ARTICLES

Information as Verb: Re-conceptualizing Information for Cognitive and Ecological Models

DAVID G. CASAGRANDE
DEPARTMENT OF ANTHROPOLOGY
UNIVERSITY OF GEORGIA

Abstract

Current notions of information are inadequate for ecological and cognitive models because they: 1) only account for information gain that results from reducing uncertainty; 2) assume binary logic; 3) fail to account for semantics and pragmatics; and 4) can not account for shared and externalized cognition. A different model of information is presented here, which treats information as a process of state change (i.e., the term is used as a verb), rather than as a variable. The potential for information is defined to include not only stimuli, but the context of the informational moment; and is distinguished from realized information, which is the result of a state change. The proposed model also distinguishes epistemological levels of abstraction at which information takes place. Abstraction, fuzzy logic, and consensus supersede the reduction of uncertainty, and pragmatic contextual marking of information at different epistemological levels provides a basis for explaining shared and externalized cognition.

Introduction

The term "information" has been used to describe a variety of organizational forms from genetic structure to culture. Attempts to include information in ecological models have, however, mostly relied on vague or enigmatic notions of feedback. In cognitive models, information is considered to be all raw data available for processing. But it is unclear how information is screened or organized. This paper addresses the need for revision of current notions of information within ecology and the cognitive sciences. It is particularly important to bridge the gaps that exist between the scales of intrapersonal communication, interpersonal communication, and cultural consensus.

Current notions of information, and in particular attempts to quantify information, can be traced to the work of Claude Shannon (Mingers 1997). In this paper, I will focus on the work of Klir and Folger (1988), which represents the most rigorous and recent attempt to model information mathematically, and because it illustrates the prevalence of Shannon's original core concepts (Shannon and Weaver 1949).

My goal is to build upon Klir and Folger's theoretical background by first elucidating its limits, and then incorporating ideas from the disciplines of cognitive studies, linguistics, information philosophy, and information ecology to develop a new model of information. I present the case that current theory is inadequate because it: 1) only accounts for information gain that results from reducing uncertainty; 2) assumes binary logic; 3) fails to model semantics and pragmatics; and 4) has not modeled shared and externalized cognition. I present a new model, which is based upon an interdisciplinary synthesis, and forms the basis for formal development. Such a model is necessary for understanding human interaction on an interpersonal and organizational level as well as human relationships with non-human components of the environment. As such, it is intended to provide a basis for clarifying existing conceptual relationships between cognitive theory and behavioral ecology.

Problem 1. Reduction of Uncertainty as the Sole Metric of Information

The prevailing notion of information, formalized by Shannon, is the view of information as negentropy, or the reduction of uncertainty. As expressed by Weaver:

> That information be measured by entropy is, after all, natural when we remember that information, in communication theory, is associated with the amount of freedom of choice we have in constructing messages. Thus for a communication source one can say, just as he would also say it of a thermodynamic ensemble, 'This situation is highly organized, it is not characterized by a large degree of randomness or of choice—that is to say, the information (or the entropy) is low.' (Shannon and Weaver 1949:13)

A highly probable event allows no freedom of choice and therefore carries little information. An example would be the last few letters of the word "sentence" in this sentence. The "ence" in "sentence" belongs to a class that Weaver calls redundant; "that is to say, this fraction of the message is unnecessary (and hence repetitive or redundant) in the sense that if it were missing the message would still be essentially complete, or at least could be completed" (Shannon and Weaver 1949:13).

This forms the basis of Klir and Folger's more recent approach (Klir and Folger 1988), in which they propose that the reduction of uncertainty by a measurable amount indicates the gain of an equal amount of information. I shall refer to such changes as 'state changes'. The amount of information obtained by an act (or state change) may be measured by the difference in uncertainty before and after the act, and encounters with improbable entities are considered to have higher information content (Klir and Folger 1988:189).

But this unidirectional approach (Figure 1) is inadequate for human communication because it does not account for the information content of a message which greatly increases uncertainty (i.e., accelerates entropy; Brainerd and Reyna 1990). While it is true that we strive for certainty in communication, the structure of our mental state – our understanding of syntax, meaning and pragmatics – can also be significantly altered by an experience that reduces certainty. This derives largely from the

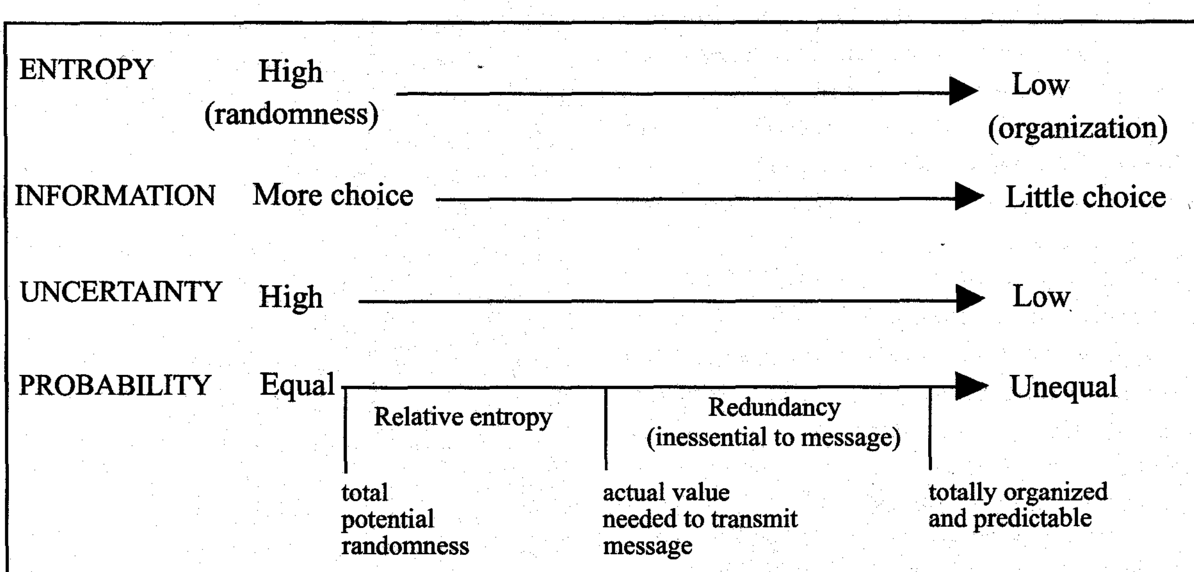

FIGURE 1: THE RELATIONSHIPS BETWEEN ENTROPY, INFORMATION, AND PROBABILITY AS PROPOSED BY SHANNON AND WEAVER (1949) AS A THEORY OF COMMUNICATION.

fact that human information is not an objective measurable entity that operates merely on a probability matrix, as described by Klir and Folger. Rather, information represents a state change (Δ) that can include greater uncertainty; or in the case of organizational communication, greater consensus regardless of uncertainty.

The unidirectional approach becomes even more problematic when describing the process of abstraction. Following the logic of Klir and Folger, in order for Δ to have a positive value as a result of abstraction, the said abstraction must increase predictability, or be perceived as such. But the resulting abstraction may or may not increase certainty. An example is the creation of gist, which is a mechanism for dealing with inundations of complex environmental stimuli (Brainerd and Reyna 1990). Certain attributes of the information set are selected in order to essentialize the message, but there is no guarantee that the selected attributes are appropriate for understanding what the sender of the message intends. In such a case, abstraction, or gist formation, would not increase certainty, although Δ will have occurred. The need to form abstractions can be considered a necessary process that increases entropy. That is, cognition can produce "information states whose structures are impoverished" (Brainerd and Reyna 1990:19).

Thus, a definition of information limited to increased certainty is inadequate for describing the short-term, near-instantaneous processes of human thought. And, as I will discuss below, such a definition also constrains attempts to describe processes of externalized cognition and interpersonal communication.

Problem 2. Limits of Binary Logic

Binary logic is another core concept from Shannon and Weaver's conceptualization of communication (1949) carried over into recent theory:

> Transformation of information into meaning involves a digitalization of the analogue. . . The importance of this for information and meaning is the argument that our perception and experiences are analogue while cognition and meaning are progressive digitalizations of this experience. . . Meaning, or the semantic content of an information source is that information *and only that information*, which is held in digital form. (Mingers 1997:81; emphasis in original)

Shannon and Weaver consider choices to be binary (either/or), and hence, information could be measured by the logarithm to the base 2 (p. 9). Binary logic is also basic to Klir and Folger's more recent theoretical revision (1988). They state that the truth value of a single proposition (i.e., that an element x belongs to a probability set) is denoted by the values 1 and 0, and the unit that characterizes full uncertainty is a bit (p. 145).

This binary approach has proven to be effective for mathematical formulations and computer science. Such an approach is not, however, always applicable to humans. We don't treat propositions as either/or decisions; instead we tend to "hedge" and introduce other variables into the decision matrix. Analyses of set inclusion indicate that we often use more than two dimensions simultaneously (D'Andrade 1995:139).

Problem 3. Failure to Model Semantics and Pragmatics

Information in human communication is commonly divided into three classes: syntactic (the relationship among signs employed in communication), semantic (the relationship between the signs and meaning), and pragmatic (the relation between entities and their utilities)(Klir and Folger 1988:188; Mingers 1997).

Schwarz' (1996) analysis of several experiments led him to conclude that pragmatics is the most important aspect of human communication. There is a "common misperception that language use has primarily to do with words and what they mean. It doesn't. It has primarily to do with people and what they mean. It is essentially about speakers' intentions;" that is, the conversational context (Schwarz 1996:7, quoting Clark and Schober 1992). In short, listeners rely more on pragmatic, rather than semantic (or syntactic), information.

Contemporary models based on Shannon's approach are, however, incapable of addressing information other than syntactic (Mingers 1997).

Klir and Folger's (1988) scheme, building upon Shannon's, applies easily to syntactic information, and has been successfully applied to computer science. But, as Klir and Folger point out, semantic and pragmatic information are also required for human communication:

> A measure of uncertainty, when adopted as a *measure of information*, does not include semantic and pragmatic aspects of information. As such, it is not adequate for dealing with information in human communication. (1988:140; emphasis in original)

Such an approach does not say anything about how or why particular interpretations are generated or selected, and "information" remains simply a metric (Mingers 1997). Reduction of uncertainty in human cognition is accomplished only when options are eliminated; this requires a pragmatic connection between the prospective outcomes of acts and the entities in which they are applied (Klir and Folger 1988:188).

Schwarz (1996) provides a pragmatic model in which a tacit assumption underlying the conduct of conversation is that "communicated information comes with a guarantee of relevance" (p. 4). Listeners assume the speaker tries to be informative, truthful, relevant, and clear. Listeners assume that speakers are to be informative and are to provide information that is new to the recipient, rather than to reiterate information that the recipient already has (p.6). If a speaker violates these assumptions, s/he runs the risk of being misunderstood, because listeners still make those assumptions. Thus, communicated information comes with at least a partial guarantee of relevance, and listeners draw on these cooperative assumptions in interpreting the speakers contributions. As a result, information that the speaker considers irrelevant (because of a focus on semantic meaning) may nevertheless be relevant for listeners, who focus on pragmatic implications (Schwarz 1996:16). This information is marked by context. Information can not be redundant per Klir and Folger; that is, it can not have zero value, because a repetition has meaning, it reiterates or reinforces (Bateson 1972:131). Repetition causes a state change, because the listener assumes the speaker intends relevance in that repetition, and/or the listener may analyze the message in terms that question the speaker's communicative abilities and social skills, or the listener may choose not to listen.

Problem 4. Limits in Conceptualizing Shared and Externalized Cognition

I will use the term "externalized cognition" to refer to those thought processes and environmental interactions that are distributed among individual organisms and their tools. Hutchins (1995) provides an example in which the cognitive processes involved in flying an airplane are distributed among the cockpit crew and their instrumentation. No single person internalizes all of the necessary flight data, no two people interpret data in the same way, nor do they attempt to do so simultaneously.

Klir and Folger's (1988) theoretical approach, as applied to interpersonal communication, is based upon an individual's matrix of *a priori* assumptions about the probability, possibility, and pragmatics of stimuli (e.g., potential acts, questions, etc.)(p. 234). The stimuli can be self-induced or of external origin. Events of lesser probability represent higher information. For example, getting a "no" answer when one expects a "yes" leads to an increase in knowledge (i.e., a revision of the probability/possibility matrix), whereas receiving a "yes" presumably changes little. But, as discussed above, events of high probability can also represent increased information (through redundancy). In Hutchins' example of the cockpit crew, such redundancy is critical for crew members to coordinate distributed cognition. Hence, state changes can occur irrespective of probability-possibility matrices.

Information in such a system of externalized cognition must be marked by the attribution of mental states. For example, a speaker may know that a listener knows something, but has only a partial theory of what the listener knows. The speaker's theory is structured by pragmatic context, and must attribute an appropriate mental state to the would be listener. Hence, the level of intensionality is specified by the pragmatic context. Information marking within contexts is fundamental to the structure of consensus, because all consensus is context-specific. Thus, Klir and Folger's model can be expanded to include changing contexts external to individual probability/possibility matrices.

The Need for a New Approach

Despite the four problems outlined here, Klir and Folger (1988) provide a rich basis for developing a more comprehensive model of information in human communication and human-ecological processes. Surprisingly, no comprehensive attempt to develop such an approach has occurred within ecology or the cognitive sciences, in spite of the widespread use of Shannon's index as a metric of ecosystemic complexity and diversity (Ludwig and Reynolds 1988).

There is little agreement about the nature of information, particularly with regard to its semantic and pragmatic aspects (Mingers 1997). A fundamental problem is that information is neither a physical entity, nor is it a quantifiable variable. This inevitably leads to the use of the word "information" to signify all data available to a processing system, and has led Gilligan to propose the paradox: "What is information?"

> How are we to say that there is a difference between data and information if there is no way to perceive that difference in the world? . . . If there is no distinction between the existence of information and believing in the existence of information, then we cannot talk about information at all. (Gilligan 1997:68)

Most attempts to define information have been attempts to arrive at some explanation for a state change, and it is possible to escape Gilligan's paradox by reverting to the original use of the word "information" as a verb:

> It will suffice (and therefore it should be the requirement) to use the word information in its original sense, that is, as a verb. There is no information: we are 'informed'. (Gilligan 1997:68)

Thus, information can be defined as a state change (Δ); that is, from a state of no difference, to a state of difference. By this definition, information is neither a constant nor a variable, and it is certainly not a physical entity. Potential information (I_p) represents the potential for a state change, not objective physical phenomena, such as photons of light or sound waves. And potential information is not limited to the existence of some symbol or sound, but includes also the context of the moment. For example, a pedestrian deep in thought might not even notice cars passing by. The cars *alone* do not represent potential to be informed (I_p). The context in this example also includes behavior and awareness (cognitive predisposition). In this case we could expect Δ to equal 0; that is, no state change. If, however, our day-dreaming pedestrian strays off the sidewalk onto the street and someone yells "look-out," this different I_p could yield a Δ much greater then zero; that is, a profound state change. In this highly simplified example, the words "look out" and their meaning can be thought of as I_p; the cars can not.

In sum, information can not occur unless the organism or consensual group is predisposed, by pragmatic context, to a change of state (Δ), which also requires the assimilation and internalization of data. By this definition, we can distinguish potential information from all the data that constitute objective reality.

Because information represents any conversion from no difference (Δ_o) to a state of difference (Δ_1), it can be entropic or neg-entropic; towards certainty or uncertainty. Our pedestrian would be plunged into the realm of uncertainty upon hearing the words "look out." Realized information (I_r) can be considered a measure of the difference between the original state and the subsequent state—the original context and the resultant context (i.e., a measure of Δ_1).

Other examples of I_r include abstractions and gists (Brainerd and Reyna 1990), which result from the need to reduce complexity and coordinate distributed cognition. Information is realized (i.e., a change of state) by screening data, and becomes marked potential information when externalized by individuals for the benefit of others within pragmatic context. The externalization of data, with the intention of creating the potential for information and/or inducing a state change can be defined as 'exformation.'

The information process can be formalized by the equation:

$$I_p \longrightarrow I_r = \Delta_I \qquad \text{(eq. 1)}$$

By this definition, knowledge is not synonymous with information. Information is the change, whereas knowledge is one result of the change (I_r). Knowledge can be a state of abstraction that is achieved as the result of having reduced uncertainty through information (Δ_I). But note that knowledge (I_r) can only be produced if there exists the potential for information to occur (I_p).

Also, information here does not represent genetic or ecosystemic organization. The latter is biological structure, and is the result of information (Δ_I).

A Model of Information

The information process can be represented diagrammatically (Figure 2). The concept of epistemological levels utilized here (L^0 through L^5) derives from Klir and Folger (1988: 194) and includes a synthesis of concepts from different disciplines (Table 1). At lower epistemological levels we deal only with raw data. At higher levels we increase the level of analysis of relations among data; sets become fuzzier and thinking is increasingly more abstract.

Realized information at any epistemological level ($I_{r(L^i)}$) is a function of state changes at lower levels, and is potential information at higher levels ($I_{p(L^{i+1})}$) if, and only if, there is a state change $\Delta_{I(L^{i-1})}$ at the next lower level, that is greater than 0. Stated formally:

$$I_{r(L^i)} = \Delta_{I(L^{0...i})} \qquad \text{(eq. 2)}$$

and

$$I_{r(L^i)} = I_{p(L^{i+1})} \text{ if and only if } \Delta_{(L^i)} > 0 \qquad \text{(eq. 3)}$$

Table 1: Relations Between Disciplinary Views of Epistemological Levels.

Level	Information theory (Klir & Folger 1988)	Cognitive & psycholinguistic studies (D'Andrade 1995)	Information ecology
L^5	meta metasystem	shared cognition	shared & externalized cognition
L^4	metasystem	cultural models refined, checked, & abstracted; meta models	collective of deductions, integration with externalized cognition
L^3	structure system	schema & cultural models constructed, pragmatic (intention)	imagination, deduction, prediction
L^2	generative system (refined data)	domain construction, gist (syntactic)	relational induction
L^1	raw data system	meaning, short-term memory (semantic)	states of variables
L^o	source system		screen/filter; variables defined

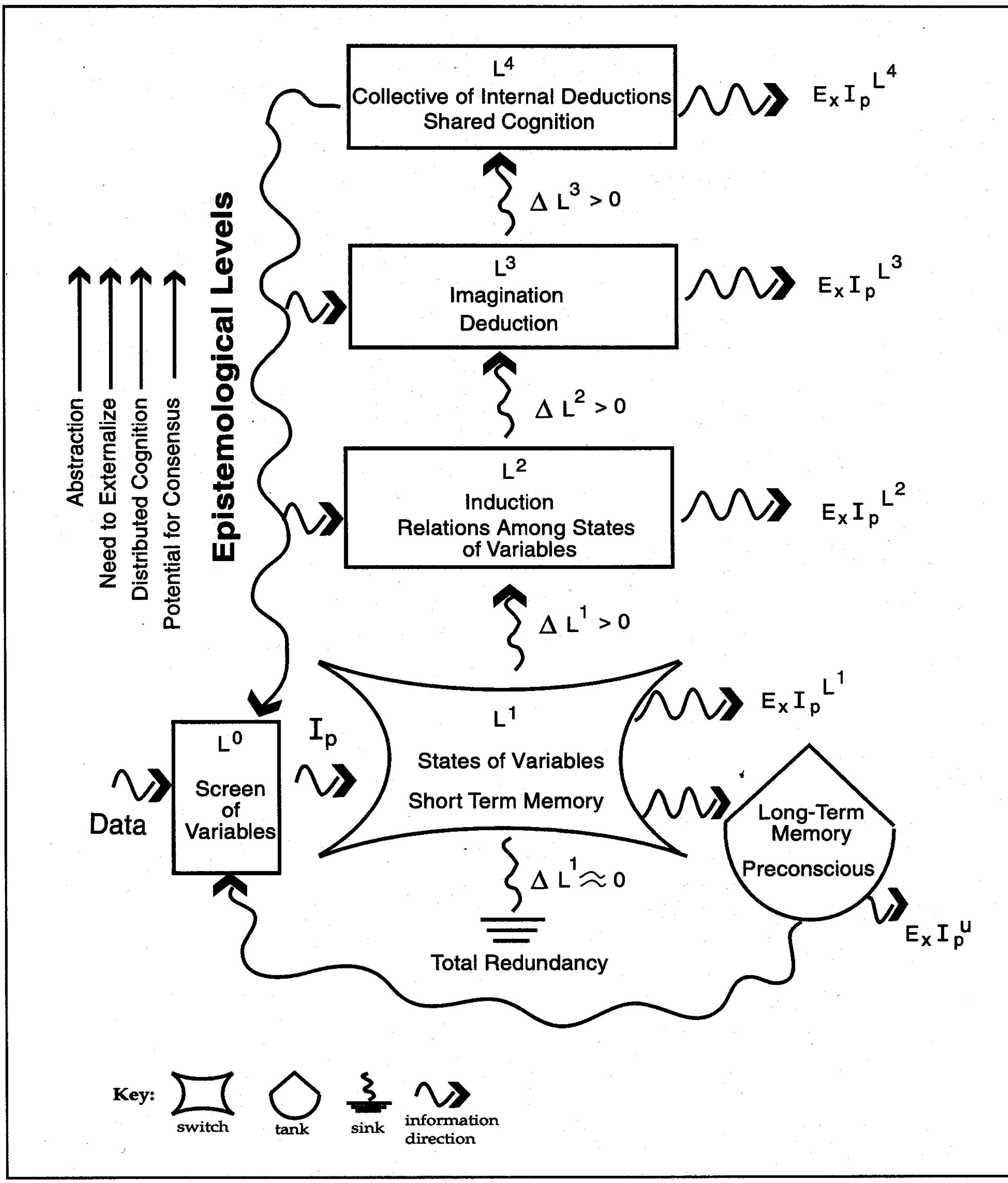

Figure 2: Partial Model of Information Process and its Distribution over Epistemological Levels.

Informational upgrading occurs when I_p that results from a state change at one level results in a state change at a higher level (Figure 2). The potential for this to occur depends on both a sender's and receiver's understanding of how a pattern is marked. Informational downgrading occurs when a state change at a higher level feeds back and causes a state change at a lower level.

The text within each box in Figure 2 includes the terminology that I have chosen for describing each epistemological level in this synthesis. L^0 represents a screen of variables. That is, data will only be accepted for certain variables, such as sounds and gestures in conversation. The variable set that exists at any given moment depends on context and feedback from higher epistemological levels. Thus, intentionality can be partly defined as the ability to choose variables at any level (L^i).

L^1 is a switch (Figure 2) within which data (i.e., states of variables or patterns) are evaluated. Messages that yield no state change ($\Delta^1 \approx 0$) are ignored, they represent absolute redundancy. But as discussed below, context rarely allows for this condition because all messages are marked by cognitive predisposition. Messages that produce state changes that are small, that is approaching zero (Δ^1 -> 0), can be relegated to long-term memory, or the preconscious (Bateson 1972). An example is rote repetition of bodily movements, such as when playing a musical instrument. Other messages, which produce state changes at L^1, are potential information (I_p) for higher epistemological levels.

In the earlier example of the daydreaming pedestrian, the words "look out" would force a downgrading from a higher, more abstract level, to a complete revision of the variables being screened at L^o. In particular, the pedestrian would begin screening for visual and audio variables that might indicate danger. A state of a variable in L^1 might include a car approaching quickly.

So far, this discussion has focused on the intrapersonal relationships of the model in Figure 2. As for interpersonal communication, we can consider the ability to induce a change of state in another person or group to result from potential exformation (E_xI_p; i.e., the intentional creation of I_p). Because exformation, as a human product, represents a potential to induce a state change at a specific epistemological level, it can be said to be marked ($E_xI_p^i$). As noted by Schwarz:

> In general, determining the intended meaning of an utterance requires extensive inferences on the part of listeners. Similarly, designing an utterance to be understood by a given listener requires extensive inferences on the side of the speaker. In making these inferences, speakers and listeners rely on a set of tacit assumptions that govern the conduct of conversation in everyday life. (Schwarz 1996:7)

Messages can be marked by the context for relevance, obscurity, ambiguity, quantity and quality (Schwarz 1996:9). Such marking induces a revision of variables at L^o.

> The concept of information applies not to the individual messages (as the concept of meaning would), but rather to the situation as a whole, the unit information indicating that in this situation one has an amount of freedom of choice, in selecting a message, which it is convenient to regard as a standard or unit amount. (Shannon and Weaver 1949: 9)

To reiterate an earlier theme of this paper, language use has primarily to do with people's intentions (Schwarz 1996). It is through such pragmatically contextualized intention that messages are marked. The notion of marked messages provides an informational foundation for consensus theory (Romney et al. 1986). As defined in this paper, the potential to inform (I_p) can include patterns perceivable by the senses, as well as neurophysiologically coded patterns that have resulted from previous informational events (i.e., ontogenic development and experience), or which are genetically determined.

An example of neurophysiologically coded I_p is that some plants look more like each other than other plants, and people from all cultures tend to distinguish the same pattern of groups – what western systematists call the genus (Rosch 1978; Berlin 1992). Such natural discontinuities in nature provide an example of patterns with the potential to inform, but only at certain epistemological levels. The ability to draw inferences from those patterns (induction), or make predictions (deduction), depends on the degree to which the patterns are marked by experience.

Marking can occur by one person rearranging patterns of meaning (e.g., color and form in the case of modern art) so that they are only recognizable at a specific epistemological level. Marking also relies on the ability of the receiver to recognize meaning in patterns. Greater abstraction, when communicated, leads to a greater state change in the recipient because state changes are precipitated at all lower epistemological levels.

Potential exformation ($E_x I_p$), therefore, represents the potential for a state change of groups of individuals to greater consensus. But total consensus is elusive, so we must generalize, essentialize or gist in order to enhance predictability about the behavior of others (Hallpike 1986). For example, writing is essentialization of spoken language, whereas money is the essentialization that allows for complex trade and exchange. Realized information (I_r) could represent the degree to which money increases consensus.

Robb (1997) alludes to the basic psychological need for predictability:

> The fact that . . . we continue to employ so-called causal, or predictive, models in singular cases suggests to me that they satisfy some deep-seated psychological needs which are surfaced under the stress of decision taking. As individuals we try to think that we are doing the 'right' thing and so we pass some of the responsibility for many decisions on to an 'expert' or 'guru' or to a body of accepted specialist professional knowledge. How often we must benefit from the 'placebo effect' of employing *argumentum ad verecundiam.* (Robb 1997:15)

This describes how we substitute predictability of human behavior for unpredictability of natural phenomena through consensus. A process which can only occur by pragmatically marking information at the appropriate epistemological levels.

Conclusion

If information is conceptualized as a process of changing states (i.e., as a verb, and not a variable) we are able to distinguish between the potential to be informed, or to inform others, and the ambient raw data that consitutes all objective reality. By also recognizing that the potential for information to occur is marked by pragmatic context, and at various levels of abstraction, or epistemological levels, we can begin to bridge the current theoretical gap between intrapersonal and interpersonal cognition. The concept of pragmatic contextual marking allows us to better understand the trade-offs between dealing with inundations of stimuli via abstraction or gist, the human need to increase predictability, and the need to establish group consensus.

References Cited

Bateson, G.
1972 *Steps to an ecology of mind.* San Francisco: Chandler Publishing Co.

Berlin, B.
1992 *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies.* Princeton, NJ: Princeton University Press.

Brainerd, C. J., and V. F. Reyna.
1990 Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review* 10:3-47.

D'Andrade, R.
1995 *The development of cognitive anthropology.* Cambridge: Cambridge University Press.

GILLIGAN, J.
1997 "Patterns on glass: The language games of information," in *Philosophical aspects of information systems.* Edited by R. L. Winder, S. K. Probert and I. A. Beeson, pp. 65-72. Bristol, PA: Taylor & Francis Inc.

HALLPIKE, C. R.
1986 *The principles of social evolution.* New York: Oxford University Press.

HUTCHINS, E.
1995 How a cockpit remembers its speeds. *Cognitive Science* 19:265-288.

KLIR, G. J., AND T. A. FOLGER.
1988 *Fuzzy sets, uncertainty, and information.* Englewood Cliffs, NJ: Prentice Hall.

LUDWIG, J. A., AND J. F. REYNOLDS.
1988 *Statistical ecology.* New York: John Wiley and Sons.

MINGERS, J.
1997 "The nature of information and its relationship to meaning," in *Philosophical aspects of information systems.* Edited by R. L. Winder, S. K. Probert and I. A. Beeson, pp. 73-84. Bristol, PA: Taylor & Francis Inc.

ROBB, F. F.
1997 "Some philosophical and logical aspects of information systems," in *Philosophical aspects of information systems.* Edited by R. L. Winder, S. K. Probert and I. A. Beeson, pp. 7-22. Bristol, PA: Taylor & Francis Inc.

ROMNEY, A. K., S. C. WELLER, AND W. H. BATCHELDER.
1986 Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist* 88:313-338.

ROSCH, E.
1978 "Principles of categorization," in *Cognition and categorization.* Edited by E. Rosch and B. Lloyd, pp. 28-49. Hillsdale, NJ: Lawrence Erlbaum Associates.

SCHWARZ, N.
1996 *Cognition and communication: Judgmental biases, research methods, and the logic of conversation.* Mahwah, N.J: Lawrence Erlbaum Associates.

SHANNON, C. E., AND W. WEAVER.
1949 *The mathematical theory of communication.* Urbana, IL: University of Illinois Press.