Evidence for Explanatory Patterns in Evolutionary Biology

Jorge F. Larreamendy-Joerns
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, PA 15260
larream@pitt.edu

Stellan Ohlsson
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, PA 15260
stellan@vms.cis.pitt.edu

Abstract

Students' naive conceptions of natural phenomena have been analogized to scientific theories. This *theory view* does not account for the instability and inconsistency of students' explanations. An alternative is the *schema view* according to which students construct explanations by instantiating explanatory patterns acquired in previous learning. In two previous studies eight explanatory schemas for evolutionary change were identified through content analysis of students' explanations. It was hypothesized that if the schemas have cognitive reality, data from explanation tasks ought be consistent with data from recognition tasks. In this study students were asked to sort 24 explanations that exemplified the eight different schemas in three different ways. A hierarchical cluster analysis shows that half of the students recognized five schemas and merged the remaining ones into one broader explanatory pattern, giving partial support for the schema view. Implications of the schema view for the learning of scientific ideas are discussed.

Students' Explanatory Schemas

Students' beliefs about complex phenomena in domains such as physics, astronomy and biology have been analogized to scientific theories. Albeit some novice representations do resemble theories in the history of science (McCloskey & Kargon, 1988), the adequacy of the scientific theory/naive belief analogy remains a subject of debate in developmental and cognitive psychology (Kuhn, 1989). A problem for the theory view, which underlies most studies of misconceptions, is that students' explanations, unlike their scientific counterparts, are not stable over time and lack internal consistency (diSessa, 1988).

An alternative to the theory view is the schema view (Brewer & Nakamura, 1984; Leake, 1992; Ohlsson, 1993; Schank, 1986; Schank, Kass & Riesbeck, 1994). The latter implies that students' prior knowledge should be conceptualized as a repertoire of explanatory patterns or schemas constructed in the course of everyday learning and prior educational experience. Explanations are generated by instantiating a schema vis-a-vis a target event and replacing slots with referential expressions. Unlike propositional beliefs, which carry truth values, a schema that does not fit a given situation is not rejected, simply not activated. An existing schema can also be modified ("tweaked", Schank, 1986) to fit a novel situation, and indexed in memory as a new knowledge structure. However, the schema initially retrieved is not deleted from memory.

Explanatory schemas emphasize the instrumental role of knowledge structures rather than their veridical status. Furthermore, a schema repertoire is not a belief system, naive or otherwise. Two or more alternative schemas can coexist in memory without contradicting each other in the logical sense. Hence, no coherence demands are imposed upon the mental representation, and the question about the consistency of students' mental models is consequently de-emphasized.

In previous studies we asked students to explain instances of evolutionary change (Ohlsson, 1991; Ohlsson & Bee, 1992, 1993). Although the students produced explanations that include ideas similar to some theories in the history of biology, the instability and inconsistency, across and within tasks, of their accounts suggest that the content of their explanations are not expressions of stable beliefs about evolutionary change. Eight explanatory patterns were identified in students' explanations through content analysis: intentional creation, mutation, crossbreeding, static selection, dissemination, cognition, training, and needs and advantages.

The *intentional creation schema* rejects the very occurrence of evolutionary change and therefore bypasses any explanation of its underlying mechanisms. This schema relies on the purposive action of some agent who created the different species. The *static selection schema* explains adaptations via natural selection, but without gradualism; the selection happens within a single generation (hence its static nature). This schema does not account for the source of intra-species variation, nor does it specify the relation between selective processes and changes in reproduction rates. According to the *crossbreeding schema*, new species arise from a crossing between organisms with different traits. This schema does not rule out the possibility of inter-species breeding, nor does it constrain the kind of changes likely to occur through intra-species crossbreeding. The *mutation schema* postulates a sudden genetic change that brings to existence complex anatomical features, while the *dissemination* schema postulates an unexplained appearance of an adaptation in one or more members of the species, followed by the spread of that adaptation throughout the entire population. The *needs and advantages schema* explains adaptations as the direct result of an organism's needs to meet the demands of the environment, without explicating the mechanisms underlying the emergence of new traits. According to the *cognition schema*, adaptations are the outcome of cognitive processes such as decision, discovery and learning. Finally, the *training schema* embodies the idea that an adaptation appears as a consequence of an animal's repetitive behavior.

As this repertoire demonstrates, students do not have one major misconception (e.g., Lamarkianism) about evolution but a variety of schemas for explaining biological change. Variability exists within as well as between students. This repertoire of schemas is not a classification of students because individual students typically use more than one schema. Within the students' conception of evolution, these schemas satisfy different explanatory goals; they account for the source of change (e.g., genetic mutations, creation), the mechanism of change (e.g., selection, breeding, purposive actions, needs), or the type of change (e.g., gradual, sudden).

One of the central assumptions of schema theory is that schemas are used both generatively and interpretatively, i.e., both to produce discourse and to comprehend discourse (Brewer & Nakamura, 1984). If explanatory schemas do have psychological reality, and subjects have learned them well, then data from production tasks should be consistent with data from recognition tasks. This is the rationale behind the use of sorting tasks to assess cognitive representations (Benysh, Koubek, & Calvez, 1993; Chi, Feltovich, & Glaser, 1981).

The purpose of this study was to gather evidence to support the psychological reality of the eight explanatory patterns mentioned above, as a step toward understanding how explanations of complex phenomena are generated. We asked students to sort evolutionary explanations into categories on the basis of similarities. Each of the eight explanatory patterns mentioned above was instantiated for three different species. Thus, the 24 explanations were classifiable in two ways: First, according to explanatory patterns, in which case we would expect students to sort the explanations into eight categories with three explanations in each. Second, according to species, in which case we would expect students to sort the explanations into three categories with eight explanations in each. If students did not sort the explanations on the basis of either patterns or species, then the pattern of their groupings would deviate from either of these two ideal patterns.

Method

Subjects

Forty-nine first-year undergraduate students were recruited for this study. Students' prior knowledge of evolutionary biology was assessed by asking them to summarize Darwin's theory. Data from the summary task are reported elsewhere (Larreamendy-Joerns & Ohlsson, 1994). Students showed little or no understanding of the Darwinian theory of evolution.

Materials and Procedure

Students were instructed to sort 24 evolutionary explanations into piles on the basis of similarities and to write down the reasons for their grouping. The 24 explanations instantiated schemas that were extracted in the two prior studies (Ohlsson, 1991; Ohlsson & Bee, 1992).

In the present study, the eight explanatory schemas (i.e., needs & advantages, mutation, crossbreeding, dissemination, training, cognition, static selection, and intentional creation) were instantiated three times: one for dinosaurs' size, one for tigers' stripes, and one for birds' wings. The explanations were edited to control for word familiarity and length. Each explanation was presented on an individual card. Table 1 shows three examples.

Results

The data were analyzed with the purpose of deciding whether students sorted the explanations on the basis of explanatory patterns, species or something else. Category size ranged from 2 to 12 explanation cards (mean = 6, SD = 2.09). The frequency distribution of number of explanation cards per grouping suggests that neither species alone, nor explanatory patterns alone were the criteria guiding most students' sorts.

Table 1: Dinosaur explanations based on the schemas of needs & advantages, mutation, and static selection

Schema	Explanation
Needs & advantages	During the Jurassic period, ferns and trees were extremely large and tall, and the ground was frequently flooded. There were few plants on ground level. Dinosaurs had to reach food high up or in difficult-to-reach places, so they became larger to adapt and survive.
Mutation	Millions of years ago there were only small dinosaurs. These dinosaurs were dying because they couldn't protect themselves against large predators. Then, a mutation occurred and large size dinosaurs evolved. This mutation was highly favorable for survival.
Static Selection	Each species has different characteristics. Accordingly, some dinosaurs were bigger than others, some were smaller, some had long necks, and others had short necks. Eventually, the shorter ones perished because they couldn't adapt to the environment, leaving only the larger dinosaurs able to survive.

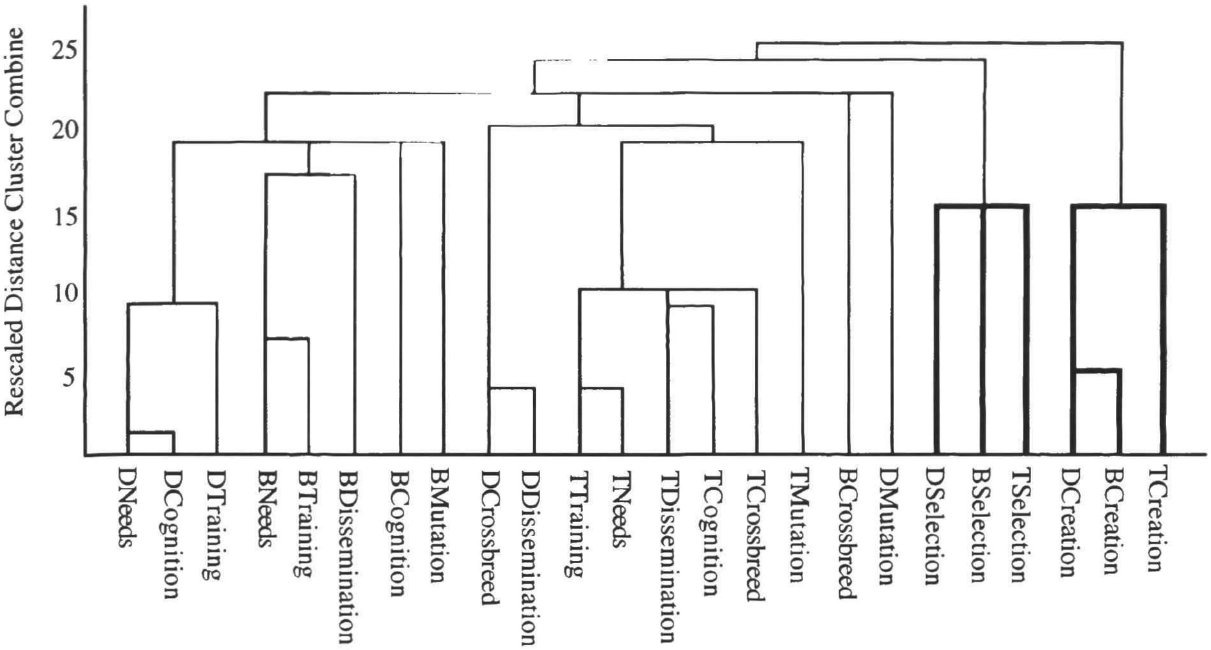

Figure 1: Dendrogram for all subjects (N = 49)

From a total of 292 categories formed by all 49 subjects 49 groupings (16.8%) were exactly in accordance with one of the eight explanatory patterns. The remaining 243 groupings were formed either by explanations pertaining to a single species (3.4%), or by a combination of explanation cards from different explanatory patterns and species (79.8%).

To get a more fine-grained picture of the sorts, the data were transformed into individual co-occurrence matrices and analyzed using the SSPS 4.1 Hierarchical Cluster Analysis Program's single linkage method (nearest neighbor). A whole-sample matrix was constructed by adding all 49 individual matrices. Figure 1 shows the dendrogram for this matrix (explanation cards are identified by a capital letter referring to the species, and the name of the explanatory schema. Letter D stands for dinosaurs; letter T for tigers; and, letter B for birds).

The dendrogram shows the explanation cards collapsed into two major higher-level nodes: the creation schema (at the right hand edge in the figure), and a large, complex node comprising the remaining 21 explanations. Albeit not all subjects recognized it, the creation schema was the most salient type of explanations for the total sample. The node comprising the 21 remaining explanations branches out into two major nodes. The first one is the static selection schema which was almost as salient as the creation schema (the second group of three from the right). The second node comprises the remaining 18 explanations which tended to be clustered on the basis of species similarities. However, within each species there was a high co-occurrence among training, cognition, and needs & advantages explanations. Mutation explanations were not seen as similar either to other types of explanations, nor to each other.

To investigate whether there were subgroups of students within the total sample who sorted according to different criteria, the whole sample was split into three groups in the following way. Each subject was assigned to one of the following groups: (a) those whose sorting actions (i.e., assignment of an explanation card to a category) were mostly based on species similarities; (b) those whose sorting actions were mostly based on similarities between explanatory patterns; and (c) those whose sorting actions were not predominantly based on species similarities, or schema similarities. To assign subjects to groups, each co-occurrence frequency in the individual matrices was classified into one of the following categories: (a) co-occurrence between explanations of the same species; (b) co-occurrence between explanations of the same pattern; and (c) co-occurrence between explanations of different schema and different species. The proportion per subject for each of these categories was computed and each subject assigned to one of three groups based on his/her greatest proportion. For each group a co-occurrence matrix was constructed and a cluster analysis performed. Since the cluster analyses of schema-oriented subjects and combined-criteria subjects yielded similar results, those two groups were merged and the analysis repeated for the merged group.

Figure 2 shows the dendrogram for the cluster analysis performed on the co-occurrence matrix for the twenty-six subjects whose sorting actions were mainly schema-oriented. Like the whole-sample dendrogram, this dendrogram shows the explanation cards collapsed into two major nodes. First, the creation schema (at the leftmost side of the figure). Second, a higher-level node comprising four clearly differentiated schemas (i.e., mutation, crossbreeding, selection, and dissemination), and a rather undifferentiated cluster formed

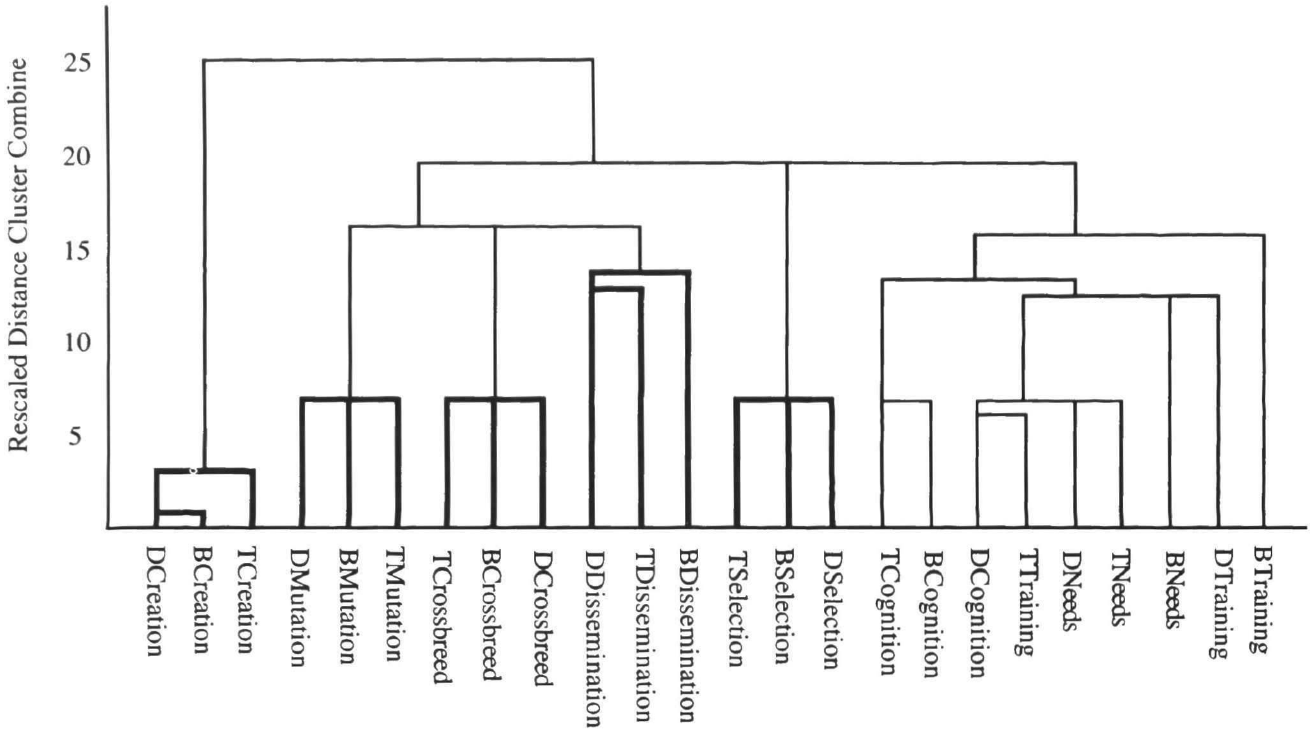

Figure 2: Dendrogram for schema-oriented subjects (N = 26)

by needs, training and cognition explanations. In short, this analysis indicates that these students were able to recognize five schemas: creation, mutation, crossbreeding, dissemination and static selection (see groups at the left edge in the figure). The clustering of needs, training and cognition explanations into a single node is consistent with the co-occurrences observed in the whole-sample cluster analysis.

Figure 3 shows the dendrogram for the cluster analysis performed on the co-occurrence matrix for twenty-three students whose sorting actions were mainly guided by species similarities. This dendrogram consists of two major clusters. The first one comprises all tiger explanations, and stands apart in the dendrogram (see the cluster to the right in the figure). The other cluster branches out into two nodes, one of them grouping all dinosaur explanations, and the other all bird explanations. Although no single explanatory pattern was fully recognized by this group of students, the dendrogram shows the three creation explanations standing relatively apart within their respective species clusters, which suggests that some students sorted together this type of explanations.

Discussion

The sorting data provides mixed evidence for the hypothesized explanatory patterns. Only one subject produced a sort that was exactly in accordance with all eight patterns. However, if we look at the sample as a whole, the students recognized the creation schema with relative ease. This schema was clearly isolated in two of the three cluster analyses. Although less salient, the explanatory schemas of selection, crossbreeding, mutation, and dissemination were also recognized. Explanations based on the schemas of cognition, training, and needs were systematically collapsed into a single explanatory category. This co-occurrence occurred both across species and within species.

These results suggest that some of the schemas that we initially attributed to students were too fine-grained. In particular, a broader pattern, comprising training, cognition, and needs explanations conforms better to the students' idea that the environment poses demands of various sorts and organisms respond by changing in such a way as to deal with those demands.

The results show that patterns that emerge in content analyses and think-aloud protocols may or may not correspond to actual knowledge structures in students' memory. Hypotheses drawn from discourse analyses should be validated against additional data. In particular, mental models, misconceptions, and explanatory patterns derived from production tasks should be checked against data from recognition tasks.

Given that there is partial evidence for some explanatory patterns, a number of issues need to be addressed in future research. First, on-line data should be collected to determine how students use these explanatory patterns. Following Schank and co-workers (Leake, 1992; Schank, 1986; Schank, Kass, & Reisbeck, 1994), we suggest that the production of evolutionary explanations by students involves three major cognitive processes: schema retrieval, schema instantiation, and explanation evaluation. Retrieval processes are particularly important because the activation of one or more patterns is a function of the degree to which both a subject's representation of evolution and his/her ex-

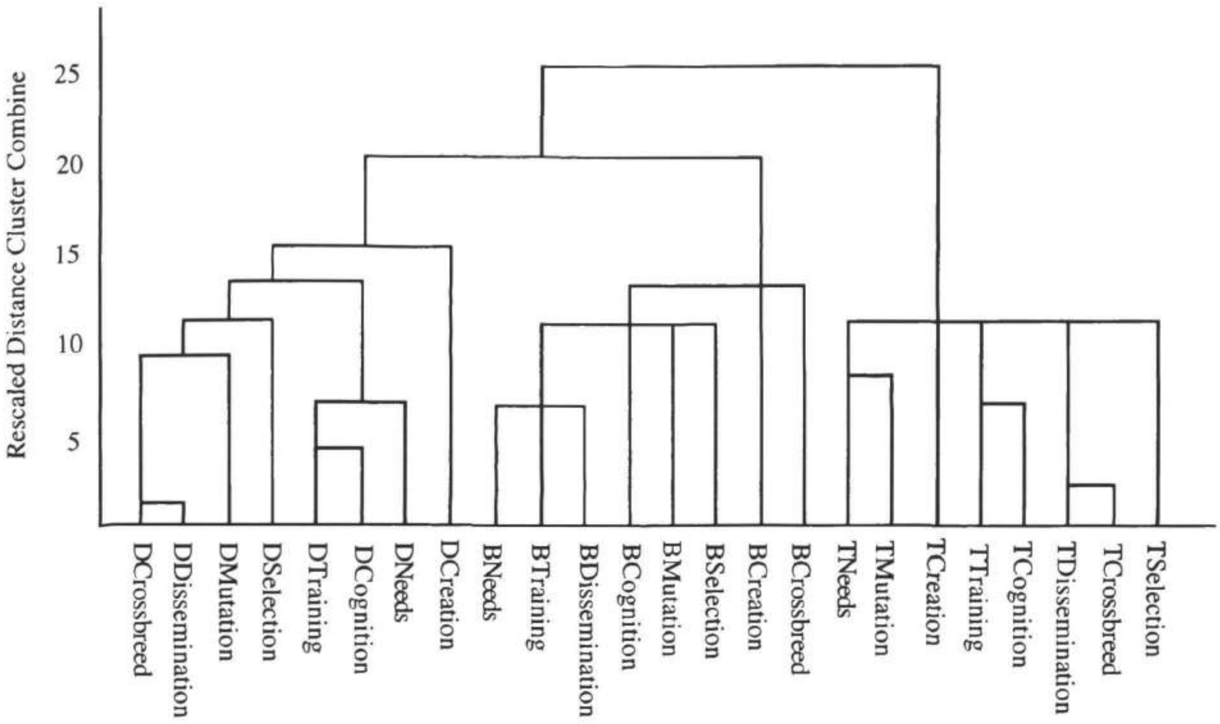

Figure 3: Dendrogram for species-oriented subjects (N = 23)

planatory goals match the way explanation patterns are indexed in memory.

For example, the co-occurrence of patterns in some explanations reveals the students' attempt to construct explanations involving multiple explanatory goals. Thus, a subject who claims that dinosaurs became gigantic because they needed to adapt and then a single mutation occurred, is focusing on the source of change (explanatory goal 1), as well as on the mechanism of change (explanatory goal 2). The co-occurrence of contradictory explanatory patterns within one and same explanation can be attributed to the absence of coherence revision during the evaluation phase of the explanatory process, or to the lack of domain-specific knowledge that may constrain the retrieval of certain schemas in the light of specific explanatory goals (e.g., use of creation and crossbreeding schemas to explain evolutionary change).

A second question is how the correct explanation pattern is acquired. The schema view implies three learning processes that we refer to as *relevance determination*, *concept differentiation*, and *coordination*. First, students enter a science course with a vast repertoire of schemas acquired in everyday experience as well as previous study. The majority of those schemas are irrelevant rather than incorrect. The students cannot know in advance which schemas are relevant for the intellectual game they are about to learn. Hence, schemas will be activated with a probability that is some function of their past usefulness. In the course of learning, students have to discover which of the activated schemas constitute valid moves in the intellectual game they are trying to master. The first step in learning is therefore to round up the relevant components of prior knowledge.

Second, the schema repertoire suggests that learning requires explication of what Carey (1992) calls undifferentiated categories. The concept of undifferentiated categories is well illustrated by the training schema. The key feature of this schema is the lack of distinction between phenotypic changes --ontogeny, maturation, seasonal variations, aging, effects of athletic training-- and evolutionary changes. Because many phenotypic changes are driven by activity, so is evolution. Without a clear distinction between acquired and inherited traits, the contrast between Lamarckianism and Darwinism that opens most textbook chapters on evolutionary theory cannot be understood. Correct understanding of the theory of evolution presupposes that biological categories are differentiated well beyond the level of semantic precision that is typical of common sense discourse.

Third, students' explanations are typically simpler than Darwinian explanations. Where students tend to appeal either to mutations or the environment or selection or dissemination, an expert explanation intertwines these ideas into a more complex mental model of evolution. Hence, the schema hypothesis implies that one fundamental dimension of learning is to assemble existing schemas into more complex schemas. This *assembly hypothesis* is related to the coordination process proposed by Piaget (1985). The relation between coordinated schemas is not mere association. Two schemas, each of which is well established already, are hooked together in such a way that their coordination acquires a new cognitive function. For example, one stepping stone on the way to correct understanding of Darwinism is probably to coordinate the ideas of selection and repetition into a new schema for repeated selection. This schema is more complex but also more powerful than either selection or repetition by themselves.

The view that novice competence consists of a repertoire of schemas has very different implications for the nature of learning than the misconception view (Hewson & Hewson, 1984; Lawson & Weser, 1990; Strike & Posner, 1985). Inspecting the schema repertoire we have uncovered in our studies shows that students' prior knowledge is often irrelevant, sometimes fuzzy and always too simple; it is not obviously incorrect or false. Instead of rejecting their prior knowledge, students have to determine the relevance of existing schemas, explicate undifferentiated concepts by making the relevant distinctions, and construct more complex schemas by coordinating existing schemas.

Although data from this study are far from conclusive as to the validity of the schema view, the results invite exploration of alternatives to the naive theory view of misconceptions in science learning. If students retrieve, instantiate, and coordinate explanatory patterns during learning effective pedagogical techniques should support these processes.

Acknowledgments

This research was supported, in part, by a grant to The National Research Center on Student Learning from the Office of Educational Research and Improvement (OERI), US Ministry of Education, and in part by fellowships from the Fulbright Commission and the Pan American Leo S. Rowe Fund, Organization of American States, to the first author.

References

Benysh, D. V., Koubek, R. J., & Calvez, V. (1993). A comparative review of knowledge structure measurement techniques for interface design. *International Journal of Human-Computer Interaction, 5* (3), 211-237.

Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. In R. Wyer & T. Srull (Eds.), *Handbook of social cognition* (pp. 119-160). Hillsdale: Erlbaum.

Carey, S. (1992). The origin and evolution of everyday concepts. In R. N. Giere (Ed.), *Minnesota studies in the philosophy of science. Volume XV. Cognitive models of science* (pp. 89-128). Minneapolis, MN: University of Minnesota Press.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5* (2), 121-152.

diSessa, A.A. (1988). Knowledge in pieces. In G. Forman & P.B. Pufall (Eds.), *Constructivism in the computer age* (pp. 49-70). Hillsdale, NJ: Erlbaum.

Hewson, P. W., & Hewson, M. G. A. (1984). The role of conceptual conflict in conceptual change and the design of science instruction. *Instructional Science, 13*, 1-13.

Kuhn, D. (1989) Children and adults as intuitive scientists. *Psychological Review, 96*, 674-689.

Larreamendy-Joerns, J. & Ohlsson, S. (1994). *The psychological reality of explanatory patterns in evolutionary biology* (Technical Report). Pittsburgh, PA: University of Pittsburgh.

Lawson, A. E., & Weser, J. (1990). The rejection of non-scientific beliefs about life: Effects of instruction and reasoning skills. *Journal of Research in Science Teaching, 27*, 589-606.

Leake, D. (1992). *Evaluating explanations: A content theory*. Hillsdale, NJ: Erlbaum.

McCloskey, M. & Kargon, R. (1988). The meaning and use of historical models in the study of intuitive physics. In S. Strauss (Ed.), *Ontogeny, phylogeny, and historical development* (pp. 49-67). Norwood, NJ: Ablex.

Ohlsson, S. (1993). Abstract schemas. *Educational Psychologist, 28* (1), 51-66.

Ohlsson, S. (1991). *Young adults understanding of evolutionary explanations: Preliminary observations*. Technical Report. LRDC. University of Pittsburgh.

Ohlsson, S., & Bee, N. (1992). *The effect of expository text on students' explanations of biological evolution* (Technical Report). Pittsburgh, PA: University of Pittsburgh.

Ohlsson, S., & Bee, N. (1993). *Constructing and reading evolutionary explanations improve (some) students' understanding of Darwin's theory* (Technical Report). Pittsburgh, PA: University of Pittsburgh.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. Chicago, IL: The University of Chicago Press.

Schank, R. (1986). *Explanatory patterns: Understanding mechanically and creatively*. Hillsdale, NJ: Erlbaum.

Schank, R., Kass, A., & Riesbeck, C. K., (Eds.) (1994). *Inside case-based explanation*. Hillsdale, NJ: Lawrence Erlbaum.

Strike, K. A., & Posner, G. J. (1985). A conceptual change view of learning and understanding. In L. West and L. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 211-231). New York: Academic Press.