts it, the passage to normatively regulated action cannot be adapted so economically to the conventional stage of interaction. According to him, preconventional modes of coordinating action come under pressure in areas of behavior not dominated by competition, wherein deception is precluded. In these situations the sociocognitive inventory does require a global reconstruction to make room for a mechanism of nonstrategic coordination of action. As Habermas explains, this mechanism must be independent of both authority relations to an actual reference person and of direct links to self-interests, so that "this stage of conventional but nonstrategic action requires basic sociocognitive concepts revolving around the notion of a suprapersonal will" (ibid, p. 152). Habermas then goes on to discuss the structural breaks underlying his justification of the developmental sequence associated with the emergence of different concepts and institutions embodying the idea of a suprapersonal authority, such as loyalty to social roles and legitimacy of rules (see Table 1). Concepts and intuitions of this kind provide the elements for constituting a social world of legitimately ordered interpersonal relations and for judging actions according to whether or not they conform to or violate socially recognized norms. At the conventional level, these judgments connect in turn with the justice concepts of conformity to roles and conformity to systems of norms, as shown in Table 1.

Table 1 – Stages of interaction, social perspectives, moral stages, and value systems

Types of action (Habermas)	COGNITIVE STRUCTURES					Value systems (Graves)
	Perspective Structure	Concept of authority	Concept of motivation	Social perspective/ concept of justice	Stage of moral reasoning (Kohlberg)	
Preconventional Interaction controlled by authority ——— Cooperation based on self-interest	Reciprocal interlocking of action perspectives (Selman's Level 2)	Authority of reference persons: externally sanctioning will	Loyalty to reference persons: orientation toward reward & punishment	Egocentric / Complementarity of order and obedience ——— Egocentric / Symmetry of compensation	1. Punishment and obedience ——— 2. Naïve instrumental hedonism	3rd → 4th (egocentric) 3rd → 4th
Conventional Role behavior ——— Normatively governed interaction	Coordination of participant and observer perspectives (Selman's Level 3)	Internalized authority of supra-individual will: loyalty ——— Internalized authority of and impersonal collective will: legitimacy	Duty *versus* inclination	Primary group perspective / Conformity to roles ——— Collectivity's perspective / Conformity to existing systems of norms	3. Good boy good girl morality ——— 4. Law and order morality	4th (authoritarian) 4th → 5th
Postconventional Discourse	Integration of speaker and world perspectives (Habermas's *decentered understanding of the world* orientation)	Ideals *versus* social validity	Autonomy *versus* heteronomy	Principled perspective / Orientation toward principles of justice ——— Procedural perspective / Orientation toward procedures for justifying norms	5. Morality of democratic contract ——— 6. Morality of individual principles	5th (strategic) 6th → 7th → (consensus → ecological)

Source: Author's configuration, adapted from Habermas (1990, p. 166) and Graves (2005, p. 443).

At this point, Habermas indicates that the complex structure of perspectives underpinning normatively regulated interactions—which includes the differentiation of a formal three-world reference system (objective, social, and subjective) to which correspond three different attitudes toward the world (objectifying, norm-conformative, and expressive), and the three basic modes of language use (first, second, and third-person communicative roles of speakers and hearers)—satisfy the structural preconditions of a communicative action in which individual plans of action are coordinated by means a mechanism for reaching understanding through communication (ibid, p. 158).

According to Habermas, the third stage of interaction, i.e., discourse (Table 1), takes form only when communicative action becomes fully reflexive. At this stage, the complexity of the perspective structure undertakes a further growth in order to make room for the hypothetical attitude that characterizes the decentered understanding of the world and allows participants in argumentation to leave behind the horizon of unquestioned, intersubjectively shared, nonthematized certitudes of a quasi-natural social world in order to focus on and test validity claims that are initially raised implicitly in communicative action and are naively carried out along with it. As Habermas explains, the structural leap is marked by the synthesis of the two systems of world perspectives and speaker perspectives. "On the one hand, the system of world perspectives, which has been refracted, as it were, by the hypothetical attitude, is [now] constitutive of claims of validity that are thematized in argumentations. On the other hand, the system of fully reversible speaker perspectives is constitutive of the framework within which participants in argumentation can reach rationally motivated agreement" (p. 159).

In discourse, then, the two systems that had been fully developed at the second stage of the conventional level are put in relationship to one another. This enables participants in communicative action that want to reach a shared understanding about something in the objective, social, or subjective worlds to adopt, when necessary, an objectifying attitude to a given state of affair, a norm-conformative attitude to legitimately ordered interpersonal relations, and a expressive attitude to their own lived experiences, but also vary these attitudes in relation to each of the three words. This flexibility makes it possible to confront external nature not only in a objectifying attitude but also in a norm-conformative or an expressive one, to confront society not only in a norm-conformative attitude but also in a objectifying or an expressive one, and to confront inner nature not only in an expressive attitude but also in a objectifying or a norm-conformative one (ibid, p. 138-9). At the same time, the prior polarity involving communicative action and strategic action is overcome in discourse, as the success-orientation of competitors is assimilated into argumentation. As Habermas explains, what happens in argumentation is that "proponents and opponents engage in a competition with arguments in order to convince one another, that is, in order to reach a consensus" (ibid, p. 160). Actually, the condition that arguments are not regressively reduced to mere means of influencing each other, as often presumed along with an exclusionary instrumental conception of rationality, is what distinguishes the communicative from the strategic use of communication. "In discourse," Habermas says, "what is called the force of the better argument is wholly unforced" (ibid).

Thus, "In the light of hypothetical claims to validity the world of existing states of affairs is theorized, that is, becomes matter of theory, and the world of legitimately ordered relations is moralized, that is, becomes a matter of morality" (ibid, p. 161). This "moralization of society" undermines the normative power of the factual, so that

institutions that have lost their quasi-natural character can be turned into "so many instances of problematic justice" (ibid). A new reorganization of the fundamental sociocognitive concepts available at the stage of role behavior and normatively governed interaction becomes necessary in order to rationally justify the "uprooted and now free-flowing systems of norms" (ibid). At the postconventional level, norms of action are subordinated to principles, or higher-order norms. "The notion of the legitimacy of norms of action is now divided into the components of mere de facto recognitions and worthiness to be recognized" (ibid). Correspondingly, a parallel differentiation occurs in the concept of duty, where "the respect for the law is no longer considered an ethical motive per se" (ibid). To dependence on existing norms, is opposed "the demand that the agent make the validity rather than the social currency of a norm the determining ground of his action" (ibid). That is, to heteronomy it is opposed autonomy (Table 1).

In short, Habermas claims that a strict (cognitivist, universalist, formalist) concept of morality can evolve only at postconventional stage, for "only at the postconventional stage is the social world uncoupled from the stream of cultural givens" (ibid, p. 162). To be sure, it is precisely the sight of plural relativism, which comes into view at the postconventional stage, which makes the autonomous justification of morality an unavoidable problem (ibid).

Now, if Habermas's action-theoretic account of development of the sought-after moral point of view admittedly requires distinctions not easy to operationalize, the difficulty to understand how the conceptions of justice emerge from the sociocognitive inventory of the corresponding stages of interaction can be facilitated by a key insight. This insight, which Habermas properly attributes to Durkheim, is that there is no specific socialization process through which agents acquire moral dispositions. As

Habermas puts it, "In trying to understand this process, one has to take into account that the normatively regulated fabric of social relations is moral in and of itself, as Durkheim has shown" (ibid, p. 164, emphasis in the original). In Heath's (2001) words, "This means that acquiring the competences required to manage routine social interactions amounts to acquiring the dispositions and personality structures that we understand to be essential elements of moral agency" (p. 8).

As Habermas plainly recognizes, a hypothetical reconstruction of the type sketched above can serve at best as a guide for further research. With this intent we bring Clare Graves's emergent-stage conception of adult personality systems development in Table 1. The connecting points were provided by Graves`s own correlation analysis involving the stages advanced in his conception and those described in Kohlberg's model (cf. Graves, 2005, p. 443).

Without going on describing the substance of these correlations, the next section is aimed only at illustrating the implications of our suggested integration of the cognitive-developmental account of moral agency into the analysis of institutional change and development in the context of social dilemmas

BEYOND FOUNDATIONALISM: INTEGRATING DISCOURSE ETHICS AND STAGES OF MORAL AGENCY WITH BLOOMINGTON SCHOOL

As discussed previously, insofar as it remains locked in epistemological foundationalism the cognitivist version of institutionalism does not account for the rational basis of moral agency. The interpretation of the role of communication remains confined to an objectifying attitude aimed at obtaining useful information about a given state of affair. In turn, both the norm-conformative and expressive attitudes of the participants are reduced to utilitarian calculus mediated by strategic reasoning alone.

In Figure 2, we simply substitute the exclusively instrumental view of the role of communication for this more general conception. We include a pictographic representation of Clare Graves's model (1970) as the Spiral Dynamics to further integrate the developmental psychology perspective into the image. The small spiral superposed on Habermas's typology (in the upper left corner in Figure 5) represents Habermas's own developmental account of the human capacity to coordinate interaction using different standards of choice (Table 1), and according to which communicative action emerges only in later (postconventional) stages of interaction.

According to the summarization presented in the previous section, one can reasonably hypothesize that, in face of the conflict of interests that characterizes the social dilemmas, individuals centered at the preconventional stages interaction—third stage in Graves`s model (Table 1)—will use communication as an opportunity to trick others, quite in line with Hardin`s account and the standard game-theoretic prediction. On the other hand, individuals centered at the conventional stages are expected to use the communication to reinforce conformity to roles and existing systems of norms, whereas those centered at the discourse stage (postconventional mode of coordinating action) are expected to use communication as an opportunity to discuss alternative governance rules and to commit themselves to pursue their goals only on the condition of an agreement.[4] The presence of these individuals is probably what explains the regular findings attesting the effect of communication in enhancing cooperation in both laboratory and field experiments.

---

[4] This spiral also reminds us of Graves's own research on perceptual readiness and validates his point of view (Graves; Huntley; LaBier, 1965). According to the study results, the attention mechanism that Frohlich and Oppenheimer (2001) identify as driving the selection of information and the perception of the action situation are similarly affected by the dynamics of inner growth; both the vividness and the salience of the perceived objects depend on the developmental stage of the observer.

In the center of the Figure 2, the spiral suggests that the internal drives of human development impose restrictions on how the individuals are able to revise their "mental models" of the action situation. On the one hand, centralizations in animistic (2nd stage), egocentric (3rd) or authoritarian (4th) modes of thinking and interacting may explain the "persistence throughout history of dogmas, myths, superstitions, and ideologies" that baffles the proponents of the cognitive institutionalism. On the other hand, the flexibility which makes possible for the individuals at the discourse stage of interaction to combine the basic attitudes corresponding to the objective, the social, and the subjective worlds, broadens the interpretation of the "mental models" and the revision process illustrated in the Figure 2. One can now think not only of revisions anchored in an objectifying attitude in face of the action situation but also of revision of norm-conformative and expressive attitudes toward it.

The spiral next to the "culture" box stands for both the phylogenetic aspect of Habermas`s historical reconstruction of the development of the modes of interaction in
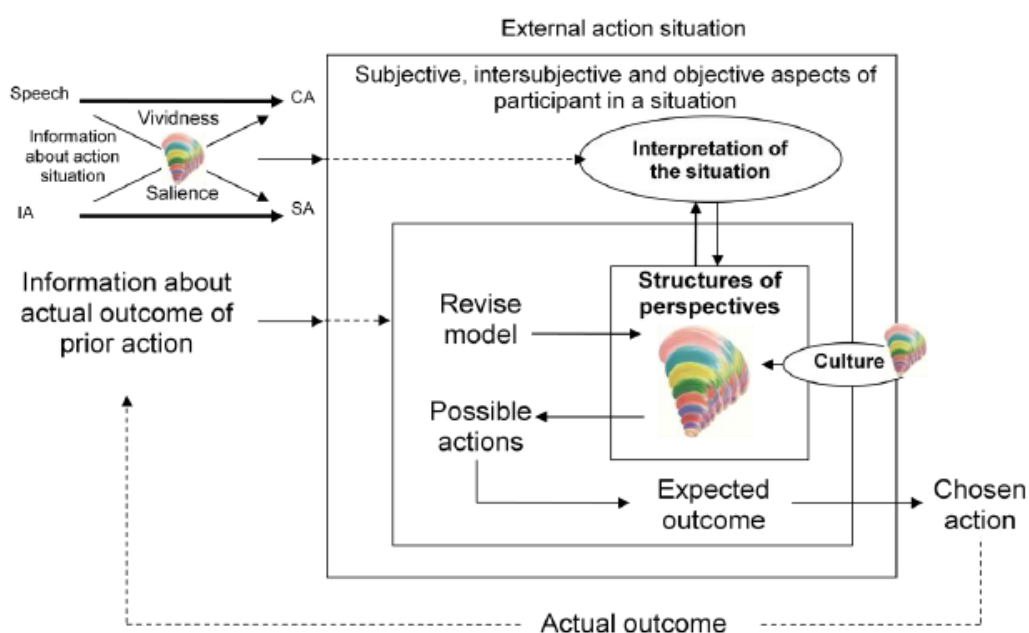

Figure 2 - Internal drives and stages affect both the revision of perspectives regarding the action situation and the participants' attitudes toward communication opportunities.
Source: Authors' configuration, adapted from Ostrom (2005)

24

human societies and the collective aspect of the ontological recapitulations in the development of the individuals. The relevant aspect to discern is that while the culture certainly influences and often restricts the individual movement up spiral of interior growth it is not the sole determinant of it. As Wilber (2000) emphasizes, the multiple inner structures configuring human interiority have their own stages of growth and development and are not changed from the outside. The analytical result is that information about cultural values and beliefs systems are not a substitute for individual information.

In summary, the process through which individuals revise their perceptions of an action situation and then chose which action to seek a particular outcome (including mutual understanding and normative commitments) can allow for the existence of a structured interiority. New limitations and potential to change both perceptions and actions arise from the recognition of the laws of inner transformation. The importance of acknowledging multiple dispositions and forms of rationality in analyzing social dilemmas increases with the recognition that open-ended, multi-stream, complex interior growth is nonetheless a process that involves a continuing decline in egocentrism, increasing autonomy and the increasing ability to take other people, places, and things into account when making decisions that affect the well-being of others (cf. Wilber, 2000, 2001).

CONCLUDING REMARKS: ON FUTURE RESEARCH AND THE ROLE OF EDUCATION

That collective action problems involve a fundamentally moral dilemma is manifest in the opening pages of *The Tragedy of the Commons*, which situate the dilemma within a class of problems whose solution will require a change in human values and ideas of morality. Still, Hardin's standpoint on moral judgments is purely

relativist. He refers to Fletcher's (1966) Situation Ethics to reveal what he regards "a not generally recognized principle of morality, namely: the morality of an act is a function of the state of the system at the time it is performed" (p. 1246).

From this perspective, it appears that the role of education is "to reveal to all" the need to abandon the belief in the freedom of the commons as long as the "state of the system" requires it. Once this necessity is recognized, "mutual coercion mutually agreed upon" is Hardin's proposed solution to the problem.

We see no issue with drawing the rationality of a moral choice from the recognition of a contingent necessity. Nevertheless, we believe that the analysis presented here expands the range of alternatives and suggests additional directions for education and research on social dilemmas. Insofar as both communicative rationality and genuine care for others indicate interior dispositions that emerge later in the path of human development, future research should verify the existence of these traits in social dilemmas using empirical testing. If there is evidence indicating that groups of individuals centered at further stages of interior development are better able to cope with their social dilemmas successfully, then the role of education could include the creation of favorable conditions for humanity's progress up the spiral of interior growth. These conditions would eventually eliminate the need to agree upon the necessity of mutual coercion.

REFERENCES

Aligica, P.D.; Boettke, P.J. (2009). *Challenging institutional analysis and development: the Bloomington School.* London and New York: Routledge.

Arlin, P.K. (1975). Cognitive development in adulthood: a fifth stage? *Developmental Psychology* 11, 602-606

Bateson, G. (1972). *Steps to an ecology of mind*. New York: Ballantine, 1972.
_____ (1979). *Mind and nature: a necessary unity*. New York: Dutton, 1979.

Basseches, M. (1980). Dialectical schemata: a framework for the empirical study of the development of dialectical thinking. *Human Development*, 23, 400-421.

Cárdenas, J.C. (2000). Thinking globally and getting others to act locally: polycentricity and the conservation of biodiversity. Paper presented... *Workshop in Political Theory and Policy Analysis*, Bloomington: Indiana University.

Cárdenas J.C., Ostrom E (2004) *Agr Syst* 82:307–326.

Carolan, M. S. (2004). Ontological politics: Mapping a complex environmental problem. *Environmental Values*, 13, 497-522.

Casari M, Plott CR (2003). *Journal of Economic Behavior and Organizations*. 51:217–47.

Commons, M.L. (1981). A comparison and synthesis of Kohlberg's cognitive-developmental and Gewirtz's learning-developmental attachment theories. In: GEWIRTZ, J.L.; KURTINES, W.M. (Eds.). Intersections with attachment. Hillsdale, NJ: Erlbaum, pp. 257-291.

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

Denzau, A.T.; North, D.C. (2000). Shared mental models: ideologies and institutions. In: Lupia, A.; McCubbins, M.D.; Popkin, S.L. (Eds.). *Elements of reason – cognition, choice, and the bounds of rationality*. 3.ed. Readings: Cambridge University Press, p. 23-46.

Flavell, J.H., et al (1968). *The Development of Role-Taking and Communication Skills in Children*. New York.

Fletcher, J. (1966). Situation Ethics: Westminster, Philadelphia.

Forsyth, T. (2001). Critical realism and political ecology. In A. Stainer & G. Lopez (Eds.), *After postmodernism: An introduction to critical realism*. London: Athlone Press.

Frohlich, N.; Oppenheimer, J.A. (2001). *Choosing: a cognitive model of economic and political choice*. Winnipeg: University of Manitoba, Faculty of Management, 2001. (Working Paper).

Graves, C.W. (1970). Levels of existence: an open system theory of values. *Journal of Humanistic Psychology*, v. 10, n. 2, p. 131-155, 1970.

_____ (1971). *A systems conception of personality*: remarks by Clare W. Graves on his levels of existence theory. Santa Barbara, CA: ECLET Publishing (Presented at the Washington School of Psychiatry, Washington, DC).

_____ (1981). *Summary statement*: the emergent cyclical, double-helix model of the adult human biopychosocial systems. Boston, 1981.

_____ (2005). "The Never Ending Quest: A Treatise on an Emergent Cyclical Conception of Adult Behavioral Systems and Their Development." Edited and compiled by Christopher C. Cowan and Natasha Todorovic. Santa Barbara, CA: ECLET Publishing.

Graves, W. C. Huntley,W.C.; LaBier, D. (1965). *Personality Structure and Perceptual Readiness: An Investigation of their Relationship to Hypothesized Levels of Human Existence*, Union College. Retrieved May 12, 2013, from the Historical Collection of the work of Dr. Clare W. Graves at http://www.clarewgraves.com/articles_content/1965_GHL/1965_GHL1.html

Habermas, J. (1984). The *theory of communicative action: reason and rationalization of society*. Boston, MA: Beacon Press. v. 1.

Habermas, J. (1987). *The theory of communicative action: the critique of functionalist reason*. Boston, MA: Beacon Press. v. 2.

Habermas, J. (1990). *Moral consciousness and communicative action*. Cambridge, MA: MIT Press.

_____ (1993). *Justification and application: remarks on discourse ethics*. Cambridge, MA: MIT Press.

Hardin, G. (1968) The tragedy of the commons. *Science*, v. 162, p. 1243-1248, 1968.

Harsanyi, J.C.; Selten, R. (1988). *A general theory of equilibrium selection in games*. Cambridge: MIT Press, 1988.

Heath, J. (2001). *Communicative action and rational choice*. Cambridge, MA: MIT Press, 2001.

Hobbes, T. (1991). *Leviathan* [1951]. Cambridge: Cambridge University Press, 1991.

Hurwicz, L. (2008) *But who will guard the guardians?* (Nobel Prize Lecture). *American Economic Review*, 98:3. 577-585. Retrieved May 12, 2013, from http://www.nobelprize.org/nobel_prizes/economics/laureates/2007/hurwicz_lecture.pdf

Inhelder, B.; Piaget, J. (1955). De la logique de l'enfant à la logique de l'adolescent. Paris: Presses Universitaires de France. (English version: The growth of logical thinking from childhood to adolescence. London: Routledge, 1958)

Law, J. & Urry, J. (2003). *Enacting the Social*, published by the Department of Sociology and the Centre for Science Studies, Lancaster University, Lancaster LA1 4YN, UK, Retrieved May 12, 2013, from http://www.comp.lancs.ac.uk/sociology/papers/Law-Urry-Enacting-the-Social.pdf

Kramer, D.A. (1983). Postformal operations? A need for further conceptualization. *Human Development*, 26, 91-105.

Mantzavinos, C.; North, D.C.; Shariq, S. (2004) Learning, institutions, and economic performance. *Perspectives on Politics*, v. 2, n. 1, 2004.

Marchand, H. (2005) Some reflections on postFormal thought. *The Genetic Epistemologist*, v. 29, n. 3.

Morin, E. (1999). *Seven Complex Lessons in Education for the Future* [Original title: Les sept savoirs necessaires a I'education du future] Translated by Nidra Poller. Unesco, EDP-99/W.

Nash, J.F. (1950). Equilibrium points of n-person games. *Proceedings of the National Academy of Science*, v. 36, p. 48-49.

Nash, J.F. (1951). Non-cooperative games. *Annals of Mathematics*, v. 54, p. 286-295.

Ostrom, E. (1990). *Governing the commons*. Cambridge: Cambridge University Press.

_____ (2005). *Understanding institutional diversity*. Princeton, NJ: Princeton University Press.

_____ (2007). A diagnostic approach for going beyond panaceas. *PNAS,* v. 104, n. 39, p. 15181-15187.

_____ (2009). Complexity and sustainability of social-ecological systems. *Workshop in Political Theory and Policy Analysis*, Proceedings... Bloomington: Indiana University. (Working Paper, 9).

Ostrom, E.; Gardner, R.; Walker, J.M. (1994). *Rules, games, and common-pool resources*. Ann Arbor: The University of Michigan Press.

Ostrom, V. (1980). Artisanship and artifact. *Public Administration Review*, v. 40, n. 4, p. 309-317.

_____ (1982). The human condition, workshop archives. Working Paper n° 16. *Workshop in Political Theory and Policy Analysis*, Bloomington, IN. Indiana University..

_____ (1993). The Place of Language in the Political Economy of Life in Human Societies. Working Paper n^o W93-6, *Workshop in Political Theory and Policy Analysis*, Bloomington, IN. Indiana University.

Pascual-Leone, J. (1984). Attentional, dialectical, and mental effort: toward an organismic theory of life stages. In M. L. Commons, F. A. Richards & C. Armon (Eds.), *Beyond formal operations* (pp. 182-216). New York: Praeger.

Piaget, J. (1970/1972). Intellectual evolution from adolescence to adulthood. Human Development, v. 15, p. 1-12.

Richards, F.; Commons, M. (1990). Postformal cognitive-developmental theory and research: A review of its currents status. In: ALEXANDER, C.; LANGER, E. (Eds.). Higher stages of human development. New York: Oxford University Press, 1990. p. 139-161.

Riegel, K. (1973). Dialectic operations. The final period of cognitive development. *Human Development*, 16, 346-370.

Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*, New York: Cambridge University Press.

Sinnott, J. D. (1984). Postformal reasoning: The relativistic stage. In M. L. Commons, F. A. Richards & C. Armon (Eds.), *Beyond formal operations* (pp. 298-325). New York: Praeger.

_____ (1989) Adult differences in use of postformal operations. In M. L. Commons, J. D. Sinnott, F. A. Richards & C. Armon (Eds.), *Beyond formal operations* II (pp. 239-278). Westport: Praeger.

Tarnas, R. (1999). *A epopéia do pensamento ocidental*. Rio de Janeiro: Bertrand Brasil.

Wilber, K. (2000). *Integral psychology: consciousness, spirit, psychology, therapy*. Boston, Mass. Shambhala.

_____ (2001). *A Theory of Everything*. Boston , Mass: Shambhala.