

Willingness to Cooperate and Stages of Moral Reasoning: Evidences from Common-Pool Resource Experiments with “Nonbinding” Communication and Sanctioning Conditions

Leandro F. F. Meyer*

*Socio-Environmental Institute, Federal Rural University of the Amazon
Belém-PA, Brazil*

Marcelo J. Braga

*Department of Agricultural Economics, Federal University of Viçosa
Viçosa-MG, Brazil*

Abstract

As economists increasingly recognize the limits of the canonical assumption of self-interest, the lack of a theory of human valuation that clearly specifies how individuals reach utility judgments renders the prediction of rational action in intersubjective, morally relevant conflicts of action virtually impossible. Resting on fundamental assumptions about the cognitive content of the moral judgment, we examine the explanatory power derived from a structuralist-constructivist theory of adult development which presents real analytical significance for understanding behavioral diversity in situations where the individual and the collective interests collide. Experimental results suggest that the theoretical constructs built in the selected model provide reliable basis for predicting participants' behavior in a common-pool resource dilemma under diverse institutional conditions, including communication and sanctioning mechanism. These are selected results from a broader experimental research in which the same developmental model proved useful for elucidating the interplay among value judgments, motivations, internationalities, and decisions in both public-goods and ultimatum game experiments.

JEL classification: C72; C92; D74

Keywords: social dilemmas; experimental economics; sociocognitive and moral reasoning; adult development

*Corresponding author

Leandro Frederico Ferraz Meyer
Universidade Federal Rural da Amazônia
Avenida Presidente Tancredo Neves, Nº 2501
66.077-530 Caixa Postal: 917
Phone: (55)(91) 3210-5140 / 3272-7228
webmaster@ufra.edu.br
leandro.meyer@ufra.edu.br / lffmeyer@terra.com.br

1. Introduction

Although a fully articulated, general theory of how institutions affect the incentives confronting individuals and their resultant behavior in collective-action situations does not yet exist, there is today among social and political scientists interested in social dilemmas a growing consensus that the conventional economic theory of externalities conveys a special case of a more general theoretical structure (Ostrom, 2007; Frohlich & Oppenheimer, 2001).

Clearly, as ordinarily understood, the essence of social dilemmas such as in decisions involving the appropriation from a common-pool resource or the provision of a public good is inseparable from the existence of *market* externalities and the resulting payoff structure, which embroils the individuals in a situation where the rationality that is in the individuals' interest is not favorable from the group perspective, and vice versa. Without questioning the market reference for defining externalities and the generalisability of the economic postulate of rationality in this context, the conventional theory presents the individuals involved in dilemmas of this type as being trapped in the "inherent logic" of the situation (Hardin, 1968). Accordingly, they are said to face a "social dilemma" in that they would all be better off if they found a way of cooperating together, but no one acting alone is assumed to have an incentive to bear the costs of such cooperation (Ostrom, 2007).¹

As a result, the conventional theory repeatedly advocates that institutions aimed at preventing the "tragedy of the commons" should, first, address the fundamental problem about the specification of the property rights—either in the form of private property or government ownership and control—and, second, that such regulations have to be *imposed* from external authorities, which are assumed to act in the public interest and have enough information to devise suitable institutions so as to induce socially optimal behavior (Ostrom, 1990; 2007).

¹ This is basically because a cooperating rule that could solve the dilemma has the character of a *public good*, meaning that anyone included in the community benefits from that rule, whether they contribute for its provision or not. Clearly, under the canonical assumption of self-interest, the provision of such rule configures a second-level, same-type dilemma, placed on-top of the initial one. Hence, it is not consistent with the conventional theory that the same "helpless" participants, trapped by the inherent logic of the commons, solve a second-level dilemma in order to address the first-level dilemma under analysis.

However, among the lessons scholars have learned from the vast empirical research of the last several decades is the recognition that problems of overharvesting and misuse of ecological systems are rarely due to a single cause (Ostrom, 2007). Both field and laboratory research on social dilemmas have shown that individuals' behaviors in these situations are affected not only by the structural characteristics of the outcomes (high or low, certain or probable, et cetera), and the structural characteristics of the group of people involved (large or small group, with or without communication, with or without a leader, et cetera), but also by the specific content or context of the interaction situation (investment decisions, social events, environmental issues, et cetera) (Kollock, 1998; Komorita & Parks, 1995; Kopelman et al., 2002; Lepyard, 1995; Van Lange, Liebrand, Messick, & Wilke, 1992; Ostrom, Gardner & Walker, 1994; Poppe, 2005).

In view of that, much of research on social dilemmas has been focusing on identifying sets of variables that enhance the likelihood of users organizing themselves to avoid the social losses associated with the basic conflict between the individual and the collective rationalities or interests, in face of the externalities (Ostrom 1990; Schlager 1990; McKean 1992; Tang 1994; Ostrom et al. 1994; Wade 1996; Baland & Platteau 1996; Agrawal 2001). Yet, the puzzle is that these structural and contextual variables interact also with the characteristics of the *individuals* involved (economic, sociocultural, psychological, age, et cetera), so that different individuals may respond differently to objectively *similar* incentive structures and contexts of action. Thus, for example, even considering a given decision structure in a specific context (e.g. "costal fishery"), different individuals often have different attitudes towards existing information, perceived uncertainty and risk, as well as towards communication opportunities, and the influence of authorities and leaders.

In particular, the distinctively *moral* nature of the choices in situations where the individual and the collective interests collide makes it important to understand how individuals reach utility or value judgments when addressing social dilemmas.² Basically, the

² As Heath (2008) indicates, while there are many aspects of morality that are puzzling, perhaps the most puzzling is that it often requires us to act in ways that are contrary to our self-interest. "We may find ourselves

moral element of the dilemma renders simply inadequate to keep the *aprioristic* notion that all preferences are self-regarding, as in the standard theory. As Ostrom (2005) explains, while keeping the standard assumption might be a reasonable scholarly strategy for modeling behavior in highly competitive market settings, it is not so for addressing most social dilemmas because this type of situation often evoke *internal values* for participants that are *not* monotonically related to the objective payoff involved (see also, e.g. Axelrod, 1997; Gintis, 2000; Camerer, 2003).

In addition, once the fact that intrinsic values matter when addressing social dilemmas is recognized, one is supposed to acknowledge also that the situation is one of incomplete, rather than, complete information, because the agents cannot know exactly how the others are valuing the alternative actions and outcomes affecting the well-being of each other (Ostrom, 2005). Finally, the uncertainty regarding others' valuations is hardly amenable to risk analysis, so that one should admit, as well, that observed behavior usually proceeds from discrete rules of thumb, heuristics, or normative orientations, rather than, maximizing calculations of the expected utility (cf. Heiner, 1983).

This is possibly why Ostrom (2005) suggests that the major theoretical challenge facing scholars interested in social dilemmas today is developing an appropriate family of assumptions to make about the intrinsic values individuals place on actions and outcomes—particularly outcomes obtained by others. According to her, “Without further progress in developing our theories and models of human valuation in social dilemma situations, those convinced that all human behavior can be explained using rational egoist models will continue to recommend Leviathan-like remedies for overcoming all social dilemmas.”

In this study, we suggest that the theories and findings pertaining to the structuralist-constructivist conception of human development are of real relevance for advancing our understanding of human valuation, and then for devising better institutions addressing social dilemmas. Actually, a central tenet in developmental psychology is that, in order to produce

wanting something, but feeling that morality prohibits us from doing what is necessary to obtain it. Morality therefore presents itself to us in the form of a duty to refrain from the pursuit of individual advantage, or to use the more technical term, in the form of a *deontic constraint*.”

the expected results, the incentive structure should be tuned to the characteristics of each psychosocial centralization stage due to unique motivational needs, aims, and means that differ within each stage. Further, the substantive significance of developmental framework for addressing social dilemmas is accentuated by a growing consensus among developmental psychologists around the understanding that (i) *psychological development is not upper-bounded*, i.e., it is not limited to the childhood and adolescence, as traditionally assumed, and (ii) open-ended, multi-stream, complex interior growth is a process involving, for the most part, a *continuing decline in egocentrism, increasing autonomy and increasing ability of taking other people, places, and things into account* when making decisions that affect the well-being of others (cf. Wilber, 2000, 2001).

Despite the clear indications drawn in these convergent findings and central tenets of developmental psychology, the serious investigation of the implications of the developmental view on morals for addressing the social dilemmas has been hindered by the widely accepted idea that “value judgments” or “moral questions” have no cognitive content (a position known as *moral noncognitivism*), and are therefore rationally undecidable. As Heath (2001) explains, such hindrance is epistemological in nature, as 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*” (p. 2, see also Habermas, 1993 and 2003). As an ordinary source of *moral relativism* and *skepticism*, the foundationalist epistemology has definite implications in determining the way we are supposed to explain rational choice in morally relevant conflicts of action, and in particular the role of communication in producing normative commitment in those situations. Basically, moral argumentation is assumed to be incapable of producing any greater level of agreement than that which agents already bring with them. As a result, linguistic communication is reduced to a background for strategic interactions, providing agents with the information against which they determine maximizing strategies—such as common knowledge of preferences, action alternatives and action-outcome linkages, other participants’ reputation as cooperators, expectations of sanctions, and so on—without

exercising any constraint on the range of action alternative available to them (Heath, 2001, p. 25; see also Habermas 1984/1987 and 1990).

Clearly, although strategic interaction does not itself presuppose or depend upon any sort of moral noncognitivism (Heath, 2001), this “nonbinding” character of *instrumental* communication is what explains why, in sticking with the canonical assumption of self-interest, much of contemporary, noncooperative game theory treats Nash’s (1950, 1951) condition of no communication for distinguishing between cooperative and noncooperative games as superfluous, as Ostrom *et al* (1994) have observed. But even if we admit that preferences are not homogenously self-regarding, the tacit assumption that value judgments are not rationally justifiable still obscures, and mystifies, in fact, the understanding of the role of communication in fostering normative commitment in social dilemmas.

Having developed this epistemological discussion in another place (Meyer, 2009), this research is meant as a step towards assessing the empirical robustness of alternative theories and models of adult development for explaining and predicting behavior in morally relevant conflicts of action. Rooted in the constructs of a selected developmental theory which is particularly suitable for examining behavior in situations where the individuals and the collective interests collide, we began by assessing the psychosocial profiles of 322 Brazilian, potential participants in (1) a laboratory common-pool resource appropriation dilemma, including “nonbinding” communication and sanctioning conditions, (2) a stepwise public-goods provision dilemma with variable levels of required contribution, and (3) a equity-risk dilemma as represented in the standard Ultimatum game. In order to pretest the cross-cultural robustness of the theoretical constructs set in the chosen developmental model (Graves, 1970) we carried out a factor analysis on the survey data. The procedure resulted in three quite meaningful principal components standing for three principal psychosocial centralization stages present in our sample. Participants’ behaviors in the different experimental situations were then analyzed to find out whether they agree with a set of theoretical expectations derived from the selected theory, as well as with the general features of human interior development, mentioned previously. The experimental results

thus far suggest that the theoretical constructs built in the chosen developmental model provide reliable basis for predicting behavior in the situations we have examined, and that greater willingness to cooperate is indeed associated to higher stages of psychosocial development, as described in the model.

In this paper we present selected results from the common-pool resource (CPR) experiment. Results from the public-goods experiment can be found in Meyer and Braga (2009). The complete set of results is described in Meyer (2006). We begin with a brief outline of the chosen theory: namely, Clare Graves's Emergent-Cyclical Levels of Existence Theory (ECLET) of adult personality systems development and cultural institutions (Graves, 1970, 2005). We then summarize the method and the procedures, accompanied by the specification of the theoretical expectations associated to the constructs postulated in Graves's model in the different experimental conditions of the CPR experiment. The results are presented and discussed in the sequence. In the concluding section we point to the policy oriented implication of our findings.

2. The *biopsychosocial* waves of agency and communion: outline of Graves's theory

Rather than purely a psychological study, the developmental theory we examine here postulates that the *biopsychosocial* development of human beings arises from the interaction of a double-helix complex of two sets of determining forces: the environmental social determinants, and the neuropsychological equipment of the organism for living. Out of about a decade of careful empirical research, Graves (1970) conceptualized eight emergent stages or waves of interior growth which provide a description of states of biopsychosocial equilibrium, comprising a perception of the environment, a reciprocal state of neurochemical balance, reflected in a social construction that then influences those mental states of equilibrium, as part of the environment perceived. In Graves's words:

The psychology of the adult human being is an unfolding, ever-emergent process marked by subordination of older behavior systems to newer, higher order systems. The mature person tends to change his psychology continuously as the conditions of his existence change. Each

successive stage or level of existence is a state through which people may pass on the way to other states of equilibrium. When a person is centralized in one of the states of equilibrium, he has a psychology which is particular to that state. His emotions, ethics and values, biochemistry, state of neurological activation, learning-systems, preference for education, management and psychotherapy are all appropriate to that state. If he were centralized in some other state he would think, feel and be motivated in manners appropriate to that state. He would have biochemical characteristics and state of neurological activation particular to it. When in a certain state, he would have opened only certain systems for coping and learning. Thus, he would respond most positively to education, management, and therapy which are congruent with that state. And he would have to respond negatively to forms of education, management and therapy not appropriate to the state of his centralization (Graves 2005, p.29-30).

Acknowledged, Graves's sweeping statement is subject to the currently accepted understanding that most of the multiple lines or streams of consciousness comprising human interiority (e.g. cognitive, interpersonal, psychosexual, emotional, moral, et cetera) configure partially decomposable subsystems, so that they develop in a relatively independent fashion (Wilber, 2001, p. 44). As a result, a person can be very advanced in some lines, medium in others, and low in still other—all at the same time. Hence, it is not quite appropriate to talk about general "levels of existence," as no sequential development can possibly be devised when considering the sum total of all these different lines. However, as Wilber (2001, p. 28) reports, "the bulk of research has continued to find that each developmental line itself tends to unfold in a sequential, holarchical fashion," meaning (i) that higher stages in each line tend to build upon or incorporate the earlier stages, (ii) that no stage can be skipped, and (iii) that the stages emerge in an order that cannot be altered by environmental conditioning or social reinforcement.

Keeping the mentioned caveat in mind, we point that the special significance of Graves's theory for advancing our understanding of the interplay among cognition, valuations, and institutions, and how they affect behavior and the likelihood of self-organization in collective-action settings derives from his focus and methodology, and is rooted in the very structure of the model. First, Graves's constructs describe a sequence of emergent worldviews and value-based behavior systems resulting from research with adult

subjects, ranging from 18 to 61 years of age, both male and female (Graves, 1971, p. 8). Graves's focus on adult thinking and *argumentation* revolving on conceptions of health, and mature personality led him to pay considerable attention to both psychodynamic aspects and environmental factors that may either arrest or promote the development of increasingly integrated value based behavioral systems. As part of the conceptualization of the developmental order puts forth in his model, Graves examined how change in publicly stated values and conceptions of mature personality occurs by observing whether his experimental subjects stick with a defense of their original standpoints or revise them in face of both authority and peer argumentation on behalf of different values and conceptions. Clearly, we claim that Graves's method renders his model particularly suited for examining role of communication in fostering cooperative agreements in collective-action settings.

The substance of Graves's constructs resides precisely on revealing the different set of values individuals may place on actions and outcomes affecting the well-being of others. In this regard, though each behavioral pattern associated to the series of values systems described by Graves must be viewed with a different premise, out of their own specific aims and means, Graves's model puts forward that people tend to oscillate back and forth between two fundamental stances, much like the relative position of a pendulum in its arc between "me" (agency) and "we" (communion) orientations (Cowan and Todorovic, 2005). According to Graves's model, this cyclical turn involving the agency and the communion capacities of the self produces two basic families of behavioral systems: namely, *express-self* systems and *sacrifice-self* systems (Table 1), which have manifest implication for the analysis of situation where the individual and the collective interests collide.

Further, the postulated nested hierarchical or holarchical organization unfolding along the developmental path means that interior awakening brings about new, emergent capacities marked by broader perspectives, wherein higher-order structures gradually replace the lower ones while preserving them in a reorganized form. Putting this structural feature together with the cyclical turn between agency and communion the result is a sequence showing *decreasing egocentrism* and *increasing behavioral freedom*.

Table 1. Cyclical aspect, way of thinking and themes of the selected Gravesian stages or waves of interior development

Stage or wave	Cyclical aspect	Thinking	Basic theme
8 th	Sacrifice-self (communion)	Holistic	<i>Adjust to the realities of one's existence and accept the existential dichotomies as they are and go on living.</i>
7 th	Express-self (agency)	Ecological	<i>Express self for what self desires, but never at the expenses of others and in a manner that all life, not just my life, will profit.</i>
6 th	Sacrifice-self (communion)	Consensus	<i>Sacrifice now in order for all to get now.</i>
5 th	Express-self (agency)	Strategic	<i>Express self for what self desires, but in a fashion calculated not to bring down the wrath of others.</i>
4 th	Sacrifice-self (communion)	Authority	<i>Sacrifice self now to receive reward later.</i>
3 rd	Express-self (agency)	Egocentric	<i>Express self, to hell with others and the consequences, lest one suffer the torment of unbearable shame.</i>
2 nd	Sacrifice-self (communion)	Animistic	<i>Sacrifice self to the way of your elders.</i>
1 st	Express-self (agency)	Instinctive	<i>Express self as just another animal according to the dictates of one's psychological needs and the environmental possibilities.</i>

Source: Author's configuration based on Graves (2005) and Beck and Cowan (1996)

As indicated, these features do *not* convey eccentric attributes of Graves's model. Rather, they express what is possible the most marked convergent characteristic in the field of developmental psychology (Wilber, 2000, 2001). In Graves's scheme, that egocentrism is decreasing is noticeable by comparing the basic themes corresponding to the 3rd, 5th, and 7th stages (Table 1). It should be noticed also that the whole scheme implies a widening of the moral embrace, i.e., of those who are considered worth of moral concern.

Of course, the different worldviews and value systems brought about along the path of psychosocial development can underpin aggressive behavior as well. Hence, insofar as we can rely on Graves's orienting generalizations vis-à-vis the average features of different societies, cultures and subcultures, the theory allow us to distinguish different motivations and justifications underlying both individuals and nations engagement in wars and conflicts (Table 2). Again, the puzzling diversity of worldviews, cultural values, and forms of life we

observe can be accounted for, in part, by the partial decomposability of the multiple lines of interior development, and, in part, by the holarchical organization of the behavioral patterns displayed within emergent levels of increasing complexity and awareness.

As Beck (1999) explains, since the previously awakened levels do not disappear, but stay active within the worldview stacks, they continue impacting the nature and form of the more complex systems. Thus, as he says, “many of the same issues we confront on the West Bank can be found in South Central Los Angeles. One can experience the animistic worldview on Bourbon Street as well as in Zaire. Matters brought before city council in Minneapolis are not unlike the debates in front of governing bodies in the Netherlands. Countries and cultures are mosaics of multiple [value] MEME codes.”

Similarly, this structural foundation can explain both the limits and the potentials of human adaptability in face of diverse incentive structures and contexts of action. On the one hand, behavioral freedom and autonomy increase as new capacities are added to the previous ones along the developmental path. On the other hand, each stage of psychosocial centralization brings about certain behavioral patterns that represent preferred ways of coping with the certain problems. Thus, contrary to what the terms stages or levels might suggest, it is important to keep in mind that stages of development are *not* rigid levels but flowing waves, which much overlap and interwaving, resulting in a meshwork or dynamic spiral of consciousness unfolding (Wilber, 2001; see also Beck and Cowan, 1996).

In spite to the subtleness and caveats involving Graves’s conception, his model has evident implications for the institutional analysis of social dilemmas. While the proposition specific behavioral hypotheses must wait until we have presented the experimental conditions, it is clear that we should expect more cooperative dispositions on the part of individuals centralized at some of the sacrifice-self systems (2nd, 4th, 6th, and 8th), or at the 7th express-self system. In contrast, it seems similarly apparent that we can expect opportunistic behavior on the part of the individuals centralized at the 1st, 3rd and the 5th stages/waves, in situations where the individual and the collective interest collide.

Table 2. Political form and deepest motivation and justification for aggressive behavior

Stage or wave	Political Form	Deepest motivation and "bottom line" justification for aggressive behavior
6 th	Value Communities	to punish those who commit "crimes against humanity" and protect the victims.
5 th	Corporate States	to advance economic spheres of influence, or access to raw materials and markets.
4 th	Ancient Nations	to protect borders, homelands, hearth, preserve way of life, defend "holy" cause.
3 rd	Feudal Empires	to dominate, gain the spoils, and earn the right to rape, pillage, and plunder.
2 nd	Ethnic Tribes	to protect the myths, ancestral traditions, rights of kinship, and sacred places.
1 st	Survival Clans	to keep a place in the survival niche, as in the movie <i>The Quest for Fire</i> .

Source: Adapted from Beck (1999)

3. Moral judgment, stages of interaction and attitudes toward communication

As suggested, the foundationalist epistemology creates a critical hindrance for acknowledging cognitivist approaches in moral theory, and then for any developmental theory of the capacity for moral judgment, which certainly must presuppose the possibility of distinguishing between right and wrong moral judgments (Habermas 1990, p. 120). Without duplicating here a prior discussion respecting the epistemological matter—one that shows how moral stages can be grounded in a logic of development (see Meyer, 2009; based on Habermas 1990)—we should at least reinstate the implications resulting from bringing, as does Habermas, a nonfoundationalist and noninstrumental conception of rationality to the task of understanding the logic of social action, and in particular groups' diverse capacities to coordinate interaction through communication.

Thus, bearing in mind the present investigation on the interplay among stages of moral judgments and "nonbinding" communication, and how it affects the sustainability of normative commitment in collective-action settings, Figure 1 summarizes Habermas's *typology* of action and introduces communicative action as a second form of social action, alternative to strategic action. The first thing the figure indicates is that Habermas does not

reject the instrumental conception of rationality and replace it with an alternative, “communicative” conception. As Heath (ibid) remarks, Habermas takes as his point of departure that agents have available to them a set of different, often incommensurable standards of choice. Communicative action turns out to be action governed by a particular standard, namely, that of reaching *understanding*, while instrumental action is action governed by a different standard: that of reaching *success* in respect of the intended consequences or outcomes of a chosen action.³

According to Habermas’s typology, *instrumental action* and *speech acts* form two “elementary forms of action.” From this, the introduction of a second agent generates *social action*, understood as a complex phenomenon constructed out of the interaction of the two elementary forms of action. According to this view, rational agents engaged in social action are always in a position where they face a 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 or outcomes of their actions, social action takes the form of *strategic action*, in the standard game-theoretic sense. But when speech is used to resolve the coordination problem derived from 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 the Figure 1. The upward oblique line indicates that communicative action is *not* the same as speech, for it also presupposes the basic teleological structure of action inasmuch as the actors are assumed to continue interestedly in carrying out their plans and bring about certain states of affairs in the world. Habermas’s discourse theory of ethics relies on transcendental arguments to show that the resort to communication oriented toward reaching understanding is

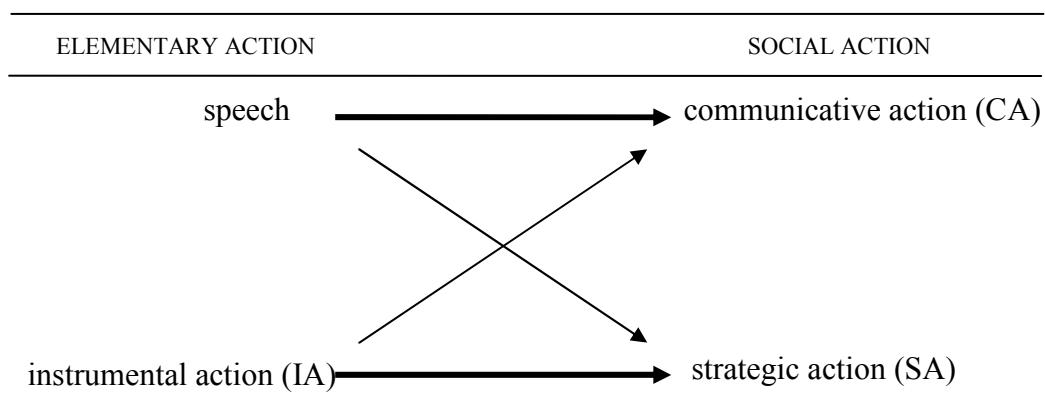
³At this point, it is worth mentioning Heath’s (2001) observation that the foundationalist epistemology underlies the “widespread tendency among social theorists to assume that instrumental action is the only form of rational action, and that norm-governed action must have some kind of nonrational source, such as conditioning, socialization, or habit” (p. 2). He further points to how the presumption of non-rationality of the normative orientations 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 (ibid).

inescapable, and that the mere use of language as an explicit coordination mechanism (rather than consistently aligning beliefs through strategic reasoning) imposes morally relevant constraints on both action alternatives and outcomes.⁴

On the other hand, the descending oblique line in Figure 1 indicates that strategic action is *not* a simple generalization of instrumental action. Like communicative action, it relies upon linguistic recourses as well. However, in this case, language provides only the background for information exchange, so that the intralinguistic objectives of speech acts are subordinated to each agent's individual projects and maximizing strategies. Hence, in the model of strategic action, communication "does not exercise any constraint on the range of action alternative available to agents" (Heath, 2001, p. 24).

Now, with regard to the significance of the developmental framework for the analysis of the limits and potentialities of communication opportunities in fostering normative commitment in collective-action settings, the reasoning motivating the present research rests on the proposition that stages of moral judgment have a bearing on the orientations participants will preferably assume in the face to the communication opportunities allowed in

Figure 1: Elementary action types combine to produce social action types.



Source: Adapted from Heath (2001, p. 25)

⁴ In this regard, it is worth mentioning that, by means of limitative results, Heath (2001) provides a straightforward game-theoretic demonstration in support of Habermas's rationale for introducing communicative action as a second type of rational action. According to him "As far as the instrumental model of rationality is concerned, the fact that agents are able to communicate successfully is completely mysterious," and "This means that any attempts to expand the notion of rational action to account for communication starts out with a certain prima facie plausibility" (p. 81).

the common-pool resource experiment. Habermas himself develops a length argument—the details of which we will not pursue here—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,⁵ are recapitulated in the ontogenesis of our capacity to speech and action, and isomorphic to the stages described in Laurence Kohlberg’s model of the development of sociocognitive and moral reasoning.⁶ 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. “The point of this chain of argument is to connect structures of moral judgment to structures of social interaction in such a way that their developmental-logical features stand out more clearly” (McCarthy, 1990, p. ix).

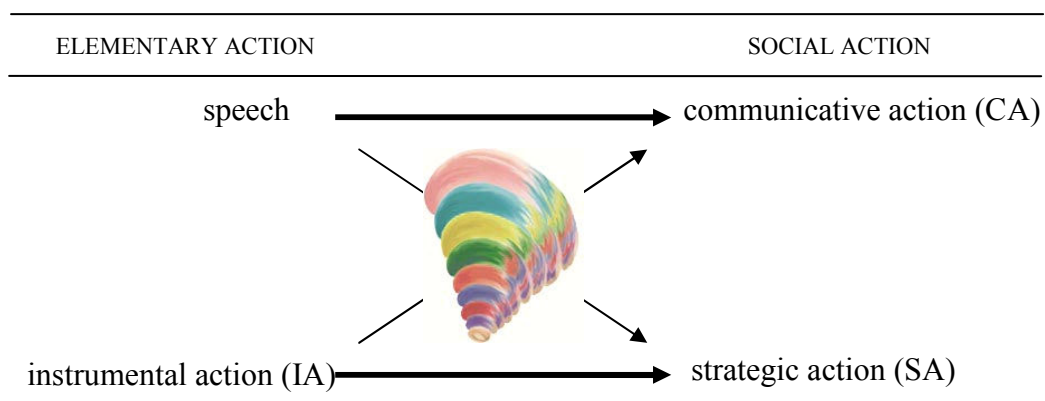
Respecting our immediate research interest, the point is that Habermas action-theoretic approach lends support to the view that increased willingness to cooperate and to commit to normative agreements set forth thorough “nonbinding” communication in social dilemmas situations are associated to higher stages of sociocognitive and moral reasoning. In Habermas’s account, this is due to the existence of an asymmetry between the developmental requisites for strategic action and action oriented toward reaching understanding. Again, without going over Habermas’s argument, we simply illustrate the basic idea in Figure 2 by inserting a pictographic representation of Graves’s model as a dynamic spiral (due to Beck and Cowan, 1996) into Heath’s illustration of Habermas’s typology of rational action. In this representation, the color code (from beige to turquoise) stands for the eight stages show in Table 1. Respecting the behavioral hypothesis to be specified in the next session, the general expectation is that speech acts shall imply the normative commitments draw in communicative action only if participants in communication are centered at the sixth (green) or seventh (yellow) stage of Graves’s model. As presumed,

⁵ This phylogenetic account is in the fifth chapter of *The Theory of Communicative Action* (1984/1987, v. 2).

⁶ 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).

participants centered at the fourth stage (blue) shall assume a merely conformative stance toward the moral content of utterances in support of normatively governed interaction, while individuals centered at either the third (red) or the fifth (orange) stages are expected to engage in strategic action, wherein the content of speech acts provide only the information information—such as common knowledge of preferences, action alternatives and action-outcome linkages, other participants' reputation as cooperators, expectations of sanctions, and so on—against which the agents determine maximizing strategies.

Figure 2: Stages of value judgment affect the way how elementary action types combine to produce social action types.



Source: Author's configuration

4. Method

In order to test value of Graves's constructs to explain and predict behavior in collective-action dilemmas, the psychosocial profiles of 322 Brazilian participants (44% females and 56% males) were assessed by means of an authorized Portuguese translation of the assessment tool, originally developed by Hurlbut (1979).

Only a selected number of those who have filled out the survey actually took part in either one or both experiments. The selection of the actual participants was not intentional. It occurred according to participants' possibilities to attend the experimental schedule. In the common-pool resource (CPR) appropriation dilemma we had 200 participants (94 female

and 106 male), divided in 25 groups of 8 participants each. In the public-goods (PGG) provision dilemma there were 127 participants (one group).

4.1. Recruitment procedures and characteristics of the sample

For the recruitment of the potential participants we used the service of the student office to send out a message announcing the experiment to all students who had email account in the University Federal of Viçosa (virtually all students). We also had the support of the Associations of Graduate Students, University's Employees and Professor Staff to email the announcement to their respective lists of contacts. We also fixed a printed version of the announcement into virtually all buildings in the University Federal of Viçosa and in various spots distributed all around the downtown.

The respondents were for the most part, but not exclusively, undergraduate and graduate students from various major degrees. Respecting the ages, 43% were between 18 and 21 years old and 52% between 22 and 29 years old. We have only 10 people in their thirties, 3 in their forties and just one man in his fifties. Concerning birthplaces, 92% were from the Minas Gerais state, 3% from São Paulo, and 2% from Rio de Janeiro. The few remaining were from Bahia, Espírito Santo, Brasília (D.F), and Pará.

4.2. Experimental settings and general procedures

The experiments were conducted in a classroom of the Department of Agricultural Economics, at the Federal University of Viçosa, Brazil. In order to link the individuals' behaviors in the experiments with their psychosocial profiles, participants ought to write their names in the corresponding record sheets, as when filling out the survey. At the experimental setting, participants were accommodated so as to guarantee the privacy of the individuals' decisions. To preserve anonymity in the CPR experiment, where individual decisions were published in the blackboard, identification numbers were randomly ("pick one") distributed before beginning.

Since the research involved sequential experimental sections (25 repetitions), the possibility of post-experiment communication among acquainted participants could not be

avoided. In order to neutralize the effect of such communications, the whole experimental design was revealed in advance. In addition, the existence of a communication condition in the CPR experiment and the relatively large size of the groups suggest that the eventuality of post-experiment communication is not a concern in explaining individuals' choices in the experiments.

4.3. On the assessment tool

The assessment tool consists of forty multiple choice questions in the *Most Like Me/Least Like Me* format. It was designed to reveal a person's psychosocial profile (from 2nd to 7th stage)⁷ with reference to his/her overall lifestyle and not any compartmentalized area of life such as his/her professional occupation, family life, religious beliefs, etc.. Also, while the holarchical structuration means that a person may have mixtures of different systems and functioning, Hurlbut's test was designed to reveal a person's overall dominant value system, secondary value system, and so on.

Table 3 offers an illustration of a representative statement associated to each of the six selected stages of development, in response to one of the questions in the survey. Without knowing to which stage of development each statement is associated (as the respondents have no information whatsoever about the developmental theory under the survey), they ought to indicate for each of the forty questions both one statement they regard as "most like me" and one they regard as "least like me." The distribution of frequencies of "most like me/least like me" responses associated with each stage of development constitute the basic quantitative information to be used in the statistical analysis described ahead.

In view of the present application, it is worth mentioning that Hurlbut's test was expected to be valid only for speakers of Standard American English, and for persons with

⁷ Both the 1st and the 8th stages or waves are not covered by assessment tool. In the first case, the behavioral systems associated to the 1st level it is conjectured to hold for approximately 0.1 percent of the world adult population (Beck & Linscott, 1991; Wilber, 2001), as alternatively found in senile elderly, late stage Alzheimer's victims, mentally ill people, and starving masses. Likewise, the 8th level is thought to be relatively rare and considered to represent the present "leading edge" of *collective* human evolution. As examples of some 8th level's intellectual products, Beck and Cowan (1996) mention items that include Teilhard de Chardin's concept of noosphere, the growth of transpersonal psychology, chaos and complexity theories, integral-holistic systems thinking, and Gandhi's and Mandela's pluralist integration.

the equivalent of a high school education. However, Graves regarded his biopsychosocial conception to be fundamentally cross-cultural valid, so that we carried out a Portuguese translation with the consent and in collaboration with the National Values Center in order to test Graves's claim in a new cultural context.

Table 3. Illustration of the representative statements on one general question in the survey associated to each stage of psychosocial development

FOR ME THE WORLD IS...			
Stage or wave	Cyclical aspect	Thinking	Representative statement
7 th	Express-self (agency)	Ecological	<i>a chaotic organism forged by differences and change</i>
6 th	Sacrifice-self (communion)	Consensus	<i>a human habitat in which we share life's experiences</i>
5 th	Express-self (agency)	Strategic	<i>a market place full of possibilities and opportunities</i>
4 th	Sacrifice-self (communion)	Authority	<i>an ordered existence under the control of the ultimate truth</i>
3 rd	Express-self (agency)	Egocentric	<i>a jungle where the strongest and most cunning survive</i>
2 nd	Sacrifice-self (communion)	Animistic	<i>a magical place alive with spirit beings and mystical signs</i>

Source: Adapted from Beck (1999)

4.4. Factor Analysis: pretest of the cross-cultural robustness of Graves's constructs

The burden of proof on a theory that grounds a conception of human valuation in anything more universal than the "settled convictions" of cultures is enormous. In the way of verification, the main contribution of our experimental work consists in testing a series of theoretical expectations derived from Graves's constructs vis-à-vis participants' behavior in the different experimental conditions.

However, the wave-like, flowing phenomenon underlying the interior growth means that Graves's stages or "levels of existence" impart only *nodal* positions, or "centers of psychosocial gravity," not the *total* systemic manifestations of concrete individuals. Yet, these nodal positions must follow an invariant sequence, as posited, so that the interwaving of value systems does not occur at random, but must, instead, put across

certain logic of mutual evaluative perspectives vis-à-vis the multiple stages. This means that results from factor analysis can indicate the cross-cultural robustness of Graves's scheme, because the factor loadings in the component matrix must convey a nonarbitrary pattern of correlations between the "most like me/least like me" variables in the survey data, if it is going to generate interpretable or theoretically meaningful components. In other words, we suggest that the extraction of theoretical sound factors is a sign of both the quality of the translation and of the cross-cultural robustness of Graves's general scheme.

4.5. Common-pool resource (CPR) appropriation dilemma

4.5.1. Experiment summary and design

Following Ostrom *et al* (1994), the experiment consisted of thirty rounds in which the eight participants must decide how to allocate an endowment of fifteen tokens between two alternative investment opportunities. The alternative opportunities were generically labeled as Market 1 and Market 2.

Market 1 is a safe, outside activity in which each token yields a constant rate of output and each unit of output yields a return. Market 2 (the CPR) is a market that yields a rate of output per token dependent upon the total number of tokens invested by the entire group. The individuals' share of the total output produced in the Market 2 is equivalent to the percentage of total group tokens invested in that alternative. Both the production function and allocation rule create that fundamental conflict between the individual and the collective interests that characterize the dilemma.

The experimental design involved three different conditions, ran in the following sequence: (1) ten rounds where decision were made without communication (open-access, baseline); (2) ten rounds following ten minutes of costless communication (communication), and (3) ten rounds with a costly sanctioning mechanism supplied (sanctioning) (Figure 1).

During the communication condition, participants' have the opportunity to devise a joint strategy that may raise the collective appropriation, but commitment with such strategy is strictly nonenforceable. A second nonobligatory communication opportunity between the

15th and the 16th rounds is the only reinforcing mechanism during the communication condition. The third condition introduces a costly sanctioning mechanism, so that the joint strategy may be reinforced if the participant decides to incur into a given fee in order to impose a corresponding fine on defectors. The value of both fees and fines are discounted from participants' final payment. Decisions whether or not to fine defectors are taken immediately after the publication of the individuals' investment decisions in each round of the sanctioning conditions. Prior to subsequent rounds, all participants are *privately* informed respecting the fines eventually imposed on them. Anonymity is guaranteed all through the experiment and afterward.

4.5.2. Theoretical expectations

The production function employed in the CPR experiment (Ostrom *et al*, 1994) determines that the appropriation dilemma consists in reducing the total group investment from 64 tokens (unique and symmetric Nash equilibrium) to 36 tokens (Pareto optimum).⁸ According to standard game theory, the Nash equilibrium is supposed to obtain in all the three conditions (open-access, nonbinding communication, and costly sanctioning), because insofar as the groups' norm is nonenforceable communication is supposed to have no consequence whatsoever (cf. e.g. Harsanyi & Selten, 1988, p. 3), and "rational players" are

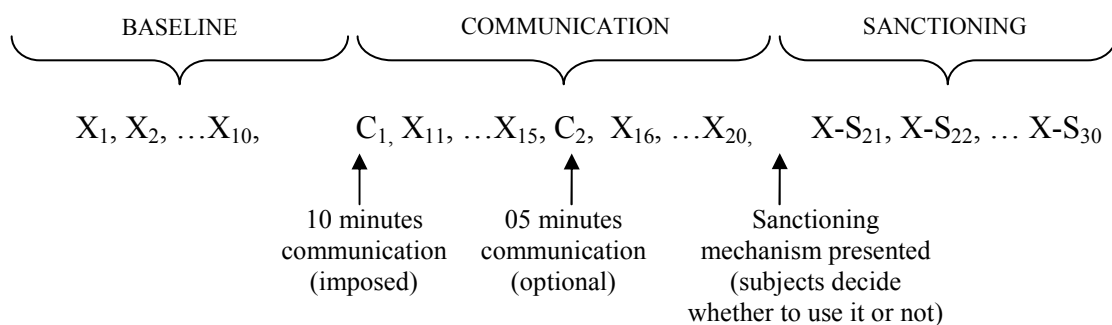


Figure 1. CPR experimental design

⁸ The symmetric rules totalizing either 32 or 40 tokens yield the same and slightly suboptimum collective outcome.

supposed not to engage in costly sanctioning (because of its nature of public-goods), regardless if the game is one-shot or finitely repeated (Ostrom et al., 1994).

However, repeated evidences have shown both the significant effect of communication in increasing the selection of cooperative strategies in repeated interaction settings (e.g. Braver & Wilson, 1986; Bornstein, & Rapoport, 1988; Hackett et al, 1994; Ostrom et al 1994), and that many people, but not all, are willing to incur in personal cost in order to reinforce social norms and collective agreements (e.g. Henrich & Boyd, 2000; Barr, 2004; Fehr, Fischbacher & Gächter, 2002).

Our purpose is to examine the helpfulness of Graves's constructs in explaining individuals' *different* responses in face of both the communication and sanctioning conditions. Hence, we suggest that the *opportunism* associated to the fragility of nonbinding communication is characteristic *not* of rationality itself, as implied in the standard explanation, but of the intentional states and value judgments associated to both 3rd and 5th stages of interior development, as described by Graves's model (Table 1). On the other hand, we suggest that the positive effect of communication opportunities is *not* an effect of communication per se, but a combined effect of communication and the interior dispositions of individuals centralized at the 2nd, 4th, 6th or the 7th stages of psychosocial development. Clearly, different motivations and reasoning underlie similar behavioral responses associated to different phases of psychosocial development. Nonetheless, with regard to our behavioral observations we raise the following hypothesis:

CPR_ communication: *positive associations between the frequency of defection during the communication condition and participants' psychosocial profiles must indicate either 3rd or 5th psychosocial centralization stages, whereas negative associations must be indicative of 2nd, 4th, 6th or 7th stages of psychosocial development.*

With regard to the sanctioning condition, the information in Table 2 should remind us that individuals' centralized at any stage of development may be willing to incur into sacrifices, but for different reasons and with different motivations. In the present experimental context, where the decisions were suggested to involve "investments" in two

generically labeled market alternatives, we expect the decisions to punish defectors to be associated to the factor scores in the 2nd, 4th or the 7th stage. However, only associations with the 7th stage are interpreted as “altruistic punishment,” meaning that the disposition to incur in the costs is fully motivated by the intent of reinforcing a norm for the sake of the collective good. In the two other cases, even if there is the intent of reinforcing the social norm, Graves’s descriptions suggest that a certain amount of taste in punishing deviators subjectively discount the actual cost of the sanctioning. Similarly, we do not expect that decisions to sanction be associated to the 3rd, 5th or the 6th stages, but for different reasons. While taste for punishment is definitely set in the 3rd stage of psychosocial development, we believe that the opportunity of free ride on the sanctioning provided by others should dominate the 3rd stage’s motivational complex in the experimental context. Opportunism is undoubtedly the explanation for the behavior associated to the 5th level. On the contrary, while the psychosocial profile of the 6th stage certainly allows us to expect it to be willing to incur into the costs of safeguarding the collective good, we do not expect it to be particularly prone to make use of the provided sanctioning mechanism out of its convicted distaste for coercive means. In summary, considering the behavioral observations open in the sanctioning condition we raise the following hypothesis:

CPR_ sanctioning: *positive associations between the expenditures with fees (which are required in order to fine defectors) during the sanctioning condition and participants’ psychosocial profiles must indicate 2nd, 4th or 7th psychosocial centralization stages, whereas negative associations must be indicative of 3rd, 5th, or 6th stages of psychosocial development.*

5. Results

5.1. Factor analysis suggests that Graves’s scheme in cross-culturally robust

Table 4 presents the component matrix from the factor analysis applied on the survey data. The moderate communalities indicate that the scores in the survey are considerably scattered along the tridimensional space defined by the principal axes. Such dispersion reflects the natural overlapping and interwaving of the different psychosocial

positions or perspectives, as discussed in the theoretical section. What matters most is that the correlation pattern in the component matrix reveals three theoretically sound, plainly meaningful principal components (statistically significant correlations are detached in bold).

Skipping comments on the 2nd stage of psychosocial development, which plausibly showed to be less useful for distinguishing the *nodal* positions present in a sample of undergraduate and graduate students, the first component (Factor 1) clearly expresses the *egocentric wave*. We see this in the acceptance of the values and worldview associated to the 3rd stage (positive correlations with 3th_most, and negative correlations with 3th_least), which is expected to be accompanied, as in Table 4, by a rejection of those values and worldviews linked to both the 4th stage and the 6th stage (positive correlations with 4th_least and 6th_least, and negative correlations with 4th_most and 6th_most). On the other hand, the identification with the representative statements linked to the 4th stage, as in the Factor 2, is expected to be accompanied by a sense of strangeness and strong rejection of the worldview hold at the 7th stage (see Table 3). We thus label the Factor 2 as the *absolutistic wave*, in reference to the 4th stage's thinking (Table 1). Finally, the Factor 3 expresses the *sociocentric wave*, showing identification with the statements associated to the 6th stage and strong rejection, as expected, of the positions linked to the 5th stage.

Taken together, the three factors explain about 70% of the total variance in the sample (Factor 1 = 23.44%; Factor 2 = 22.15%; and Factor 3 = 24.56%, after Varimax rotation⁹). We claim that the theoretical meaningfulness of principal components is an indication of the cross-cultural robustness of Graves's constructs. Otherwise, there is no reason, other than chance, to expect theoretically sound mutual evaluations involving Graves's constructs in a sample of Brazilian participants, assessed about three decades following their original conception in the United States. This claim is supported by the associations between the factors scores and the behavioral observations produced in the experimental conditions, as we shall see.

⁹ It should be mentioned that when the Varimax rotation is done the *maximum variance property* of the original components is destroyed. The rotation essentially reallocates the factor loadings and, thus, the first rotated factor will no longer *necessarily* account for the maximum amount of variance.

Table 4
Factor Analysis: Rotated Component Matrix^a

Variable	Rescaled communalities	Rescaled Component (Factor)		
		1	2	3
2 nd _most	.466	.369	.476	.335
2 nd _least	.592	.002	-.464	-.614
3 rd _most	.538	.722	.072	-.109
3 rd _least	.851	-.897	.124	.176
4 th _most	.733	-.346	.783	-.007
4 th _least	.733	.727	-.453	.020
5 th _most	.696	.092	-.052	-.827
5 th _least	.532	-.069	-.246	.683
6 th _most	.788	-.423	-.069	.774
6 th _least	.518	.569	.256	-.359
7 th _most	.766	.109	-.850	-.176
7 th _least	.521	.215	.682	.104

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

The correlations detached in bold are all statistically significant

a. Rotation converged in 7 interactions.

Source: Research results

5.2. Communication and sanctioning helps a lot: renewed evidences

With regard to the impact of the institutional changes, our results just replicate consistent findings shown in other experimental works. They indicate that the efficiency of the collective appropriation in the CPR experiment is significantly higher ($p < .0001$) in both the communication and the sanctioning conditions, when compared with the collective appropriation pattern during the baseline (unregulated) condition.

However, these are average effects, and do not reveal the variety of individual responses to those different incentive structures. Here we take advantage of the existing variability in order to verify whether Graves's constructs help to understand both the limits and the potentialities of communication opportunities and the likelihood that a costly sanctioning mechanism might be provided.

5.3 Similar incentives, different responses: the value of the psychosocial information

Concerning the communication condition, we find that “unbinding” normative commitment is clearly supported by the *sociocentric* wave (Factor 3) and hazard by the *egocentric* wave (Factor 1). Results from multivariate linear regression models wherein the three principal or nodal waves of existence figure as explanatory variables suggest that the higher the scores in *sociocentrism* the lower the total investment after communication ($F_{3,192} = 2.245$, $p = .019$), and the lower the frequency of defections during the 10 rounds of the communication condition ($F_{3,192} = 7.883$, $p = .0001$). As indicated in the previous section, this result expresses the acceptance of value systems and social perspectives linked to the 6th stage and the rejection of those linked to the 5th stage of psychosocial development, as defined in Graves’s model.

Conversely, the frequency of defections is positively associated ($p = .042$) with the scores in *egocentrism*. Since these defections may reflect not only opportunism but also a punitive reaction to prior deviations of others, we set two binary logistic regressions in order to distinguish the opportunistic behavior (primary or uncalled defections) from the steadfastness in supporting the collective agreement. Results suggests that the probability of finding an individual who have stuck to the group agreement all through the 10 rounds of the communication condition, even in the face of the defections of others, is statistically higher ($p = .001$) if the individual scores higher in *sociocentrism* (Factor 3). Conversely, the probability of finding an individual that have taken the initiative of cheating is higher ($p = 0.051$) if the individual scores higher in the *egocentric* wave (Factor 1).

These results, which summarize a series of theoretically sound, significant bivariate (Pearson) correlations between the experimental observation and the original survey data, are in line with the hypothesis **CPR_ communication**. As we see them, these findings indicate that the weakness of the noncoercive agreements, commonly attributed to the *homo homini lupus* conception of the human nature, or else to the “inherent logic” of the commons, can be better explained as ensuing from shared centralization at the 3rd stage of

psychosocial development, as well as at the 5th stage, to the extent that the social perspective of the latter contrasts with that of the 6th stage in the *sociocentric* wave (Factor 3).

Concerning the sanctioning condition, a bivariate regression model supports the expectation that the *sociocentric* wave would not engage in punishment (**CPR_sanctioning** hypothesis). The results show a negative relationship ($F_{1,194} = 2.804$, $p = .096$) between the scores in sociocentrism (explanatory variable) and the total expenditure with fees. Also, a negative relationship ($F_{1,194} = 3.097$, $p = .08$) between the scores in sociocentrism and total outflow due to fines restates the sociocentric adherence to the collective agreement. The interpretation of this observation as expressing 6th stage's distaste for retaliatory means and preference for consensus, instead of an attempt to free ride on sanctioning provided by others (compare Tables 1, 2 and 3), as in the justification of the **CPR_sanctioning** hypothesis, is allowed by results in the Public-Goods experiment, as we shall see in a moment.

With regard to the *absolutistic* wave (Factor 2), no statically significant associations were found in the CPR experiment (or in the PGG experiment). Yet, rather than being a disinteresting sign, we attribute this fact precisely to the dramatic change in the way as the self expresses its desires when psychosocial development achieves the 7th stage (Table 1). What happens is that, as shown in the correlations built in the Factor 2 (Table 4), the perspective and motivations hold at the 4th stage contrasts with those associated to the 7th stage, but, at the same time, the expectations regarding the behavioral responses ensuing from both psychosocial systems in both experimental conditions (as in the PGG experiment) converge (see the **CPR_communication**, **CPR_sanctioning**, and the **PGG_contribution** hypotheses). These opposing forces then collaborate to fade the statistical associations involving the Factor 2. Still, the examination of the bivariate (Pearson) correlations using the original scores from the survey data allows us to provide some evidence of the *altruistic* disposition to incur in costly sanctioning linked to the 7th stage. On the one hand, the correlations show that the *acceptance* of the 7th stage's values (7th_most) is *positively*

associated ($p = .064$) with the total expenditures with fees, whereas their *rejection* (7th_least) correlates *positively* ($p = .060$) with the *frequency of defections*. These associations suggest that those who defect more are *not* centralized at the 7th stage of psychosocial development. Additional findings reinforcing this interpretation come from the Public-Goods experiment (below) and the Ultimatum game experiment (not shown).

With regard to the 5th stage, in addition to the interpretations allowed on the basis of the rejection of its stance, as set in the Factor 3 (Table 4), a more direct observation of the 5th stage's go-getting character is provided by a significant ($p < 0.10$) *positive* correlation (Pearson) between the *acceptance* of 5th stage's representative statements (5th_most) and the total investment in the Market 2 (the CPR) during the *baseline* condition. This relationship is confirmed by a *negative* and significant ($p < 0.05$) correlation between the total investment in the Market 2 and the *rejection* of the representative statements of the 5th stage (5th_least).

In spite of being just a small piece of additional information, we regard these correlations with 5th stage remarkable because only in the baseline condition the players have no constraint, other than their ability to think strategically, to behave according to their own individual bent. In this circumstances, the observation that the appropriation behavior linked to the 5th stage perspective stand above the sample offers an alternative interpretation for the "tragedy of the commons:" namely: instead of assuming that "Each man in locked into a system [i.e. external situation] that compels him to increase his herd [i.e. the investments in Market 2] without limit," as Hardin (1968, p. 1244) suggests, one might well admit that "man" is actually locked into a interior psychosocial system that compels him to see the world as "a market place full of possibilities and opportunities" (Table 3), and where he should "express self for what self desires," since there is no constraint that could "bring down the wrath of others" (Table 1).

6. Conclusion and implications

Even though one observes today among institutional and behavioral economists an increasing recognition of the limits of the standard model of individual in economic theory, and a growing consensus around a number of structural and contextual variables enhancing the likelihood of co-governance in collective-action settings, the lack of a theory of human valuation explaining how individuals reach diverse utility judgments in face of similar incentive structures and contexts of action hinders the advancement of a more general theory of how alternative institutions affect the incentives confronting individuals and their resultant behavior in social dilemmas. This theoretical deficit, in turn, underpins on the part of those convinced that all human behavior can be explained using rational egoist models the continual recommendation of not always effective, and principally uninspiring Leviathan-like remedies to overcome all social dilemmas.

We suggested that the paradigm of developmental structuralism offers a series of empirically grounded theories and model of the development of the sociocognitive and moral reasoning that can be tested to predict *diverse* behavioral responses to similar incentive structures in social dilemma situations. Our results point towards the worthiness of one of them: Graves's *biopsychosocial* conception of organism and emerge-cyclical model of personality systems development and cultural institutions. Designed to rationally reconstruct the pretheoretical knowledge of competently judging subjects, Graves's constructs showed useful to predict and interpret participants' diverse behaviors in the CPR appropriation dilemma with both "nonbinding" communication and sanctioning conditions. The usefulness of Graves's constructs to predict behavior in a public-goods experiment under variable institutional conditions is reported in Meyer and Braga (2009b).

The policy implication linked to the developmental point of view ensues from both the structural features involved and the qualitative changes brought about along the developmental path. On the one hand, the holarchical unfolding of the stages means that the interior perspectives emerge in an order that cannot be altered by external conditioning or social reinforcement, because the drives of interior development is not simply imported from

the external world but follows also its own internal laws of transformation. On the other hand, there is a great amount of consistence among alternative developmental models in describing such process has involving, for the most part, a continuous decline in egocentrism, increase in autonomy, and increase in the moral embrace. If one takes these features seriously when analyzing and recommending institutions aimed at overcoming social dilemmas, one cannot avoid the conclusion, reached by Graves himself, that the prime goal of institutional designers should be to devise institutions designed first and foremost to promote human movement up the spiral of consciousness unfolding.

We hope our experimental work might foster a renewed awareness of the significance of an alternative approach to the interplay among rationality, norm-governed action, and institution—one anchored in constructivist developmental thinking.

7. References

- AGRAWAL, A. (2001). Common Property Institutions and Sustainable Governance of Resources. *World Development*, v. 29, p. 1649-1672
- ANDERSON, J.R. (1980). *Cognitive Psychology and Its Implications*, San Francisco: W. H. Free-man.
- AXELROD, R. (1997). Advancing the art of simulation. In: CONTE, R.; HEGSELMANN, R.; TERNA, P. (Eds.). *Simulation social phenomena*. Berlin: Springer, 1997. p. 21-47
- BALAND, J.M.; PLATTEAU, J.P. 1996. *Halting degradation of natural resources: is there a role for rural communities?* Oxford: Clarendon Press, 1996.
- BARR, A. (2004). Kinship. Familiarity. and Trust. In J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, & H. Gintis (Eds.). *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.
- BECK, E.D.; LINSKOTT, G. (1991). *The crucible: forging South Africa's future*. Johannesburg, South Africa: The New Paradigm Press.
- BECK, D. (1999). The Search for Cohesion in the Age of Fragmentation From the New World Order to the Next Global Mesh. In http://spiralodynamics.net/DrDonBeck/essays/search_for_cohesion.htm.
- BECK, E.D.; COWAN, C.C. (1996). *Spiral Dynamics. Mastering Values, Leadership and Change. Exploring the New Science of Memetics*. London : Blackwell Publishers.
- BORNSTEIN. G.; RAPOPORT, A. (1988). Intergroup competition for the provision of step-level public goods: effects of preplay communication. *European Journal of Social Psychology* 18, 125-142.
- BRAVER, S.; WILSON, L. (1986). Choices in social dilemmas: effect of communication within subgroups. *Journal of Conflict Resolution* 20(1), 51-62.
- CAMERER, C. (2003). *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton University Press.
- 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. (2000). Hierarchical complexity of contingencies of reinforcement and of stimulus discriminations. Washington, D.C. (Presented at the Society for Quantitative Analysis of Behavior).
- 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.
- COWAN, C.; TODOROVIC, N. (2005). *Spiral Dynamics I: Certification*, NVCC, Santa Barbara.
- FEHR, E., FISCHBACHER U., GÄCHTER S. (2002). Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms. *Human Nature* 13, 1-25.
- FEHR, E.; FISCHBACHER, U.; GÄCHTER, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, v. 13, p. 1-25.
- FROHLICH, N.; OPPENHEIMER, (2001). J.A. Choosing: a cognitive model of economic and political choice. Winnipeg: University of Manitoba, Faculty of Management. (Working Paper).
- GINTIS, H. (2000). Beyond Homo economicus: evidence from experimental economics. *Ecological Economics*, 35(special issue), 311-322.
- GRAVES, C.W. (1970). Levels of Existence: An Open System Theory of Values. *Journal of Humanistic Psychology*, Fall 1970, Vol. 10, No. 2, pp. 131-155.
- GRAVES, C.W. (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.
- HABERMAS, J. (1993) *Justification and Application: Remarks on Discourse Ethics*; translated by Ciaran Cronin. MIT Press, Cambridge, Massachusetts, and London, England.
- HACKETT, S., SCHLAGER, E.; WALKER, J.M. (1994). The role of communication in resolving commons dilemmas: experimental evidence with heterogeneous appropriators. *Journal of Environmental Economics and Management*, 27(2), 99-126.
- HARSANYI, J.C.; SELTEN, R. (1988). *A general theory of equilibrium selection in games*. Cambridge, Mass.: MIT Press.
- HEATH, J. (2008). *Following the rules: practical rationality and deontic constraint*. Oxford and New York: Oxford University Press.
- HEINER, R.A. 1983. The origin of predictable behavior. *The American Economic Review*, v. 73, n. 4, p. 560-595.
- HENRICH, J.; BOYD, R. (2000). Why People Punish Defectors. Weak Conformist Transmission can Stabilize the Costly Enforcement of Norms in Cooperative Dilemmas. *Journal of Theoretical Biology*, 208, 79-89
- HURLBUT, M. A. (1979). *Clare Graves's Levels of Psychological Existence: A Test Design*. North Texas State University, Denton, Texas. PhD dissertation.
- KOHLBERG L. (1969). Stage and Sequence: The Cognitive-Developmental Approach to Socialization, in D. A. Goshn, ed., *Handbook of Socialization Theory and Research*, New York: Rand McNally, pp. 347-480.
- KOHLBERG L. (1976). Moral stages and moralization: the cognitive developmental approach. In: LICKONA, T. (Ed.). *Moral development behavior*. New York: Holt, Rinehart and Winston.
- KOLLOCK, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology*, 24, 183-214.
- KOMORITA, S. S., & PARKS, C. D. (1995). Interpersonal relations: Mixed motive interaction. *Annual Review of Psychology*, 46, 183-207.
- KOPELMAN, S., WEBER, J. M., & MESSICK, D. M. (2002). Factors influencing cooperation in commons dilemmas: A review of experimental psychological research. In E. Ostrom, T. Dietz, N. Dolsak, P. C.
- LEPYARD, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 111-194). Princeton, NJ: Princeton University Press.
- MANTZAVINOS, C., NORTH, D.C. and SHARIQ, S. (2004). Learning, Institutions, and Economic Performance. *Perspectives on Politics*, vol. 2, nr. 1.

- MCKEAN, M. A. (1992). Management of Traditional Common Lands (Iriaichi) in Japan. In *Making the Commons Work: Theory, Practice, and Policy*, ed. Daniel W. Bromley, 63-98. San Francisco: Institute of Contemporary Studies.
- MEYER, L.F.F. (2009) Cognition and Norms: Toward a Developmental Theory Linking Trust, Reciprocity, and Willingness to Cooperate. Working paper at the Fourth Workshop on the Workshop in Political Theory and Policy Analysis (WOW4). Indiana University, Digital Library of the Commons.
- MEYER, L.F.F.; BRAGA, M.J. (2009). Fear or greed? Duty or solidarity? Motivations and stages of moral reasoning: experimental evidences from public-goods provision dilemmas. Working paper at the Fourth Workshop on the Workshop in Political Theory and Policy Analysis (WOW4). Indiana University, Digital Library of the Commons.
- OSTROM, E. 1990. *Governing the commons*. Cambridge: Cambridge University Press.
- OSTROM, E. 2005. *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press.
- OSTROM, E.; GARDNER, R.; WALKER, J.M. 1994. *Rules, games, and common-pool resources*. University of Michigan Press, Ann Arbor, MI.
- PIAGET, J. (1932). *The moral judgment of the child*. New York: Free Press.
- PIAGET, J. (1970/1972). Intellectual evolution from adolescence to adulthood. *Human Development*, v. 15, p. 1-12.
- POPPE, M. (2005). The specificity of social dilemmas situations. *Journal of Economic Psychology*, v. 26, p. 431-441.
- POPPE, M.; ZWIKKER, M. (1996). The effect of threshold level on greed, fear, and cooperation in step-level give-some and take-some games. In W. B. G. Liebrand & D. M. Messick (Eds.), *Frontiers in social dilemma research* (pp. 185–204). Berlin: Springer.
- SCHLAGER, E. (1990). Model specification and policy analysis: the governance of coastal fisheries. [PhD Thesis] Indiana University.
- SEP, (2003). *Stanford Encyclopedia of Philosophy* <http://plato.stanford.edu/entries/game-evolutionary/>
- SIMON, H. (1962). A "The Architecture of Complexity: Hierarchic Systems", *Proceedings of the American Philosophical Society*, vol. 106: 467—482
- SIMON, H.A. (1979) *Models of Thought*. New Haven: Yale University Press.
- TANG, S. Y. (1994). "Institutions and Performance in Irrigation Systems." In *Rules, Games, and Common-Pool Resources*, ed. Elinor Ostrom, Roy Gardner, and James Walker, 225–45. Ann Arbor: University of Michigan Press.
- VAN LANGE, P. A. M., LIEBRAND, W. B. G., MESSICK, D. M., & WILKE, H. A. M. (1992). Introduction and literature review. In W. B. G. Liebrand, D. M. Messick, & H. A. M. Wilke (Eds.), *Social dilemmas* (pp. 3–28). Oxford: Pergamon Press
- WADE, J. (1996). *Changes of Mind*. Suny Series in the Philosophy of Psychology. Albany, State University of New York Press
- WILBER, K. (2000). *Integral Psychology: Consciousness, Spirit, Psychology, Therapy*. Boston, Mass.: Shambhala.
- WILBER, K. (2001). *A Theory of Everything*. Boston, Mass.: Shambhala.