The Peculiar Logic of Value*

RAY JACKENDOFF**

ABSTRACT

A system of values plays an important intermediary role in the human conceptual system. An individual associates a value – an abstract valence and quantity – with a past, present, or contemplated object or action in the environment, and uses values to help determine what actions to take. Value can be categorized into a number of different types, the most important of which for the purposes of the present article are *affective* value (how good the action is *for* someone), normative value (how good the action is *of* someone), and *esteem* (how good a reputation someone has). Normative value in turn divides into several subdomains such as moral value, legality, etiquette, and religious value. In addition, values can be divided along the orthogonal dimension of *objective* (value tout court) versus *subjective* (value in the eyes of a particular person). Each type of value plays its own characteristic role in affecting choice of action in interaction with the others.

Given an explicit formulation of systems of value, it then becomes possible to work out more precise accounts of value-laden systems of concepts. Two are explored here: fairness and freely chosen reciprocity, the latter including retaliation, restitution, honoring, shaming, and apologizing.

The overall hypothesis that emerges from the investigation is that the basic categories of value and the inference rules built on them are human universals. Cultures' value systems differ primarily in (a) what actions and objects are assigned what values and (b) the relative weighting of different sources of value in cases where they interact.

KEYWORDS

value, morality, reciprocity, fairness

* This article is excerpted from Jackendoff (to appear), in which many of the principles stated here receive more formal treatment, and which in addition treats a wider range of phenomena and ties values in with other social and mental predicates such as *intend*, *interesting*, and *obligation*. I have benefited greatly from discussions of these ideas with Hildy Dvorak, Beth Jackendoff, Steve Umans, Marion Smiley, Janet McIntosh, and members of my class Cognition of Society and Culture, taught jointly with Janet McIntosh at Brandeis University in spring 2005.

** Tufts University, Medford, MA, 02155 U.S.A. ray.jackendoff@tufts.edu

© Koninklijke Brill NV, Leiden, 2006

1. Overview

The overall question to be addressed here is what a system of values is, and what components make it up. The basic approach will not be to ask what values "are in the real world." They are not *anything* independently of the people who conceptualize them. Rather, in consonance with the goals of conceptualist semantics (Jackendoff 1983; 2002, chapters 9-12), the question is how humans conceptualize values (especially unconsciously[1]) and how values play a role in governing people's judgments and behavior. This question will be approached in part by using linguistic expressions of value as clues for the structures we are seeking. The inquiry will be validated by the extent that, by positing a relatively constrained set of conceptual building blocks, we can describe a rich variety of linguistic expressions and commonplace intuitions involving value. Through understanding the conceptual structures in which values are embedded, it is to be hoped that we are in a better position to inquire into the evolutionary and cultural roots of systems of value.

One thing I will not aspire to here is to decide what value system we should adopt, that is, to make value judgments over value systems. Nevertheless, I concur with Greene 2003 that investigation of the present sort can be useful in addressing such issues.

My overall hypothesis is that value is a conceptualized abstract property attributed to (conceptualized) objects, persons, and actions. It is abstract because it is not directly perceptible. The value of an entity plays a role in various rules of inference which affect the way one reasons about the entity. Thus value serves as an intermediary in a system of logic – not logic in any standard propositional/formal sense, but in the sense of a conceptual/heuristic logic, as it were a "folk logic". The rules of this logic form a sort of internal accounting system that helps connect many sorts of disparate objects, actions, and persons. Cultures vary in what value is conventionally assigned to what sort of entity and in what context, but, according to my hypothesis, the basic parameters

[1] See Barth 1993 for discussion of how (in the present terms) conscious, verbalizable values may be quite different from implicit (unconscious) values, which can be detected through regularities of behavior.

of the logic of value – that is, the internal accounting system – are to some degree universal: these are the terms in which judgments of value and inferences based on value are framed.

There are three basic elements to a system of values. First, in order for objects, persons, and actions to have values, there must be principles that assign values to entities on various grounds, i.e. rules in which value appears in the consequent of the rule ("If such-and-such takes place, then such-and-such a value is assigned."). These "input rules" are the entryways into the value system. Second, just assigning values is of little use unless values have some effect on behavior. So there must also be principles that favor performing certain actions on the basis of values, i.e. rules in which value appears in the antecedent of the rule – the "outputs" of the value system ("If such-and-such an action has such-and-such a value, then do it"). Third, in between input and output there may be many inferences that involve values in both the antecedent and consequent, that is, reasoning internal to the value system.

An example of such an internal accounting system may make the idea clearer. Consider the conceptual status of points in a game (a case discussed by Searle 1995). Such-and-such a physical action in the game leads to the assignment of so-and-so many points to a player: this is the input into the system. Within the system, points are totaled by addition as they are assigned; this is an operation that has no physical counterpart. The output of the system is the rule that says that the winner is the player with the most points at the end (or, depending on the game, the player with the least points, or the player to reach a prescribed total first). The points therefore serve as an inferential intermediary between the actions in the game and the outcome; without this intermediary, the outcome could not be determined. The points have no significance except within the context of the game: it is senseless to say, out of the blue, "I have 35 points," as if it were like "I have a new car."

Values are more complex than points for a variety of reasons. First, a value has two dimensions: a *valence* (good, i.e. positive, or bad, i.e. negative) and a *magnitude* (better or worse). Except in the case of monetary value, the magnitude is not a numerical quantity, but a relative quantity, perhaps measured by the basic mammalian magnitude system (Dehaene 1997, Hauser 2000). Thus values can be compared and combined approximately, but there is a lot of room for slop (Weber's Law error).

A second source of complexity in values is that there turn out to be several different sorts of value, each of which pertains to different entities and plays its own role in rules of inference. My impression is that crossdisciplinary discussions of value have often foundered because psychology deals primarily with one sort, economics with another, and moral philosophy with yet another, and ordinary language conflates them, calling them all *value*. A third source of complexity is that most of these sorts of value appear in two different versions, which I will call the *objective* and the *subjective* versions. In the objective version, the judger attributes value to something in his/her conceptualized world, e.g. *X is of value*. In the subjective version, the judger attributes value to something in his/her conceptualized world, relative to some observer (which may be the judger him/herself): *X is of value to Y/to me*. A final difference between values and points is that the rules of inference internal to the value system admit many more possibilities than simple addition. In working these rules out, I hope to establish that they are intuitively plausible, since after all they are meant to capture basic generalizations about how we reason with values.

2. Six kinds of value

This section will outline the dimensions of the value system, so as to set a context for the more detailed discussion to follow.

2.1. *Affective value*. The first type of value might be called *affective value* (or *A-value*). An event or situation has affective value for someone if it has a positive or negative effect on them, if it yields pleasure or suffering, a benefit or a cost. A simple expression of A-value is *good/bad for so-and-so*, as in (1).

(1) a. Eating your dinner will be good for you.
    b. Being overweight is bad for Max.

Obviously, the same event may simultaneously be beneficial to one person and harmful to another; for example the action of *revealing the name of the thief* might be good for the victim and bad for the thief. Hence

affective value arises not from the event per se but from its effect on individual participants in it.

Affective value is the type of value most connected to biological issues. Questions of preference, likes and dislikes, and approach and avoidance can be couched in terms of the A-value of the situations and actions involved. Thus this is the kind of value apparently of greatest concern to psychologists, e.g. Herrnstein 1993 and Mandler 1993.

There are expressions of affective value which do not name a participant, for instance (2).

(2) a. Drinking milk is good.
    b. Being overweight is bad.

Here the affected participant is an implicit generic individual identified with the subject of the generic subordinate clause: a close paraphrase is (3).

(3) a. Drinking milk is good for one/for people/for you.
    b. Being overweight is bad for one/for people/for you
       (where *you* is the generic personal pronoun, as in
       You don't see ads for DeSotos much anymore.)

When the implicit generic individual is implicit, as in (2), the assertion of value presents itself as an objective property of the generic event or situation.

I'll call examples like (1) expressions of *subjective* A-value, and those like (2) expressions of *objective* A-value. Now comes an absolutely crucial point: *Objective A-value is still not value in the world, independent of observers.* You may think that it's objectively good to drink milk, and I may not. But each of us conceptualizes this value as a property of drinking milk. What makes the value objective in the present sense is that we are in disagreement about the value of drinking milk independent of any particular person. By contrast, if you think drinking milk is good for Bill, and I think it's not good for Harry, i.e. if we're talking in terms of subjective A-value, we have no disagreement. Importantly, judgments of objective A-value do not rely on Theory of Mind, empathy, or identification with the other, whereas judgments of subjective A-value do, or at least purport to.

2.2. *Resource value.* A second type of value is what I will call *resource value* (or R-value). An object has a resource value if it is *good for someone to have*; a simpler expression is just that the object is *valuable.* One reason something may be good to have is that it offers the potential (or affordance) for an event with affective value. For a simple case, food has R-value because it offers the potential of being eaten, which is in turn an action of A-value to the eater. Similarly, a house has R-value because it offers the potential of being lived in. Another reason something may be good to have is that it adds to the esteem of its owner, as in the R-value of a famous painting. For another prominent case, money has R-value to its holder because it offers the potential of being exchanged either for other objects with R-value or for the performance of actions with A-value to the holder. Economists, whose basic data are exchanges, are therefore most concerned with resource values and those A-valued actions for which exchanges for R-valued objects are possible, e.g. labor and services (e.g. Scitovsky 1993 and Akerlof and Yellen 1993).

Resource value, like affective value, comes in subjective and objective varieties. All the examples so far are objective, in the sense that the object has simply has R-value rather than R-value *to so-and-so*. The subjective/objective contrast is hard to express using the word *good*, but it turns up in expressions like (4).

(4) a. This piece of land is very valuable/worth a lot.
[objective]
b. This piece of land is very valuable/worth a lot to Harry.
[subjective]

In (4a) the sense is that anyone will value the land highly; (4b) leaves the question open of whether it means anything to anyone else.

Resource value is defined in terms of its affordance for A-valued actions. But there is a secondary interaction between R-value and A-value. To the extent that having available resources reduces anxiety, the situation of having things with resource value can itself be of affective value: *having stuff is good for you; lacking stuff is bad for you.* Of course, the strength of this interaction varies from person to person.

2.3. *Quality.* An object or event can be valued in terms of its quality relative to other objects or events of the same type. I'll call this *Q-value.*

(5) a. This is a good/terrible computer. [object]
b. That was an excellent/miserable back dive. [action]

Typically, when objects are rated for quality, it is in terms of a particular action for which the object is to be used.

(6) a. This spatula is good for frosting cakes with.
b. This book is good for learning about deconstructionist phonetics.

When the *for*-phrase is absent (*a good spatula, a good book*), there is still an implicit purpose, as observed by Katz 1966 and Pustejovsky 1995. The default interpretation is that the object in question has quality with respect to performing its *proper function* – what the object is *for* (in the sense of Millikan 1984, or the concept's *telic quale* in the sense of Pustejovsky 1995). A good spatula is one that is good for scraping and spreading viscous materials (usually food-related), and a good book is one that is good for reading.

The adjective *excellent* is natural in expressions of Q-value; by contrast, it is somewhat awkward in expressions of A-value:

(7) a. This spatula is excellent for frosting cakes with. [Q-value]
b. ?? Drinking milk is excellent for you. [A-value]

An extension of Q-value concerns the use of some object for the function normally played by something else (Aronoff 1980). Expressions like (8), in particular the ... *makes a good X* construction, are characteristic of this reading.

(8) a. This rock is/makes a good table.
b. This table is/*makes a good table.

Alternatively, the purpose may be inferred from conversational context: one might for instance say *THAT cloud's good* in the context of comparing clouds for their resemblance to cows.

2.4. *Prowess*. Related to Q-value, there is a type of value that rates quality of an individual's performance, e.g. (9a). A second syntactic form, (9b), parallels the attribution of Q-value to artifacts (9c). We might call the sense in (9a,b) *prowess* or P-value. Note that like quality, prowess can be expressed by using the adjective *excellent*.

(9) a. Harry is good/excellent (at singing). [Prowess]
b. Harry is a good/excellent singer. [Prowess]
c. This is a good/excellent knife. [Quality]

And, with proper contextual support, there is yet another reading, in which Harry is construed as a resource: 'Harry is a good choice for the job.' Like Q-value, prowess is in effect an affordance for some task.

2.5. *Normative value.* A more complex kind of value is *normative value* (or N-value), which concerns conformance to social norms. Among the subvarieties of N-value are moral/ethical value, religious value, and valuation according to standards of etiquette (manners, politeness, etc.). Among these, moral/ethical value is the main sort of value of interest to philosophers, e.g. Stich 1993, Harman 2000. Unlike the previous sorts of value, this is strongly situated in the social domain: it not only has to do with people, but with people in the context of social interaction.

The subvarieties of normative value share a great deal of their linguistic expression and often apply in similar ways to similar situations, but they can still be teased apart into separate subdomains. For example, it is possible for a highly moral person to have bad manners; and conversely, a person with exemplary manners may well still be deeply immoral. Similarly, one can have a moral/ethical code independently of religion (think of honor among thieves, and perhaps desert traditions of hospitality); and many religious codes such as principles for performing rituals hardly fall in the moral/ethical domain. A particular action may have conflicting consequences in different normative domains. A classic case is the evil landlord in the melodrama, who is foreclosing on the poor widow in exercise of his contractual right, but who in so doing is acting in violation of the moral code. In the other direction is nonviolent civil disobedience along the lines of Gandhi and Martin Luther King, which violates the legal code but is in conformance with what is taken to be a higher moral code. More generally, my sense is that at best, explicit legal and religious codes are intended as imperfect incarnations of a more inchoate sense of morality; at worst, of course, they can be used to legitimate the raw exercise of power.[2]

[2] John Mikhail (2000) has stressed the way issues in legal codes mirror intuitive judgments of morality.

On the other hand, there are purposes for which it makes sense to lump all these subvarieties of normative value together, rather than, say, trying to isolate morality from "mere social convention." For instance, cultures differ in what they themselves consider to be morality as opposed to social convention, particularly with respect to issues such as sexual mores and slavery. Moreover, social convention and morality are inextricably intertwined. Consider for instance the Ten Commandments, where alongside the moral dictate "Thou shalt not kill" is what looks like a social convention: "Keep the Sabbath"; nevertheless, according to the Book of Exodus, violation of either one is punishable by death.[3] Thus the system of norms as a whole seems a more ecologically appropriate object of study. For the purposes of this article, I will therefore not distinguish the subdomains.

Characteristic expressions of normative value are shown in (10). Those in (10a,b) might be called expressions of *action-focused* N-value; that in (10c) is *person-focused* N-value. Here, unlike the other kinds of value, the words *right* and *wrong* are available in addition to *good* and *bad*.

(10) a. It is good of Harry to wash the dishes without being asked. [action-focused]
b. Taking cookies without asking is bad/wrong of Harry. [action-focused]
c. Harry is good/right to wash the dishes without being asked. [person-focused]

The attribution of N-value, as with A-value, is again focused on a relation between a person and an event. However, this time the event must be something that the person does intentionally, as in (11a) – it cannot be something that happens to the person, as in (11b,c), or that the person does accidentally, as in (11d). By contrast, A-value can be ascribed to such situations (11e-g).

(11) a. Washing the dishes is good of Harry.
b. *Being overweight isn't good of Harry.
c. *Being appointed chairman was good of Harry.
[except if this means he was good to *allow* himself to be appointed chairman, which is an intentional action on his part]

[3] Exodus XXI, 12 for murder, Exodus XXXI, 15 for working on the Sabbath.

d. *Getting lost on the way home wasn't good of Harry.
e. Being overweight isn't good for Harry.
f. Being appointed chairman was good for Harry.
g. Getting lost on the way home wasn't good for Harry.

There are two ways that attributions of N-value can be made more general. One is to omit the person from an action-focused attribution of N-value, as in (12). This implies that it would be good or bad of *anyone* to perform the action.

(12) a. It is good to wash the dishes without being asked.
b. It is bad to kill people.

The difference between (10) and (12) is that (10) is a relation between an actor and an act, while (12) is simply a valuation of a generic act.

The second way to bleach out the relational character of N-value is to omit the action from a person-focused expression of N-value, so that N-value is ascribed simply to a person, as in (13a). The sense is then that the person's generic actions, whatever they may be, are of N-value. This manipulation is impossible with A-value: a paraphrase like (13b) in terms of A-value makes no sense.

(13) a. Harry is good. [= 'Harry does things of positive N-value']
b. Harry is good. [≠ 'things happen of positive A-value for Harry']

(13) shows that *good* can be used for both A-value and N-value (as well as Q-value). There is a parallel ambiguity with *should*. The "predictive" sense (14a) does not have to do with values. But there are two readings that do involve values: there is a "prudential" sense (14b), which expresses affective value, and a "normative" sense (14c), which expresses normative value. The same possibilities exist with *ought to*.

(14) a. The bus should/ought to arrive soon. [predictive]
b. You should/ought to take an umbrella in case it rains.
[prudential: 'the A-value of taking an umbrella is positive'; 'it would be good for you to take an umbrella']
c. You should/ought to wash the dishes.
[normative: 'the N-value of washing the dishes is positive'; 'washing the dishes would be good of you']

There are interactions between N-value and A-value. For example, actions that are good *of* you (N-value) are often good *for* someone else (A-value), e.g. acts of charity; conversely, gratuitous violence is bad *of* the perpetrator and bad *for* the victim. Another interaction is that it may *feel* good/bad (A-value) to *do an act* that's good/bad (N-value): that is, performing an act with N-value may result an accompanying secondary A-value. In common language, we say that someone who experiences this interaction "has a conscience." In particular, when the value is negative, I think the secondary A-value is called *guilt*.

However, such synergy between N-value and A-value is not invariable. Consider for example the performance of religious rituals, which are of positive N-value but don't really benefit anyone directly. Similarly, illicit sex may be of positive A-value but is definitely of negative N-value (that's why we call it illicit).

All the description of N-value so far has been in "objective" terms: such-and-such an action is normatively good, such-and-such a person is normatively good. There is also a subjective version, where other people's opinions are being compared with one's own.

(15) To Joe, Harry's a good guy (– but not to me).

This is tricky to express in English, requiring the prepositional phrase *to Joe* at the beginning of the sentence, because the most expedient grammatical expression, *Harry is good to Joe*, is taken up by another meaning: 'Harry acts in a manner that's good for Joe (of positive A-value to Joe)'.

2.6. *Esteem*. A final notion of value also pertains specifically to persons and might be characterized as *esteem* (I'll call it E-value). Esteem seems to be a composite of person-focused normative value, prowess, status in the dominance hierarchy, wealth (accumulation of R-value) and perhaps other factors such as simple personal attractiveness. Group membership also plays a role: members of groups other than one's own are by default accorded less respect, as are members of generally low-status groups. However, particular individuals of low-status groups who have other highly respected qualities (wealth, prowess, leadership, etc.) may be accorded greater respect than the default (e.g. Al Capone). This shows the interactive, composite character of E-value.

Objective and subjective versions of E-value appear in (16a,b) respectively.

(16) a. Harry is prestigious/well-respected. Harry has a good reputation. [objective]
b. Joe respects Harry. [subjective]

Unlike prowess, esteem is clearly a socially rooted value. In fact, esteem or reputation is an important locus where the values system outputs to behavior. As Fehr and Fischbacher 2004 suggest, performing normatively good actions is one way to enhance esteem. And an individual who is esteemed or respected is sought out by others for cooperative interactions of mutual benefit. Hence esteem might be thought of as a sort of personal resource value: respect/a good reputation is (A-)good to have.

Table 1 sums this all up. The rest of the article deals mostly with A-value, R-value, N-value, and E-value.

3. Some inference rules: A route through the system

Let's now be a bit more specific about the principles behind some of the relations among different kinds of value. For a first step, we need an entry into the system, an inference rule whose antecedent is not about values but whose consequent is.

(17) Input rule:
An event E which affects person Y positively/negatively is of positive/negative A-value to Y.

E can just be an event that Y experiences, such as a delicious meal (positive) or the onset of a disease (negative). However, the more interesting case is when E is an action on someone else's part, in which case Y is a *beneficiary* (positive) or *patient* (negative) of the action: it's A-good to be helped and A-bad to be victimized. This is one input route into the value system.

Table 1

Varieties of value. *Entity to which value is ascribed is in* ***<u>underlined bold</u>***.

Type of value	**Applies to ontological type:**	**Subjective version**	**Objective version**
Affective (A-)value	events, situations	Situation **<u>X</u>** is good for Y	Situation **<u>X</u>** is good
Resource (R-)value	objects	Object **<u>X</u>** is valuable to Y	Object **<u>X</u>** is valuable
Quality (Q-value)	events		Event **<u>X</u>** was a good one
	objects		Object **<u>Y</u>** is good for doing X
Prowess (P-)value	persons		**<u>Y</u>** is good at doing X
Normative (N-)value	action-focused, relational:	To Z, doing **<u>X</u>** is good of Y	Doing **<u>X</u>** is good of Y
	action-focused, absolute:	To Z, doing **<u>X</u>** is good	Doing **<u>X</u>** is good
	person-focused, relational:	To Z, **<u>Y</u>** is good to do X	Person **<u>Y</u>** is good to do X
	person-focused, absolute:	To Z, **<u>Y</u>** is good	Person **<u>Y</u>** is good
Esteem (E-)value	persons	X respects **<u>Y</u>**	**<u>Y</u>** is prestigious

Now let us look at some rules that manipulate values within the system – the parallels to adding up points in a game. An intuition that emerged in the previous section can be codified a little more precisely as (18).

(18) If X acts in a way that is A-good/bad for Y, then that action is N-good/bad of X.

More colloquially, it's (N-)good to be nice to people, and (N-)bad to be mean to people.

Next comes an interesting step. The basic intuition is that if you do something N-good, you're an N-good person; if you do something N-bad, you're an N-bad person. This principle relates action-focused N-value to person-focused N-value. It is a rather peculiar principle, but, strikingly, it conforms to intuition.

(19) If X has performed some action that is N-good/bad, then X is N-good/bad.

A slightly more sophisticated version of this rule might say that good acts add to your total "goodness", and bad acts subtract from it.

(20) If X has performed some action that is N-good/bad, then X is N-better/worse for having done so.

As a result of (20), a person's normative value at any moment is related to his or her history of performing normatively valued actions. However, the result is not necessarily a sort of mathematical sum, in that the contribution of any particular action is subject to certain wild-card effects. We intuitively recognize these effects with statements such as: "I have now performed so-and-so many good actions. Is that enough to make up for all the bad things I've done?" "Even though what you just did was horrible, I'm not going to hold it against you"; "That one horrible thing you just did has wiped out my whole good opinion of you" (this one has the flavor of contamination, along the lines of Nemeroff and Rozin's (2000) proposals on contamination: a speck of bad stuff, say a cockroach leg, contaminates a whole lot of food). Such common statements show that rule (20) can be applied in highly subjective fashion; exactly what determines its outcome under different circumstances is a question far beyond the scope of the present exploration.

Just another reminder: rules (19-20) are not meant to treat X's normative value as a free-standing thing unto itself. We are not describing how X acquires "real" normative value, but rather how the judger conceptualizes X acquiring "objective" normative value. Lest this should seem a problematic stance on normative value, it should be recalled that this is the very same stance taken in studying vision, where perception is described in terms of the perceiver developing a conceptualization of the objective "world out there", in response to certain inputs to the perceptual system. In both cases we are concerned with the individual's sense of what is real. And from this point of view, goodness is as real as color.

The contribution of normative value to esteem might be encoded as follows:

(21) X's E-value (value as a person) is a function of X's N-value, X's P-value, X's wealth (accumulated R-value), X's dominance, X's attractiveness, and. . . .

The ratio of importance among these factors may be highly variable and context-dependent. (Being *notorious* is having a high E-value despite a negative N-value.) Moreover, since we are dealing with the analog magnitude system in measuring values, the idea of mathematical exactness in measuring E-value is far too specific.

One of the factors in rule (21) is commonly expressed by "You're known by the company you keep": people adjust an individual's esteem in terms of the esteem of the people he or she associates with ("Oh, my! YOU know Noam Chomsky?" – which may express a positive or negative effect on the addressee's esteem, depending on the speaker's opinion of Chomsky). This factor too is related to Nemeroff and Rozin's (2000) notion of contamination: a person's good or bad "essence" "rubs off" on people he or she is associated with. As mentioned in section 2.6, group membership is also one of the factors in rule (21): by default one accords members of one's own group higher esteem than members of other groups, and one accords members of high-status groups (whatever their other qualities) default higher status.

As pointed out in section 2.6, being esteemed isn't enough: it has to do you some good. Here is a somewhat more formal statement of the intuition in question.

(22) If Y is of high/low E-value (to X), it's A-good/bad (for X) to associate with Y.

One particular consequence of (22) concerns the esteem derived from group membership: people prefer to associate with and do business with members of their own group *and* with members of high-status groups, all things being equal. This seems just the right result.

Rules (18)-(22) are all inside the value system; they are the counterpart of adding up points in a game. In order for all this to impact on behavior, we also need an "output" rule in which the value system affects action. About the simplest possibility is that one should prefer to do actions that are better for one; that is, the A-value of a potential action affects its preferability. In order to affect one's action, this rule has to be stated in terms of the relative value of one's contemplated actions. The outcome of the rule has to be procedural, the making of a choice.

(23) "Do what's better for you":
Faced with a choice between two or more actions, do the one with highest A-value.

(23) compares objective A-values of actions and leads to a choice. The subjective version is more reflective. You can't decide on someone else's action. The best you can do is infer what they'll choose. (24) relativizes (23) to any actor: it is the Theory of Mind's description of what's going on in other people (and oneself) when their action is determined by (23).

(24) "People decide to do what's better for them"
Faced with a choice between two or more actions, X will do the one that is A-best for X.

This gives us a whole loop in and out of the system: you observe X doing something nice for somebody (17), you infer that this is good of X (18), and therefore that X is good (19/20). From this you gain respect for X (21), and so you infer that it is good for you to associate with X (22). Therefore you choose to associate with X (23). There are many

other routes through the system. For instance, in addition to rule (18) ("It's good to be nice to people"), there are many culture-specific attributions of normative values to actions, which serve as input rules to the value system: "It's normatively bad to dress such-and-such a way", "It's normatively bad to eat such-and-such", "It's normatively bad to show attraction to people of the same sex", and so on. These contribute to normative judgments of people who engage in these acts. On the output end, the system eventually has to ground out in action, perhaps always through a judgment of A-value that leads to rule (23), "Do what's better for you."

It's also worth mentioning that judgments of relative A-value are highly context-dependent. For instance, "giving in to temptation" is valuing an immediate gratification (an action with positive A-value) in preference to a potentially greater A-value to be realized over a longer time-span. (See Stevens and Hauser 2004 for discussion of this issue of "temporal discounting.")

4. Moving between subjective and objective

In an important sense, "subjective" value is more true to life. Being of value is fundamentally a relation between an object and a perceiver: there is no value without valuers. Yet experientially, "objective" value is every bit as valid: we take certain actions to be morally repulsive and certain people to be of high esteem, and this is not a fact about one's perception of them. In fact, an important fact about moral systems is that they are conceived of as objective, universal, and timeless (Berger and Luckmann 1966), which is why the term "moral relativism" is taken by many to be self-contradictory or tantamount to "amoral." Moreover, values are usually *taught* in their objective form: "This is what we (the group) do; *this* is good; *that* is bad."

Two obvious questions are: Why should we have distinct subjective and objective systems in cognition? And what does one have to do with the other? Part of the answer came up already in section 2.1. On one hand, value is experienced as a property of an action, just as objective as its duration or loudness. One's own contribution to the judgment of value is completely transparent, just like one's own contribution to

the judgment of duration or loudness. On the other hand, it is important to be able to account for individual differences in values, and this is what the subjective system allows us to do: this object is worthless to you and valuable to me, this action is good for you and of no import to me. However, it takes Theory of Mind (including of my own mind) to recognize these differences – always a cognitive stretch. Still, we may be rightly suspicious of the seeming redundancy, with identical predicates in the two systems, differing only in whether they have an experiencer argument.

In practical reasoning, we jump readily between the two systems. Something like (25) seems to be the appropriate rule of inference.

(25) Objectification and Subjectification
Object or event Y is A-good/bad for X iff$_{default}$ Y is A-good/bad.

First let us read the rule from left to right, using myself as X: If I like Y, or Y is good for me or valuable to me, then Y is objectively good or valuable. That is, my own judgments by default warrant a judgment of objective value. Alternatively, suppose I don't know anything about Y and I find out that you like it, or it's good for you or valuable to you. Then, using you as X in (25), I can conclude by default that Y is objectively good or valuable. In other words, from left to right, (25) represents the objectification of value.

Why should it be important to arrive at an objective value? The reason is that then the rule can be read from right to left to predict someone else's reactions to Y. If it's good or valuable, then it's reasonable to believe it will be good or valuable for you and me and anyone else. However, subjective values are necessary when I want to deal with the discovery that you and I react differently to an object. In general, I can't predict the object's subjective value to you without evidence about your reactions. My own reactions are of no relevance on this plane. Without such evidence, I fall back on (25) to predict your reaction.

So in practice, we slip between the two systems as convenient. We strongly prefer the more predictive objective system when possible, but we can easily drop into the subjective system when we have evidence of difference. This is hardly logical reasoning. But it's what we do.

The plot thickens, because there is a value judgment that arises from a conflict between the objective and subjective systems. Another instan-

tiation of the contrast is *Y is the case* (objective) versus *X believes Y* (subjective). And there is a pernicious inference rule along the lines of (26).

(26) "It's bad to be wrong"
If Y is the case and X believes NOT Y, then X's E-value goes down.

(26) raises a difficult practical problem: when I disagree with you, the question arises as to who has control of objective value and truth. One possibility is that I trust your judgment, say because you're an authority figure. Then (26) leads to the conclusion that there is something wrong with *me*, and *my* self-esteem goes down. The more standard situation, though, is when *I* am "in possession of objective truth and values", and I thereby think less of you. This is typically the case when two cultures encounter one another and each characterizes the other as uncultured, savage and lacking in values. There is no need to recount the unpleasant consequences, between generations, between religions, between religion and science, between the sciences and the humanities, even between subcultures of a discipline. The only way for dialogue to take place is if both protagonists are capable of switching into the subjective system for their own judgments as well as for the other's.

5. Fairness

The notion of A-value makes it straightforward to state a version of *fairness*: an action is fair if it is equally good or bad for everyone, i.e. if its A-value (positive or negative) to each of the individuals it affects is the same.

(27) Y acts fairly toward $X_1, \ldots, X_n$:
For all X_i, X_j, the A-value of Y's act to X_i equals the A-value of Y's act to X_j.

This can be played out in various ways: Y's act may be a single action that impinges on everyone at once. Or it may be a composite of multiple sub-actions at different times, each impinging on a different individual: if you do such-and-such to X_1 this time, you'd better do the same thing to X_2 the next time.

A further layer of objectivity vs. subjectivity emerges in the application of (27). Consider for example an action where Y is distributing resources to the group. In the simpler (objective) construal, each person in the group receives the same amount of resources. In the more sensitive, subjective construal, the differing needs of individuals are taken into account, that is, the subjectivity of A-value is taken seriously. Similarly when the action is collecting resources from the group: the simpler construal is a uniform tax, and the more sensitive construal is a progressive tax. (Which counts as "fair"? It depends how you frame the issue.) More generally, principle (27) undergirds the notion of "equality under the law", where what's often at issue is who counts as an X: everybody, members of one's own group, only men, only white men, only white men who own property, and so on.

Another prevalent pattern is distribution by rank: an action fits this pattern if the higher your rank, the better treatment you get.

(28) Y acts according to distribution by rank among $X_1, \ldots, X_n$:
For all X_i, X_j: If X_i outranks X_j, then the A-value of Y's act to X_i is greater than the A-value of Y's act to X_j.

In turn, "outrank" stands in for an inequality of value – either prowess (P-value), dominance, virtue (N-value), or general esteem (E-value). This mode of distribution is appropriate for awarding honors and prizes (see the next section). But it's of far broader social application: the top dog receives the best chair, the best mate, the most food, the plum work assignment, and (if he transgresses) the most lenient punishment; those on the lowest rungs get the fewest resources, the most unpleasant work, the most severe punishment, and so on. And of course this principle is amply attested in social animals.

In turn, the action patterns described in (27)-(28) can be accorded N-values: under such-and-such circumstances, distributing fairly has a positive N-value; under such-and-such other circumstances, distribution according to rank has a positive N-value. Fiske 1991 argues that a great many cultural differences fall under such normative values: what is the *right* way to divide up resources and responsibility, depending on the task and the participants?

6. Reciprocity

6.1. *Reciprocation, retaliation, and restitution.* Next let us consider the relation between two actions expressed by a certain use of the preposition *for* in English.

(29) a. Susan praised her son Sam **for** behaving nicely.
b. Fred cooked Lois dinner **for** fixing his computer.
c. Susan insulted Sam **for** behaving badly toward her.
d. Lois slashed Fred's tires **for** insulting her sister.

These sentences describe situations in which someone does something "in return" for someone else's action. Those in (29a,b) describe actions with positive values; those in (29c,d) describe actions with negative values. Such acts of reciprocity can felicitously take place only with another person, an entity that can be regarded as having values and responsibility. One cannot sanely punish one's car for getting a flat tire.

If we switch around the actions among the examples in (29) we get sentences that sound odd or perhaps ironic.

(30) a. #Susan insulted Sam for behaving nicely.
b. #Lois slashed Fred's tires for fixing her computer.

This shows that we expect a positively valued action in return for a positively valued action, and a negatively valued action in return for a negatively valued one. Reciprocity is furthermore sensitive to the (analog) magnitude of values as well: we find it odd if the two actions related by *for* do not match in quantity. The sentences in (31) convey some of this oddness:

(31) a. #Fred cooked Lois dinner for saying hello to him.
b. #Fred cooked Lois dinner for rescuing all his relatives from certain death.
c. #Fred slashed Lois's tires for eating too little at dinner.
d. #Fred slashed Lois's tires for murdering his entire family.

In (31a) and (31c), we sense Fred as overreacting, as doing something unwarranted in return for Lois's action; in (31b) and (31d), we sense

him as underreacting, as doing something that is not nearly enough to recognize the significance of Lois's action.

The intuition, then, is that a reciprocal action calls for rough equivalence of value between the two actions. Crucially, a particular action may be of different value to the participants: you may not even know that your action helped or harmed me. Thus the principle of reciprocation must be stated in terms of the particular person the action affects, i.e. subjective A-value. (32) is a first approximation. For the moment, we take the relevant notion of value to be affective value. Like many inferences with value, this one is defeasible.

(32) If Y acts in return for X's acting, Y's act is (defeasibly) as A-good/bad for X as X's was for Y.

The logic of reciprocity expressed by (32) is a cognitive elaboration of reciprocal altruism: "I'll scratch your back because you scratched mine."[4] I leave open how much of its detail can be attributed to nonhuman primates (not to mention elephants and bats).[5] What strikes me as particularly human, in any event, is the broad generality of the actions available for reciprocation.

Because (32) is neutral as to whether the value in question is positive or negative, it serves not only to express reciprocal altruism but also retaliation (or retribution). In this case, the equivalence of values amounts to a more or less formal statement of "the punishment fits the crime": this helps guide what responses are appropriate in retaliation for harmful actions.

There is another kind of reciprocation for negative actions, illustrated in (33).

(33) a. Fred cooked Lois dinner (to make up) for having embarrassed her in public.

[4] I use this slightly nontraditional phrasing deliberately. The usual phrasing, "I'll scratch your back and you scratch mine" (e.g. Dawkins 1976) suggests an explicit agreement on a joint action (in the sense of Bratman 199x, Searle 1995, and Clark 1996). An agreed-upon joint action has different inferences from reciprocal altruism: it is *nice* to reciprocate, but one is *obligated* to perform one's role in a joint action.

[5] See Stevens and Hauser 2004 for arguments that reciprocal altruism is much less common among nonhumans than usually thought.

b. Fred brought Lois flowers (to make up) for forgetting her birthday.

Here the perpetrator of the negative action is performing a positive action in *restitution*, righting the balance. Again there has to be a rough equivalence of value: notice the weirdness of (34).

(34) a. #Fred gave Lois his vast fortune (to make up) for forgetting her birthday.
(overreaction)
b. #Fred brought Lois flowers (to make up) for killing her whole family.
(underreaction)

Thus the rule for restitution might be stated as (35). Unlike reciprocity (32), it is not neutral to the valence of the original action: there is no counterpart that sanctions doing someone ill because you have been nice to them. Thus the principle must explicitly encode the valence of the original action.

(35) If X acts in restitution for having done something bad for Y, the restitutive act is (defeasibly) as good for Y as the original act was bad for Y.

Notice that if X does something harmful to Y, one sort of retribution for Y is to force X to perform restitution, perhaps through the intervention of the authority of the group. However, this sort of retribution is not equivalent to raw retaliation. For one extreme case, "Nothing you can make the murderer do will bring my son back," that is, restitution is impossible, even though retaliation might be (e.g. killing the murderer's son).[6]

Finally, (31) and (35) are stated as inferences from reciprocation to the value of the actions in question. However, this is not enough: one *should* reciprocate actions that benefit one and one *should* perform restitution for having harmed others, that is, there is a normative value

[6] The distinction between these two in conceptual development is noted by Piaget 1932; he claims the notion of restitution is established later than that of retribution.

attached to these actions. How this value plays out – what actions should be reciprocated and restituted, and what counts as appropriate reciprocation and restitution – is variable among cultures and subcultures. But the overall principle seems universal. (36) is one way to state these principles. They are special cases of rule (18), "It's (N-)good to be nice to people."

(36) a. If X does something of positive A-value to Y, then it's N-good of Y to reciprocate.
b. 'If X does something of negative A-value to Y, then it's N-good of X to perform restitution.

Depending on the circumstance, there are three possible negative counterparts of (36a). One is "you should retaliate", generalizing the left-hand side of (36a) to negative valence, as in (37a). A second is "it's all right to retaliate", which differs from (37a) only in that the normative value, instead of being positive, is neutral. The third is "you should turn the other cheek", in which case the right-hand side places a negative N-value on retaliation, as in (37c).

(37) a. If X does something of negative A-value to Y, it's N-good of Y to retaliate.
b. If X does something of negative A-value to Y, it's N-neutral of Y to retaliate.
c. "Turn the other cheek"
If X does something of negative A-value to Y, it's N-bad of Y to retaliate.

Not only do these principles conflict with each other, but in addition (37a) conflicts with rule (18), while (37c) does not. Again, a lot of cultural variation arises from how these rules are understood to apply in practice. I suspect that it's partly in the service of negotiating such conflicts that more explicit moral and legal codes arise.

There are a number of ways for slippage to be introduced into reciprocation. Perhaps the most pernicious arises from a general cognitive bias toward overestimating harm to oneself and underestimating harm to others. Hence, if I retaliate against you, you judge the harm done to you to be greater than the harm you originally did to me. You are

therefore motivated to even the scores by retaliating further, leading to escalating cycles of violence.

6.2. *Honoring and shaming.* The rules of reciprocation so far have been stated in terms of the A-value of actions to their participants. Another application of reciprocal *for*, involving normative value and prowess, appears in the examples in (38).

(38) a. Joe praised Sue **for** saving the drowning child.
b. The club honored Sue **for** her service to the community.
c. Sue was awarded a prize **for** winning the race.
d. The fans cheered Sue **for** hitting a grand slam in the ninth inning.

In these cases, Sue has done nothing to benefit the individual or organization that is acting reciprocally. Rather, she has done something that has raised her normative value (38a,b) or that has demonstrated prowess (38c,d), both of which contribute to the esteem in which she is to be held (by rule (21)). In particular, the reciprocal acts in (38b,c,d) are performative actions of esteem-creation (honoring, thanking), which have the effect of making the esteem public (and ostensibly objective). Such statements of esteem are naturally of positive A-value to Sue as well; and the resource value of increased esteem is a further benefit.

The negative counterpart of (38) is shaming: humiliation and community-sanctioned punishment for normative transgressions – even for moral transgressions that do not directly harm anyone else. Two examples of this counterpart are in (39).

(39) a. Sue scolded Bill for losing his wallet.
b. The fans booed Foulke for giving up a grand slam in the ninth inning.

A different negative counterpart is apology: restitution for harm done by offering expressions of self-humiliation before the injured party.

(40) is an attempt at stating these sorts of reciprocation, involving a very particular type of reciprocal action, expressing (or otherwise demonstrating) esteem.

(40) a. If X does something that increases X's N-value or P-value, it's N-good to express one's respect for X's E-value.
b. If X does something of negative A-value to Y, it's N-good of X to express negative self-esteem to Y (i.e. apologize) in restitution.

In these cases of reciprocation, it is hard to know what counts as equivalence in value between the original act and the reciprocal act. Perhaps the best one can do is context-dependent proportionality: first prize ought to be more valuable than second prize; greater praise should be accorded to someone who saves 80 lives than to someone who stops on the highway to help you fix your car; a bigger faux pas calls for more fervent contrition.

Note that etiquette (a type of N-value) demands that the recipient of one of the reciprocal actions in (38) respond "I don't deserve it; what I did was nothing." Why is this the case? The immediate cause seems to be a principle to the effect that you shouldn't think too highly of yourself. (41) gives two possible versions of this principle.

(41) a. "Don't have too high an opinion of yourself"
It's N-bad of X to have a subjective E-value much greater than zero.
b. "You shouldn't rate your esteem higher than it really is"
It's N-bad of X to have a subjective E-value higher than his objective E-value.

(41b) is related to (26), "It's E-bad to be wrong"; perhaps (41a) results from "playing it safe" in avoiding violation of (41b).

6.3. *Deserving.* Now we come to some very peculiar but pervasive reasoning about values. (42a) repeats the normative principle for reciprocation; (42b) is a different phrasing. In both cases the normative principle applies to the person Y helped by X's original action. However, the situation can also be construed in a stronger sense: not only would it be good for Y to reciprocate, Y has a sort of "moral obligation" to reciprocate, expressed perhaps as (42c). Now. looking at this situation from the point of view of X, we might express it as (42d,e).

(42) (Reciprocation)
If X does something of positive A-value to Y,
a. then it is N-good of Y to reciprocate. (= (36a))
b. then Y should reciprocate/reward X.
c. then Y owes it to X to reciprocate.
d. then X should be rewarded by Y.
e. then X deserves to be rewarded by Y. ("One good turn deserves another")

A counterpart for restitution is (43).

(43) (Restitution)
If X does something of negative A-value to Y,
a. then it is N-good of X to perform restitution. (= (36b))
b. then X should provide restitution.
c. then X owes it to Y to provide restitution.
d. then Y should be compensated by X.
e. then Y deserves compensation from X.

If the normative system condones retaliation, then the counterpart for retaliation is (44).

(44) (Retaliation)
If X does something of negative A-value to Y,
a. then it is N-good/N-neutral of Y to retaliate. (= (37a,b))
b. then Y should/may retaliate.
c. then Y is entitled to retaliate.
d. then X should be punished by Y.
e. then X deserves to be punished by Y.

I'm not going to try to formalize these inferences (or whatever they are) here. I bring them up, however, because they are so common in our reasoning, and they serve as a steppingstone to further normative conclusions that we typically draw, whether or not there is logical warrant for them. Taking the perspective of (42-44 d,e), the individual who deserves something, it is natural to drop the other individual out of the picture: it does not matter any more exactly who *owes* the moral debt. (45) illustrates.

(45) a. Deserved reciprocation:
If X does something A-good for Y, then X deserves to be rewarded.
b. Deserved restitution:
If X does something A-bad to Y, then Y deserves to be compensated.
c. Deserved retaliation:
If X does something A-bad to Y, then X deserves to be punished.

In the consequent clauses of (45), the reciprocal act need no longer be performed by the other character. Rather it may be performed by anyone – a generic actor will do.

We tend to go still further. Consider first reciprocation (45a). The fact that it is specifically Y that X is doing something good for is in a sense irrelevant to what X deserves. We can eliminate Y from the story by remembering that if X does something that is A-good for Y, it is N-good of X to do it (rule (18)). So instead of characterizing X's act in terms of its A-goodness for Y, we can characterize it simply in terms of its N-goodness, in which Y plays no essential role. This yields (46a). The counterpart for retaliation is (46b).

(46) a. Deserved reward:
If X does something N-good, X deserves to be rewarded for it.
b. Deserved punishment:
If X does something N-bad, X deserves to be punished for it.

The case of restitution is slightly different. In (45b), the person who deserves restitution is Y, the person affected by the original action. So what has to drop out of the left-hand side of the equation is X, the actor of the original action. What is left is some event in which Y is adversely affected, i.e. something A-bad *happens to* Y. Thus (47) is an appropriate form.

(47) Deserved restitution:
If something A-bad happens to Y, Y deserves to be compensated for it.

Since one's personal normative value is a cumulative function of the normative value of one's actions (rule (20)), one further shifty step applied to (46) akes us to (48).

(48) a. Good people deserve to be rewarded.
b. Bad people deserve to be punished.

Now it's not as though the steps leading to (46)-(48) follow from any sort of formal reasoning. But intuitively they're entirely seductive. And of course (46)-(48) are massively counterexemplified in the world: bad things happen to people all the time with no hope of compensation, wicked people frequently do very well indeed, and all too often "Nice guys finish last." Here is the existential "problem of evil." How is it to be resolved? Different traditions have different ways. Christianity's solution is to put off reward and punishment until the afterlife, conveniently linking up with the strongly held belief in the survival of the soul after death. Judaism tends to take the view that if something bad is happening to me now, it must be punishment for something bad I (or even my ancestors) did in the past. Hence the comedians' version of Jewish guilt: I must have done something wrong, and I'm sorry – but what was it?[7] Yet another solution is "virtue is its own reward" (i.e. it's A-good for you to behave in N-valued fashion), which gives up on reward coming from *outside*, and so in a way negates the spirit that leads to (48).

But who is going to carry out the acts of reward and punishment? The anticipated reciprocal acts can't depend on people, since they are intended precisely as the way of circumventing people's injustice in the real world. Enter gods: animate moral beings who lie outside the human sphere and who take care of righting the moral scales. This puts gods in the role of protectors, beings whom one can plead for justice and to whom one can express gratitude. Moreover, the rules of normative value dictate that one had better be nice to the gods as well, because if anyone is in a position to reward or retaliate, it's the gods. Thus the reasoning in this section leads to one of the important groundings for religion, one that is not as thoroughly explored as I think it deserves [sic] in recent work such as Boyer 2001 and Atran 2002 (though it is mentioned by Nemeroff and Rozin 2000).

[7] This also explains why many Jews lost faith during the Holocaust: nothing they or their ancestors had done could be bad enough to justify *this*.

7. What does this all mean?

To sum up the treatment here, the value system can be seen as an abstract calculating system that helps govern action. The calculations assign values to actions or anticipated actions, and use these values eventually to determine the value of individuals performing these actions. In turn, the value of these individuals is crucial in determining how to interact with them. The value system is multidimensional, in that there are at least six kinds of value, each with objective and subjective versions; and each type plays a different role in the system of inferences built on value.

Crucial to the present approach is that values are being approached in terms of a cognitive system: every individual capable of social interaction commands this system. What is necessary in learning a culture is acquiring the special rules that assign values to particular sorts of action, e.g. what it's good to do and what it's bad to do, which actions fall under moral prescription and which under (mere) manners.

Still, I can imagine the reader asking what the point is. Why translate lots of moral truisms into a relatively arcane system? I can offer three reasons.

The first reason is that it reveals the complexity of our value judgments and our reasoning about values. By attempting to express these judgments in systematic terms, with a relatively limited vocabulary, we can see the rich interrelationships among different notions of value. An important virtue of the present approach is that it acknowledges different kinds of value, each participating in the system in a different way. My impression is that previous approaches have been limited because they insist on a unitary notion of value, and because in many cases (especially in moral philosophy), they discount subjective value altogether.

Within the present approach, value is revealed as a complex conceptual system, rich in hierarchical abstract structure. In the context of how other cognitive systems are now understood – especially language – this should not be too surprising. This outcome is not undermined by the fact that value judgments are often quick and intuitive. In language, judgments of grammaticality and meaningfulness are quick and intuitive: the computational reasons for these judgments are deeply unconscious. In vision, judgments of spatial configuration and motion are intuitive

and present themselves to awareness as "what is the case in the world." Thus value judgments are of a piece with the rest of cognition. However, unlike linguistic and visual judgements, aspects of them are available to awareness as well – they lie on the borderline between intuitive and conscious reasoning.

A second reason that these explorations are valuable has to do with the connection between the value system and its linguistic expression. Part of the job of linguistic semantics is to explicate the meanings of words and phrases. In the domain of spatial language, linguistic semantics has benefitted from the attempt to develop formal analyses of the conceptualization of space, motion, force, and agency in terms of a limited conceptual vocabulary and combinatorial system, and large portions of language have been subsumed under such analyses. The present exploration has begun a similar undertaking with the parts of the vocabulary whose meanings incorporate notions of value.

A third reason for taking this approach seriously is that it offers the possibility of making many longstanding questions more precise. At the scale of the individual lifespan: What is the course of development of value systems in humans (a la Piàget 1932, Kohlberg 1981/84, Turiel 1983, Macnamara 1990, Premack and Premack 1994, Bloom 2004)? Over historical time: To what extent are value systems a functional outcome of what it takes to make a society work well (and interact well with other societies) (Fiske 1991, Jacobs 1994)? Over evolutionary time (Hauser 2000, deWaal 1996): How much of the basis of human value systems is innate? Of that, how much is part of our primate heritage, and how much is unique to humans? By attempting a comprehensive overview of the entire system, it's possible to pose these questions in a more ecologically sound context.

REFERENCES

Akerlof, George A., and Janet L. Yellen
1993 The Fair Wage-Effort Hypothesis in Unemployment. In Hechter et al. 1993, 107-134.

Aronoff, Mark
1980 Contextuals. *Language* 56: 744-758.

Atran, Scott
2002 *In Gods We Trust.* Oxford: Oxford University Press.

BARTH, FREDRIK
1993 Are Values Real? The Enigma of Naturalism in the Anthropological Imputation of Values. In Hechter et al. 1993, 31-46.

BERGER, PETER L. and THOMAS LUCKMANN
1966 *The Social Construction of Reality*. Garden City, NY: Doubleday.

BLOOM, PAUL
2004 *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human*. New York: Basic Books.

BOYER, PASCAL
2001 *Religion Explained*. New York: Basic Books.

BRATMAN, MICHAEL
1999 *Faces of Intention*. Cambridge: Cambridge University Press. [contains essays previously published between 1985 and 1998]

CLARK, HERBERT
1996 *Using Language*. Cambridge: Cambridge University Press.

DEHAENE, STANISLAS
1997 *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press.

DE WAAL, FRANS
1996 *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press.

FEHR, ERNST, and URS FISCHBACHER
2004 Social norms and human cooperation. *Trends in Cognitive Sciences* 8.4.

FISKE, ALAN
1991 *Structures of Social Life*. New York: Free Press.

GREENE, JOSHUA
2003 From neural 'is' to moral 'ought': What are the moral implications of neuroscientific moral psychology? *Nature Reviews/Neuroscience* 4, 847-850.

HARMAN, GILBERT
2000 *Explaining Value*. Oxford: Oxford University Press.

HAUSER, MARC
2000 *Wild Minds: What Animals Really Think*. New York: Henry Holt.

HECHTER, MICHAEL, LYNN NADEL, and RICHARD E. MICHOD (EDS.)
1993 *The Origin of Values*. New York: Aldine de Gruyter.

HERRNSTEIN, RICHARD J.
1993 Behavior, Reinforcement, and Utility. In Hechter et al. 1993, 137-152.

JACKENDOFF, RAY
1983 *Semantics and Cognition*. Cambridge, MA: MIT Press.

JACKENDOFF, RAY
2002 *Foundations of Language*. Oxford: Oxford University Press.

JACKENDOFF, RAY
to appear. *Language, Consciousness, Culture: Essays on Mental Structure*. Cambridge, MA: MIT Press.

JACOBS, JANE
1994 *Strategies of Survival*. New York: Vintage Books.

KATZ, JERROLD
1966 *The Philosophy of Language*. New York: Harper & Row.
KOHLBERG, LAWRENCE
1981/84 *The Philosophy of Moral Development* (vols. I and II). New York: Harper & Row.
MACNAMARA, JOHN
1990 The development of moral reasoning and the foundations of geometry. *Journal for the Theory of Social Behaviour* 21, 125-150.
MANDLER, GEORGE
1993 Approaches to a Psychology of Value. In Hechter et al. 1993, 229-258.
MIKHAIL, JOHN
2000 *Rawls' Linguistic Analogy*. Ph.D. dissertation, Cornell University.
MILLIKAN, RUTH
1984 *Language. Thought and other Biological Categories* Cambridge, MA: MIT Press.
NEMEROFF, CAROL, and PAUL ROZIN
2000 The makings of the magical mind. In K. S. Rosengren, C. N. Johnson, and P. L. Harris, *Imagining the Impossible: Magical, Scientific, and Religious Thinking in Children*, 1-34. Cambridge: Cambridge University Press.
PIAGET, JEAN
1932 *The Moral Judgment of the Child* (trans. M. Gabain). New York: Free Press.
PREMACK, DAVID, and ANN JAMES PREMACK
1994 Moral belief: Form versus content. In Lawrence A. Hirschfeld and Susan A. Gelman (eds.), *Mapping the Mind: Domain-Specificity in Cognition and Culture*, 149-168. New York: Cambridge University Press.
PUSTEJOVSKY, JAMES
1995 *The Generative Lexicon*. Cambridge, MA: MIT Press.
SCITOVSKY, TIBOR
1993 The Meaning, Nature, and Sources of Value in Economics. In Hechter et al. 1993, 93-105.
SEARLE, JOHN
1995 *The Construction of Social Reality*. New York: Free Press.
STEVENS, JEFFREY R. and MARC D. HAUSER
2004 Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences* 8.2.
STICH, STEPHEN P.
1993 Moral Philosophy and Mental Representation. In Hechter et al. 1993, 215-228.
TURIEL, ELLIOTT
1983 *The Development of Social Knowledge: Morality and Convention*. Cambridge: Cambridge University Press.

Copyright of Journal of Cognition & Culture is the property of Brill Academic Publishers and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.