In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society, 599-604*.

Assessing Explanatory Coherence: A New Method for Integrating Verbal Data with Models of On-line Belief Revision

Patricia Schank and Michael Ranney
Mathematics, Science, and Technology Division; Graduate School of Education
University of California; Berkeley, CA 94720
ranney@cogsci.berkeley.edu & schank@garnet.berkeley.edu

Abstract[1]

In an earlier study, we modeled subjects' beliefs in textually embedded propositions with ECHO, a computational system for simulating explanatory evaluations (Schank & Ranney, 1991). We both presumed and found that subjects' representations of the texts were not completely captured by the (*a priori*) representations generated and encoded into ECHO; extraneous knowledge likely contributed to subjects' biases toward certain hypotheses. This study builds on previous work via two questions: First, how well can ECHO predict subjects' belief evaluations when *a priori* representations are not used? To assess this, we asked subjects to predict (and explain, with alternatives) an endpoint pendular-release trajectory, while collecting believability ratings for their on-line beliefs; subjects' protocols were then "blindly" encoded and simulated with ECHO, and their ratings were compared to ECHO's resulting activations. Second, how similar are different coders' encodings of the same reasoning episode? To assess intercoder agreement, we examined the fit between ECHO's activations for coders' encodings of the same protocols. We found that intercoder correlations were acceptable, and ECHO predicted subjects' ratings well—almost as well as those from the more diminutive, constrained situations modeled by Schank and Ranney (1991).

Introduction

People often differentially evaluate the plausibility of similar or even identical beliefs when reasoning or arguing about a situation. How do people decide what description of the world is most plausible? Thagard (1989) and others characterize the plausibility of a belief as generally increasing with its increasing simplicity (e.g., fewer necessary cohypotheses), increasing breadth (i.e., more coverage of observation), and decreasing competition with alternate (especially entrenched) beliefs (cf. Johnson & Smith, 1991). These principles play important roles in evaluations of the quality of an explanation (Schank & Ranney, 1991; Read & Marcus-Newhall, 1991).

Science is a rigorous interpretive system that we overlay on our experiences to understand and use them. Although individuals' beliefs may not fit precisely into a scientific framework (i.e., their beliefs may conflict with established scientific hypotheses), people often hold their beliefs as long as they help explain many of their experiences. For example, students learning physics tend to hold strong intuitive beliefs about the physical world that tend to resist revision (e.g., Ranney 1987/1988; Hartley, Byard, & Mallen, 1991). Ranney and Thagard (1988) characterize belief revision as the result of seeking explanatory coherence between theories and observations. The Theory of Explanatory Coherence (TEC) is intended to account for a variety of explanatory evaluations. This theory has been implemented in a computational model called ECHO, based on the claim that beliefs and data are related explanatory entities, and evaluating their plausibility is an interactive, principled, coherence-seeking process (Thagard, 1989; Ranney, in press).[2]

We describe here an empirical study that extends our previous research (Schank & Ranney, 1991) by focusing on two questions: First, how well can ECHO predictively model how strongly individuals believe the assertions they make in the course of an explanation or argument? Second, how subjective is the ECHO encoding process? To further assess ECHO's predictive ability, we asked subjects to predict the path a bob follows when released from the endpoint of a pendular swing, to explain their own (and others') predictions, and to rate the strength of the beliefs used. Their verbal protocols were then encoded and simulated with ECHO, and its activations compared to the subjects' ratings. To assess intercoder reliability, multiple coders encoded the protocols, and we examined differences among ECHO's eventual activations with respect to the coders' encodings of identical protocols.

---

[1] Preparation of this article was supported by the Spencer Foundation, the Evelyn Lois Corey Fund, and a Faculty Research Grant from the University of California, Berkeley.

[2] ECHO's "theoretical/systemic" coherence differs from (and is generally orthogonal to) standard notions of "linguistic" coherence (Ranney, Schank, & Ritter, 1992). In ECHO, coherence is seen from the perspective of competing theories, where the dynamic tension represented as explicitly conflicting theories *reduces* the overall coherence of a system of propositions (compared to a single-theory network). In contrast, textual/discourse coherence is generally viewed as increasing with more explicit relations among various entities and assertions in a text, and less reliance on implicit background knowledge for making inferences (such as anaphoras). (E.g., the textual stimuli used in Schank & Ranney, 1991 were designed to be low in systemic coherence and high in linguistic coherence).

TEC and its ECHO Model

In TEC, coherence involves relations among two or more propositions that may "hold together" or "resist holding together." (We use "proposition" for something proposed, e.g., a piece of evidence or hypothesis, such as "gravity pulls objects down"—in contrast to a concept, such as "gravity.") For current purposes, the following principles establish the local pairwise relations among cohering and incohering propositions (nb. these principles are, selectively, from Schank & Ranney, 1991, Thagard, 1992, and Ranney & Thagard, 1988): (1) Coherence and incoherence are symmetric relations. (2) Hypotheses that together explain a proposition cohere with each other and with the explained proposition. (3) Simplicity: The plausibility of a belief is inversely related to the number of cohypotheses it needs to explain a proposition. (4) Data Priority: Results of observations, such as evidence and acknowledged facts, have a degree of acceptability on their own. (5) Contradictory hypotheses incohere. (6) The acceptability of a proposition depends on its coherence within the system of propositions in which it is embedded. A proposition's acceptability increases as it coheres more with other acceptable propositions and *incoheres* more with *unacceptable* propositions. (In ECHO, a proposition's acceptability is measured by its activation value, ranging from -1, complete rejection, to 1, complete acceptance.) (7) The overall coherence of a network of propositions depends on the local pairwise cohering of its propositions.

Schank and Ranney (1991) and Read and Marcus-Newhall (1991) show that these principles play important roles in explanations. They found that subjects prefer explanations that account for more data, are simpler, and involve hypotheses that can be further explained. Subjects' evaluations of explanations are also changed by the availability of competing explanations.

ECHO uses a connectionist architecture in which each node represents a proposition. Hypothesis evaluation is treated as the satisfaction of many constraints, determined from the explanatory relations and from a few parameters that provide degrees of freedom. Given declared input propositions and relations between them, node activations are updated using a simple settling scheme. For more complete descriptions of ECHO's algorithms, see Thagard (1989 & 1992), Schank and Ranney (1991), and Ranney and Thagard (1988).

Why Physics, and Why ECHO?

This study focuses on questions regarding (a) ECHO's ability to predictively model individual subjects' evaluations of their beliefs about physical motion (encoded from verbal protocols), and (b) intercoder reliability regarding ECHO. We chose to model beliefs about motion since studies have shown that individuals tend to persistently hold naive beliefs about the natural world that sometimes conflict with scientific explanations (e.g., Brewer & Chinn, 1991; Ranney, 1987/1988).

ECHO has been used (mostly *ex post facto*) to model changes in subjects' beliefs about physical motion (Ranney & Thagard, 1988), scientific and juror reasoning (Thagard, 1989, etc.) and social situations (e.g., Read & Marcus-Newhall, 1991), and to examine and foster students' scientific reasoning skills (e.g., Carlock, 1990; Ranney, in press). We used ECHO to model subjects' beliefs in textually embedded propositions, and found that subjects sometimes entertained competing hypotheses as nonexclusive and seemed to presume an implicit backing (i.e., other evidence or beliefs) that supported certain hypotheses (Schank & Ranney, 1991). Despite attempts to decontextualize texts, our subjects (not surprisingly) brought extraneous knowledge to bear when reasoning about the texts' statements. Consequently, the subjects' representations of the texts were not completely captured by the representation encoded into ECHO, and this unrepresented extraneous knowledge likely contributed to their relative biases. We coarsely modeled this tendency for subjects to presume backing behind "superordinate" hypotheses in ECHO by assigning them a fraction of *data priority* (usually reserved for evidence). Still, the study raised the question: Can ECHO predict subjects' beliefs as well or better if they make their implicit backings (coarsely modeled by Schank & Ranney, 1991) explicit?

Other computational models. Several models of explanation evaluation and belief change are compatible with TEC and ECHO. Ranney (in press) points out that TEC does not explicitly account for memorial capacity and processing limitations, inspiring Bar-On's (1991) theory of local coherence within *views,* an attempt to account for attentional and short term memory effects via limited capacities. Bar-On argues that localist connectionist models provide more appropriate levels of abstraction (than distributed models) for simulating locally coherent views. Also similar to ECHO is HEIDER, Gabrys' (1989) simulation, which seeks consistency (coherence) within its world view in the face of new information.

ECHO may initially seem less compatible with other computational models. For example, Ram and Leake (1991) argue that people best learn to accept explanations when they come with needed information, and present a goal-based computational model that focuses on finding "useful" (vs. "valid") explanations by incorporating the goals into explanatory evaluations. Okada and Klahr (1991) code subjects' naive, complex, idiosyncratic, beliefs (garnered from transcribed protocols) as a hypothesis space, but they view belief revision as a search through this space of beliefs (vs. parallel constraint satisfaction, as in ECHO). However, both of these models highlight *goal-* or *utility-based* reasoning, which ECHO does not attempt to model. These models are more comparable to MOTIV-ECHO (Thagard, 1992), a program that allows ECHO's inferences to be biased by goals.

Models of text and discourse analysis. Our previous methodology was limited in that the extraneous knowledge subjects brought to bear when reasoning about assertions in the texts were not completely captured by the representation encoded into ECHO (Schank

& Ranney, 1991). Hence, we asked subjects in the current study to *explain* their beliefs and predictions, and encoded their protocols as ECHO belief networks.

Other researchers represent mental interpretations of text and discourse by systems of interrelated propositions (e.g., Trabasso, van den Broek, & Suh, 1989; Givon, 1991). Kintsch's work (in press, cf. 1988) supports the use of higher-level (e.g., causal or explanatory) relations between belief propositions as a powerful level of abstraction. He describes text comprehension as the construction of representations consisting of primary concepts and higher level propositions, in a network with associative relations at the conceptual and propositional levels and causal (explanatory) relations at the proposition level. Including both types of links in the representation enabled more accurate predictions of subjects' immediate recall (r=.76), but causal links alone explained most of the variation (r=.61).

Method

In this study, we use ECHO encodings, produced from a qualitative analysis of the subjects' protocols, for a more ecologically valid and comprehensive test of the ECHO model and encoding schemes (cf. Ranney and Thagard, 1988). Design decisions followed the desire to represent subjects' explanations and believability ratings as completely and accurately as possible, and to assess intercoder reliability for ECHO networks. We used a novel combination of convergent methods (cf. Ranney 1987/1988) in that we collected (a) drawn trajectory predictions, (b) quantitative believability ratings (to avoid subjective bias about the strength of subjects' beliefs), and (c) verbal protocols (with evaluative comments and ratings edited out), used as the basis for the ECHO encodings.

Subjects and Procedure

Ten subjects, four men and six women, were chosen from the University of California (Berkeley) student population, from responses to an advertisement. The subjects had various backgrounds, but little or (usually) no formal physics background. During the 30-60 minute sessions, subjects were asked to make predictions about pendular–release situations, and to rate the believability of the hypotheses and evidence they verbalized as they reasoned about the task. The interviewer recorded subjects' beliefs on paper (in real-time, as they completed an utterance) using the subjects' terminology. (Subjects were later given feedback on the situations' outcome, and then again asked to rate the believability of the same, noted, propositions.) Audiotaped protocols were collected and transcribed for encoding and intercoder reliability analyses.

Tasks. Subjects were first shown an animated pendular–release situation (from Ranney, 1987/1988) in which the swinging motion of the pendulum–bob was frozen in time at the endpoint of its swing. Subjects were asked to imagine that the pendulum string broke at the extreme of the swing and/or the bob was released, and to predict (draw and explain) the subsequent trajectory of the bob. They were then shown, serially, five alternative, commonly predicted, paths (generated by subjects in Ranney, 1987/1988, and in a recent pilot study), and asked to explain why each path may or may not be correct.

As a subject reasoned out loud about the plausibility of the paths, the interviewer noted the subject's assertions. After the subject finished reasoning about the endpoint-release situation, the interviewer read back to the subject the list of beliefs she had noted. Subjects were then asked to rate (on a scale from 1, "completely unbelievable," to 9, "completely believable") how strongly they believed the propositions they had verbalized, and to rate how strongly they believed in the alternative paths (now displayed in parallel).[3]

Encoding. Subjects' stated believability ratings were edited out of copies of the transcribed protocols, as were evaluative statements that qualitatively revealed the strength of their beliefs. The edited protocols were then encoded into ECHO–style input by one to four, variously experienced, "blind" coders. (I.e., coders segmented and categorized subjects' assertions into beliefs, evidence, explanations, and contradictions.) Encodings of only the first (prediction) part of the first (pendular–release) task are reported here (see footnote 3). This portion generally accounted for over half of the transcribed sessions.

Coders 1 (who coded all of the protocols at least once) and 2 had experience encoding previous protocols (e.g., from Ranney, 1987/1988, and the pilot study), but coders 3 and 4 had virtually none. Coders discussed encoding principles at length, and incorporated agreed–upon principles into an "explanation encoding guide" that included a list of syntactic and grammatical cues that tend to indicate explanation structures (e.g., verbs that signal beliefs;

---

[3] Data from what follows are not analyzed here: After elaborating on their predictions on the endpoint pendular-release task, subjects read and made predictions about an isomorphic playground swing situation. Subjects were asked to predict the path that a child leaving the seat at the endpoint of a playground swing would follow, and to rate both their beliefs (again, noted "on-line" by the interviewer) and their predicted path.

Following their predictions in the pendular–release and playground situations, subjects were allowed to change their predictions, then given trajectory feedback: The pendular-release task's simulation was repeated and subjects were asked (1) if they chose to modify or draw a new path, and (2) to rate their beliefs (read back by the interviewer), including the alternative paths (displayed in parallel). Subjects were then shown the *actual* path (i.e., a dynamic feedback simulation of both the swing and the subsequent vertical trajectory after release; Ranney, 1987/1988). They were then asked, with hindsight, (a) to try to explain why the (usually surprising; Ranney & Thagard, 1988) feedback was correct, and (b) to again rate the strength of their beliefs. Similarly, on the isomorphic playground task, subjects could change their prediction and re-rate their beliefs before being given (verbal) feedback. After feedback (indicating a vertical, post-release trajectory), they were asked to give a hindsight explanation of the vertical trajectory, and to re-rate their beliefs.

conjunctions and verbs that signal explanations; negations, conjunctions, and verbs of conflict that signal contradictions;[4] see Table 1 for an example encoding).

ECHO simulations of the encodings were run with the parameter settings used by Schank and Ranney (1991), which were also in the midrange of those used by Ranney and Thagard (1988).[5] Each subject's belief ratings (i.e., of assertions encoded as ECHO propositions for which the subject reported a rating) were then compared with the ECHO activations from the simulations. Comparisons between ECHO's activations and subjects' ratings (just prior to the time of feedback) were made for each subject–coder pair and for each coder overall. For each subject, ECHO activations among coders were also compared.

By having coders unfamiliar with the subjects' (excised) believability ratings and qualitative evaluations encode the protocols, we could use the results of ECHO simulations on these encodings to assess how well ECHO predicts *a priori* the subjects' beliefs. Thus, we compared ECHO's activations for the various encodings with subjects' evaluations (i.e., believability ratings). We assessed intercoder agreement by comparing different encodings and simulation results of the same protocols.

Table 1. Example protocol and its encoding.

Convention:	**Indicates:**
underlines	explanations
italics	contradiction/competition
[bracket]	beliefs/data propositions

Protocol: "I've changed my mind. (no coding; monitoring statement.). I like P1[the arch path]. I think P1[the bob will fly out in an arch curve] since H1[its got motion right] and H2[up], and H3[gravity is pulling it down]. *But* H1[it has motion right] so I guess P2[the diagonal path] is possible. NP3[It won't drop straight down], though, because H1[it has motion right]. Hm. I remember E1[jumping off a swing and flying out in an arch], though, so I think P1[its going to fly up and out like it does in that arch path you have there]. Yeah, I really think P2[the diagonal path] won't happen (no coding; evaluative statement.). "

Encoding:

P1 The arch path	P2 contradicts P1 (implicit)
H1 Bob has motion right	NP3 Not straight down path
H2 Bob has motion up	H1 explains NP3
H3 Gravity is pulling down	E1 I fly out in an arch when I jump off a swing
H1, H2, and H3 explain P1	data E1
P2 Diagonal path	E1 explains P1
H1 explains P2	

Results

Figure 1 shows subjects' average ratings (prior to feedback) for each path. We computed correlations between ECHO's activations and subjects' ratings for each subject–coder pair (see Figure 2), and for each coder. Analyses of variance of the subjects' ratings were performed; we also computed correlations between ECHO's activations among the various coders for each protocol. Analyses were computed both (a) over all beliefs, and (b) for the path propositions only.

Protocol and Encoding Analyses

Encoded portions of the protocols averaged about 600 words in length. On average, about 23 propositions, 2 data (e.g., observations or remembered experiences), 13 explanations (3 of which were considered "implicit" by coders), and 21 contradictions (17 of which were considered implicit, including 15 essential contradictions between the competing release paths) were encoded per protocol. As expected, subjects did not make all of their explanations and contradictions explicit, even with prompts for elaboration (cf. Grice, 1975, on the conversational *maxim of quantity*, which predicts that people typically avoid being overly informative). The interviewer captured on–line (and hence had subjects rate) about 60% of the belief propositions that were later encoded by coders, so ratings for 60% of the encoded propositions were available for comparison with ECHO activations. Ratings for most of the subjects' key central beliefs, and all alternative trajectories, were collected. Analyses of the encodings also revealed that beliefs *not* rated played more peripheral explanatory roles, compared to their rated counterparts. (E.g., compared to rated beliefs, unrated beliefs were about twice as likely to not be part of any explicit explanation of a path, and about *thrice* as likely to not be part of *any* explicit explanation. Further, rated beliefs were about twice as likely as their unrated counterparts to be within two explanatory links from any path-prediction, and almost 50% more likely to *directly* explain a path-prediction. These differences were significant at p<.01.)

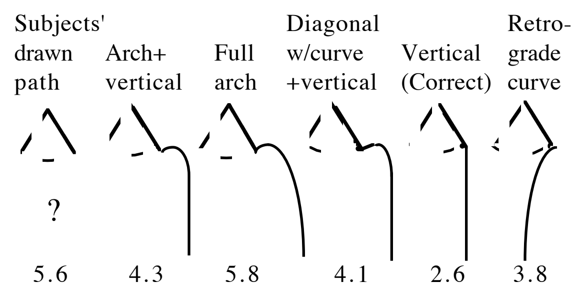

Figure 1. Mean path believability ratings (1-9 scale; prior to feedback). (Nb. Four subjects drew unique paths not among the 5 alternatives. Mean ratings for drawn paths were about 5.6 both for unique paths and for non–unique paths.)

[4] We encoded the alternative pendular–release trajectories in this study as contradictory, based on the assumption that subjects believe one unique trajectory exists.

[5] The parameter values used were: decay = .04, excitation = .03, inhibition = -.06, and data excitation = .055.

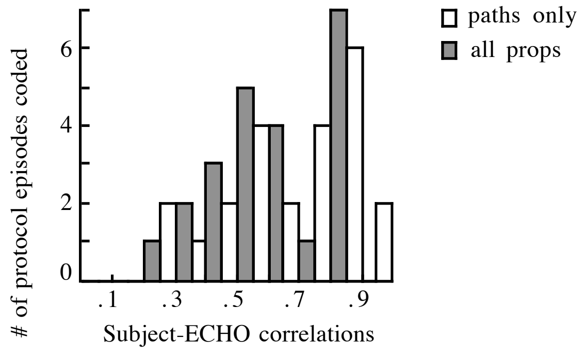

Figure 2. Distribution of correlations between ECHO's activations and subjects' ratings on *all* rated propositions (grey) and on endpoint-release paths only (white).

Prediction of Subjects' Ratings

ECHO simulations were run with the implicit explanations and contradictions, added by the coders, included. Final ECHO activations, over all rated beliefs, were positively correlated (at $p<.05$) with subjects' ratings in 14 of the 23 simulations. Correlations were as high as $r >.80$ on seven simulations, and as low as $r=.25$ on one (see Figure 2).[6] Overall, there was a significant positive correlation ($r =.56$) between ECHO's activations and the subjects' ratings ($r=.65$ for most highly correlated coder).

Correlations between ECHO activations and the subjects' ratings for the path-propositions alone mirrored those computed over all beliefs, with correlations as high as $r= .99$ and as low as $r=.23$. The overall correlation for paths was nonsignificantly higher (at $r=.61$) than the correlation over all beliefs ($r=.56$). The overall correlations were also slightly higher for most of the coders ($r=.70$ for the most highly correlated coder).[7] In general, results from coders who coded the most protocols yielded higher correlations between coding–based ECHO activations and subjects' ratings.

Intercoder Agreement

Over all beliefs, and for the paths alone, ECHO's final activations were significantly correlated with subjects' ratings for coders 1 and 2 ($r >.57$; $p<.0001$). These coders were the most experienced, and encoded more of the protocols (and thus had more data to correlate); the correlations between ECHO's final activations and subjects' ratings generally increased with encoding experience.

Comparing ECHO networks to assess intercoder agreement is difficult; tractable methods of comparing such network topologies have not, to our knowledge, been developed. Therefore, we used approximating measures to gauge intercoder agreement. We analyzed the propositions common among coders; pairs of propositions were judged to be the same if both their wording and source locations in the protocol were virtually the same. On average, about 60% of encoded propositions were judged to be the same between pairs of coders. For these, the overall intercoder correlation (i.e., between ECHO activations for common propositions) was significant ($r=.49$; $p<.001$).

Analysis of variance results for the subjects' ratings over all beliefs (with ECHO activations as a covariate) later indicated that ECHO's activations explain about 28% of the variance in the subjects' ratings, while individuals account for about 8% of the variance in the ratings (both $p<.001$). As one would hope, the ANOVAs also indicate that *none* of the variance is accounted for by the coder, so systematic coder effects are negligible. Similar results are found when only the path propositions are considered, except that the type of path considered also accounts for a significant amount of variance.

Discussion and Conclusions

Schank and Ranney (1991) raised the question of whether ECHO can predict subjects' beliefs as well as or better than they had observed ($r >.7$) if subjects make their "implied backings" explicit. However, we found slightly lower ECHO-subject correlations (about $r =.6$).[8] This might suggest that ECHO does not predict subjects' beliefs better (or perhaps even as well as) when they make their implicit backings explicit. There are several reasons to resist this conclusion. First, of the ten individuals' ratings in this study, six were predicted with a correlation of $r >.80$ (seven, when only the paths were considered); the data also suggest that the ECHO-subject correlation tends to increase when only the (more central) path beliefs are examined. Second, combined with other variables, ECHO helped account for about 40% of the variation in subjects' ratings. Third, the texts used in our prior study reflected topologies defined *a priori*, which likely constrained subjects' representations of the situation. (I.e., subjects in Schank & Ranney, 1991, were not encouraged to elaborate on their beliefs and bring other knowledge into their representations, as they were encouraged to do here). For these (and other content– and context–dependent) reasons, the task of modeling the subjects' beliefs in Schank and Ranney (1991) was, in essence, of smaller scale. Fourth, the ECHO networks generated here were, by salient measures (e.g., the number of propositions, the number of links), about two to over 20 times larger and much less explicit than the networks in the our prior

[6] Earlier versions of this article, due to a single coding error, spuriously reported two negatively correlated simulations.

[7] As predicted by Ranney and Thagard's (1988) simulation, feedback on the vertical *pendular–release* trajectory influenced subjects' believability ratings for the (uncoded) isomorphic *playground* situation: Belief in predicted, non–vertical release paths for the playground swing went from strong belief (mean = 7.1) prior to pendular feedback, to slight disbelief (mean = 4.0) after such feedback.

[8] Recall Kintsch's (in press) aforementioned $r=.61$ between his activations and his subjects' propositional recall.

study. This extra complexity may have caused difficulties for subjects who, unlike ECHO, have limited attention and memory (Ranney, in press). We plan to incorporate such limitations in future modeling efforts.

In sum, ECHO predicted subjects' ratings fairly well, and the overall intercoder agreement was acceptable (r=.49; so, the coders' simulations currently correlate better with the subjects' data than with each other). Still, the model did not fit the data as well as one might have expected, based on the more diminutive, constrained, theoretical conflicts modeled by Schank and Ranney (1991). But the correlation between ECHO's activations and subjects' believability ratings seems to increase with encoding experience, suggesting that the encoding process is successful and refinable. Further, this study's method is novel, so these nascent attempts to establish intercoder reliability will, no doubt, improve. However, the moderate intercoder correlation also suggests that, for modeling purposes, a better approach may be to have subjects encode their own representations directly into ECHO—with a user–friendly interface (e.g., Carlock 1990). Thus, we are now incorporating ECHO into a computer-based learning environment in which subjects can directly encode their own representations. We plan to use the environment ("the reasoner's workbench;" Ranney, in press) to teach coherent argumentation, and to encourage students to think about consistency and coherence as metrics of reasoning and rationality.

Acknowledgements

We thank Peter Pirolli, Paul Thagard, Margaret Recker, Marsha Lovett, Christopher Hoadley, Joshua Paley, Edouard Lagache, Michelle Million, two reviewers, and the Reasoning Group for their help and comments.

References

Bar-On, E. (1991, May). *Mental capacity and locally-coherent views: Towards a unifying theory*. Draft. Department of Science Education, Technion, Israel.

Brewer, W., & Chinn, C. (1991). Entrenched beliefs, inconsistent information, and knowledge change. In L. Birnbaum (Ed.), *Proceedings of the International Conference on the Learning Sciences*, 67-73. Charlottesville, VA: Association for the Advancement of Computing in Education (AACE).

Carlock, M. (1990). *Learning coherence through MacECHO IFE: An analysis of coherence*. Second Year Project Report, Univ. of California, Berkeley, Graduate School of Education.

Gabrys, G. (1989). HEIDER: A simulation of attitude consistency and attitude change. In S. Ohlsson (Ed.), *Aspects of cognitive conflict and cognitive change.* Univ. of Pittsburgh, Technical Report, KUL-89-04.

Givon, T. (1991, August). *Coherence: Toward a cognitive model*. Paper presented at the First Annual Meeting of the Society for Text and Discourse.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.). *Syntax and Semantics*, vol 3: Speech Acts. New York, NY: Seminar Press.

Hartley, J. R., Byard, M. J., & Mallen, C. L. (1991). Qualitative modeling and conceptual change in science students. In L. Birnbaum (Ed.), *Proceedings of the International Conference on the Learning Sciences*, 222-230. Charlottesville, VA: AACE.

Johnson, T. & Smith, J. (1991). A framework for opportunistic abductive strategies. *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, 760-764. Hillsdale, NJ: Erlbaum.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163-182.

Kintsch, W. (in press). How readers construct situation models for stories: The role of syntactic cues and causal inferences. To appear in A. F. Healy, S. Kosslyn, & R. M. Shiffrin (Eds.), *Essays in honor of William K. Estes*, vol. 2. Hillsdale, NJ: Erlbaum.

Okada, T., & Klahr, D. (1991). Searching an hypothesis space when reasoning about buoyant forces: The effect of feedback. *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, 842-846. Hillsdale, NJ: Erlbaum.

Ram, A. & Leake, D. (1991). Evaluation of explanatory hypotheses. *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, 867-871. Hillsdale, NJ: Erlbaum.

Ranney, M. (1988). Changing naive conceptions of motion (Doctoral dissertation, University of Pittsburgh, Learning Research and Development Center, 1987). *Dissertation Abstracts International, 49*, 1975B.

Ranney, M. (in press). Explorations in explanatory coherence. To appear in E. Bar-On, B. Eylon, and Z. Schertz (Eds.) *Designing intelligent learning environments: From cognitive analysis to computer implementation*. Norwood, NJ: Ablex.

Ranney, M., Schank, P., & Ritter, C. (1992, January). *Studies of explanatory coherence using text, discourse, and verbal protocols*. Paper presented at the Third Annual Winter Text Conference, Jackson, Wyoming.

Ranney, M. & Thagard, P. (1988). Explanatory coherence and belief revision in naive physics. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society,* 426-432. Hillsdale, NJ: Erlbaum.

Read, S. J. & Marcus-Newhall, A. (1991). *The role of explanatory coherence in the construction of social explanations*. Draft. Univ. of Southern California.

Schank, P., & Ranney, M. (1991). An empirical investigation of the psychological fidelity of ECHO: Modeling an experimental study of explanatory coherence. *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, 892-897. Hillsdale, NJ: Erlbaum.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12*, 435-502.

Thagard, P. (1992). *Conceptual Revolutions*. Princeton, NJ: Princeton University Press.

Trabasso, T., van den Broek, P., & Suh, S. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes, 12*, 1-25.