Comment on diSessa

William J. Clancey

A recent series of papers in *Cognition and Instruction* (Volume **10**(2-3)) has considered diSessa's p-prim epistemology of physics and its relation to situated cognition. Commentaries by Marton, Ueno, and Chi & Slotta were printed, along with diSessa's response. In this note, I seek to clarify Ueno's and diSessa's references to my paper, "A boy scout, Toto, and a bird" (in press). This position paper was written in 1991 for an AI audience, to provide advice for building robots. The argument is that recent work in artificial life makes many of the same assumptions about the nature of human knowledge and memory that underlie the symbolic modeling approach it seeks to overturn.

I find diSessa's response to be generally a model of diplomacy and scholarship. I admire his ability to acknowledge the disparate views and place his work in relation to them. The problem is that, although Ueno generally understands my points, he has omitted key parts of a cited passage. Unfortunately, diSessa then picks up the abridged quote and continues with an absurd misreading: "Clancey, according to the paper quoted by Ueno, wants to deny any internal representations at all."

Here is the full section in which the quote appears. Ueno begins with "Such representations..." omitting the preceding sentence (and qualifying phrase about storage) that indicates I was referring to representations in expert systems and symbolic cognitive models. My claim is that human knowledge is not a collection of structures in a language, such as we find in a knowledge base or cognitive model.

The section quoted by Ueno appears here in *larger italic font*. Citations have been updated to reflect publication dates.

_____

> ❏ Knowledge is an explanatory concept like energy, a capacity for interactive behavior. Knowledge can be represented, but "knowledge is never in hand" (Newell, 1984). "The map is not the territory" (Korzybski, 1941).
>
> **Knowledge-level descriptions** (e.g., prototype hierarchies, scripts, strategies) constitute an observer's model, characterizing patterns in behavior—the product of internal mechanisms and some environment—not structures or mechanisms inside the agent. Just as there is no such thing as "all in the information" in some situation, there is no such thing as completely describing an agent's behavior. Descriptions are always with respect to the frame of reference of an observer interacting with an agent-in-its-environment; descriptions embody the observer's point of view (including values and goals) and are themselves the product of interactions.

Cognitive models (including expert systems) replicate the patterns of human behavior—how it appears in recurrent interactions—without replicating the mechanism that produces human behavior. Such descriptions are necessary and valuable; they help specify what a cognitive architecture must be capable of accomplishing. In human behavior, such models, in the form of natural language grammars, disease hierarchies, operating procedures, etc. are extremely valuable for coordinating group behavior, or in general designing, controlling, diagnosing, and repairing complex systems (Clancey, 1992).

❏ Meaningful structures are not fixed, given, or static in either the environment or in human memory:

Human memory is not a place where things (e.g., schemas, categories, rules, procedures, scripts) are stored. *Such representations*—when they are not stored in the environment—*are always constructed each time they are used. Representations are not manipulated by people in a hidden way, but must be perceived to be interpreted; that is, they must be in the environment (including silent speech and imagery). Interpretation is a process of commentary, constructing secondary representations that give meaning to experiences and perceptions by placing them in a context, thus relating them to activity* (Suchman, 1987; Agre, 1988; Clancey, 1991).

Information is not given to the agent. Information is constructed by people in a process of perception; it is not selected, noticed, detected, chosen, or filtered from a set of given, static, pre-existing things (Maturana, 1983; Reeke and Edelman, 1988). Each perception is a generalization, a new construction. No category is merely retrieved or reinstantiated. In people every utterance is a new representation.

_____

"Such representations" in this passage refers primarily to formal symbol structures expressed in some language, exemplified by symbolic cognitive models such as ACT*, MOPS, Neomycin, and Repair Theory. I say that these representations of knowledge can be stored in the environment, and must otherwise be constructed, meaning that we must physically create them by saying or writing something. I also say that such representations of knowledge ("knowledge-level descriptions") model patterns of interactive *behavior*, not structures inside the agent. By these omissions, Ueno and

diSessa obscure the essential distinction between human knowledge and *computer representations of knowledge* that I make in this paper.

Indeed, the abridgment of my remarks and excision of a phrase nicely illustrate how "meaningful structures" are not merely *given*, but are partly a construction of the observer's comprehension process and activity. In this case, in excerpting and commenting on my paper, Ueno has indicated what text is meaningful to him.

On this cited page of my position paper, I briefly articulate the relation of symbolic cognitive models, human knowledge, internal mechanism, information, etc. Nevertheless, in his response diSessa says, "Clancey, according to the paper quoted by Ueno, wants to deny any internal representations at all." This interpretation of what I say is incorrect. Instead, I deny that there are internal *representations of knowledge*: Human memory is better characterized as a *capacity* than as a *repository*. A knowledge engineer may represent what someone knows in terms of formal linguistic descriptions, such as the rules in an expert system, but these rules are not literally stored in the expert's head. In fact, knowledge bases often contain models of the world that go beyond what anyone has said before (Clancey, 1989).

We can all agree with diSessa that "radical situated claims against internal representations throw the baby out with the bath water." But I would like to add, "Radical *interpretations* of situated claims about internal representations throw the baby out with the bath water."

With the broader context of my analysis in front of us, we can shed some light on diSessa's subsequent remarks: "Clancey's image is perilously close to the disturbing model both Marton and I rejected: Meaning is generated in response to meaningless external configurations in the form of other (external) forms. Instead, I believe we must come to see how meaning resides implicitly in experience." I fully agree with the last sentence, but the preceding remark is not what I intended.

In the cited passage, I was referring to the process of interpreting the meaning of an expert system rule or interpreting the meaning of something we have just said to ourselves. The process of interpreting representations involves perception, either directly in reading a text or diagram, or indirectly as we comprehend and comment on our own imagined dialogue and visualizations. That is, representing is a conscious activity, occurring in the sequence of our speech and writing, not something that happens *between* acts:

> Representations are not manipulated by people in a hidden way, but must be perceived to be interpreted; that is, they must be in the environment (including silent speech and imagery). Interpretation is a process of commentary, constructing secondary representations that give meaning to experiences and perceptions by placing them in a context, thus relating them to activity (Suchman, 1987; Agre, 1988; Clancey, 1991).

Interpreting what a representation means (e.g., saying what a symbol in an expert system rule means) involves generating other representations, which I call "commentary." This *representation of meaning* must not be equated with meaning itself. For example, a conceptual dependency representation of the meaning of a sentence must not be equated with a person's understanding of the sentence. The commentary may be about symbolic structures in the external environment or about internal experiences (when I carry on an imaginary conversation in my head). The commentary itself may be expressed in the form of symbolic structures in the environment or be an internal experience. And of course, we can comprehend text without saying what it means. With all these distinctions—inner vs. external speech, experience vs. description, neural structures vs. representational artifacts, understanding (as an activity) vs. modelling, etc.—it is no wonder we stumble in our in-print conversations.

In this passage, I wanted to emphasize the distinction between internal and external *structures*. External structures—exemplified by statements in a journal article or rules in a knowledge base—are physical symbols, strung, stored, and arranged spatially in some medium. Internal neural structures, whatever they may be, are not of the same logical type; they are not linguistic structures that are put away, indexed, and formally matched. Perhaps the passage could be improved by not using the term "representations" to refer to both external structures (e.g., rules in a knowledge base) and *experiences of representing* to ourselves. When I say to myself "The sky appears deep blue today" there is no representational artifact like a statement written on paper. I am representing, but there are no internal representations in the sense that there are rules or schemas in an expert system or symbolic cognitive model.

The paper excerpted by Ueno, as well as many of my other recent publications (Clancey, 1993a, Roschelle and Clancey, 1992), begins by claiming that the term "representation" has been used too loosely in the AI literature to refer both to structures or processes in the head and to external structures, as in an expert system, that people deliberately create, arrange, and interpret. Similarly, when I say perception is involved in the commentary process, this doesn't mean that perceiving during an imaginary conversation should be equated with perceiving sounds in the external environment. The

more general idea here is that the environment can be internal or external, as Dewey nicely expressed:

> An experience is always what it is because of a transaction taking place between an individual and what, at the time, constitutes his environment, whether the latter consists of person with whom one he is talking about some topic or event, the subject talked about being also a part of the siutation; or the toys with which he is playing; the book he is reading; or the materials of an experiment he is performing. The environment, in other words, is whatever conditions interact with personal needs, desires, purposes, and capacities to create the experience which is had. Even when a person builds a castle in the air he is interacting with the objects which he constructs in fancy. (Dewey, 1938, p. 519)

In the transactional perspective, creating and interpreting representations (e.g., knowledge bases, drawings, and journal articles) is a conscious process involving a sequence of sensorimotor coordinations, going on in our activity over time (Dewey, 1896). Understanding how perceptual, conceptual, and motor functions can arise together, as Dewey conjectured, is partly informed by recent neuroscience research (Edelman, 1992; Clancey, 1993b). On the matter of drawing upon brain science, diSessa says, "But knowledge is hopelessly complex and particular for us to relegate to brain scientists at this point (and, I conjecture, at any future point)." But multidisciplinary scholarship must be distinguished from *relegating* work to someone else.

diSessa goes on to say, "As for Clancey's outward push [to the physical and social world], intuitive physics makes it perfectly clear that what gets 'internalized' to return in action and explanation is not a mirror of the physical world." I strongly agree. Perhaps the first step in understanding situated cognition is to break away from the correspondence view of reality, in which knowledge is equated with descriptions (models of the world), and meaning is viewed as another descriptive process of mapping these descriptions to a somehow objectively known world (Tyler, 1978).

I would like to conclude by heartily endorsing diSessa's summary of the individual and social points of view about cognition in the opening pages of his response to Ueno. Perhaps Ueno's selections from my text made it difficult for diSessa to see that our research programs and philosophical perspectives are almost fully aligned.

References

Agre, P. 1988. The dynamic structure of everyday life. Dissertation in Electrical Engineering and Computer Science, MIT.

Clancey, W.J. 1989. The knowledge level reconsidered: Modeling how systems interact. Machine Learning, **4**(3/4) 285-292.

Clancey, W.J. 1992. Model construction operators. *Artificial Intelligence.* **53**(1) 1-124.

Clancey, W.J. 1993a. Situated action: A neuropsychological interpretation (Response to Vera and Simon). *Cognitive Science*. **17**(1) 8-116.

Clancey, W.J. 1993b. The biology of consciousness: Review of Rosenfield's "The Strange, Familiar, and Forgotten" and Edelman's "Bright Air, Brilliant Fire." *Artificial Intelligence* **60**:313-356, 1993.

Clancey, W.J. (in press). A boy scout, Toto, and a bird. In L. Steels and R. Brooks (eds) *The "Artificial Life" Route to "Artificial Intelligence": Building Situated Embodied Agents.* New Haven: Lawrence Erlbaum Associates.

Dewey, J. [1896] 1981. The reflex arc concept in psychology. *Psychological Review*, III:357-70, July. Reprinted in J.J. McDermott (ed), *The Philosophy of John Dewey,* Chicago: University of Chicago Press, pp. 136-148.

Dewey, J. [1938] 1981. The criteria of experience. In *Experience and Education*, New York: Macmillan Company, pp. 23-52. Reprinted in J.J. McDermott (ed), *The Philosophy of John Dewey,* Chicago: University of Chicago Press, pp. 511-523.

Edelman, G.M. 1992. *Bright Air, Brilliant Fire: On the Matter of the Mind.* New York: Basic Books.

Korzybski, A. 1941. Science and Sanity. New York: Science Press.

Maturana, H. R. 1983. What is it to see? *¿Qué es ver?* 16:255-269. Printed in Chile.

Newell, A. 1984. The knowledge level. *Artificial Intelligence* **18**(1) 87-127 .

Reeke, G.N. and Edelman, G.M. 1988. Real brains and artificial intelligence. Daedalus, **117** (1) Winter, "Artificial Intelligence" issue

Roschelle, J. and Clancey, W. J. 1992. Learning as social. and neural. *The Educational Psychologist*, **27**(4): 435-453.

Suchman, L.A. 1987. *Plans and Situated Actions: The Problem of Human-Machine Communication.* Cambridge: Cambridge Press.

Tyler, S. 1978. *The Said and the Unsaid: Mind, Meaning, and Culture.* New York: Academic Press.