Hypotheses and Evidence About Evidence and Hypotheses

Christine Diehl (CDiehl@Socrates.Berkeley.Edu)
Michael Ranney (Ranney@Cogsci.Berkeley.Edu)
Grace Lan (Graslawn@Uclink4.Berkeley.Edu)
Sergio Castro (Floyd@Uclink4.Berkeley.Edu)
Graduate School of Education; 4533 Tolman Hall
University of California, Berkeley, CA 94720 USA

Introduction

We report on two studies that investigate students' understandings of the relationship between hypothesis and evidence. The *Convince Me* software and associated reasoning curriculum developed by the ECHO Educational Project aid students in generating and analyzing scientific arguments, requiring students to identify propositions that fill the roles of evidence and hypothesis (Ranney, Schank & Diehl, 1996). The distinction between evidence (or data) and hypothesis (or theory) appears to be fundamental in scientific reasoning, yet research shows that students seem to differ from scientists in their process of differentiating hypothesis and evidence. However, even experts do not always exhibit good agreement regarding this distinction. Although training with *Convince Me* lends sophistication to students' epistemic criteria, making their categorization appear more expert-like, the categorization of individual propositions may still be disputed (Ranney, Schank, Hoadley, & Neff, 1994). We hypothesize that students rely on prototypical models of hypothesis and evidence in structuring their arguments. Prior analysis of students' *Convince Me* arguments points to linguistic markings of this prototypic representation, which we investigate with paper-and-pencil surveys.

Study 1

Sixty-three undergraduate students completed a survey asking them to list five to seven very good examples each of hypotheses and evidence. We coded the surveys for the presence of fourteen linguistic categorical features, then used a stepwise multiple regression analysis to identify the best ("prototypical") models for propositional categorization. The model for evidence explained more variance than the model for hypothesis and included only one linguistic marker, the specifier category (e.g., when, which), as well as past tense and non-sentential phrases referencing objects or actions (e.g., fingerprints, data analysis). The predictor model for hypothesis included the modal (e.g., may, can) and causal (e.g., because, results in) linguistic marker, as well as references to a particular scientific theory.

Study 2

Sixty-two undergraduate students completed a survey with twenty-one propositions for which they were asked to indicate how good an example of a *hypothesis* or of a piece of *evidence* the statement/phrase is. One-third of the propositions contained only linguistic features for the hypothesis model identified in Study 1; one-third contained only features for the evidence model; and the last third were intended to be ambiguous with respect to categorization and contained either no such features or conflicting pairs of features. A multiple regression analysis shows that proposition type is a significant predictor of both hypothesis-ness and evidence-ness.

Discussion

Our results support the claim that the presence of certain linguistic features can predict the epistemic categorization of a proposition. Propositions that contained linguistic features representative of "prototypical" hypothesis were rated as being more like hypothesis than evidence. Propositions that contained linguistic features representative of "prototypical" evidence were rated as being more like evidence than hypothesis. Subjects considered ambiguous propositions with conflicting linguistic features to be more like hypothesis and ambiguous propositions with no obvious epistemic features to be more like evidence. This result indicates that evidence may be the "default" categorization of an epistemic statement, where as a single "hypothetical" feature biases an entire statement toward seeming hypothetical in nature. Extensive use of evidentials (especially modals) to express the strength and confidence of an assertion is evident in scientific discourse and writing; therefore, it seems reasonable that students should use these linguistic features in evaluating a proposition's epistemic status and determining its role assignment in an argument's structure (Crismore & Farnsworth, 1990).

References

Crismore, A. & Farnsworth, R. (1990). Metadiscourse in popular and professional science discourse. In W. Nash (Ed.), *The writing scholar: Studies in academic discourse*. Newbury Park, CA: Sage Publications.

Ranney, M., Schank, P., & Diehl, C. (1996). Competence versus performance in critical reasoning: Reducing the gap using *Convince Me*. *Psychology Teaching Review,4*(2), 153-166.

Ranney, M., Schank, P., Hoadley, C., & Neff, J. (1994). "I know one when I see one": How (much) do hypotheses differ from evidence? In the *Proceedings of the Fifth Annual American Society for Information Science Workshop on Classification Research*.