Jun 3rd, 9:00 AM - Jun 6th, 5:00 PM

Reply to my Commentator - Hitchcock

David Hitchcock
McMaster University

Reply to my Commentator

DAVID HITCHCOCK

Department of Philosophy
McMaster University
1280 Main St. W.
Hamilton, Ontario, L8S 4K1
Canada
hitchckd@mcmaster.ca

Scott Jacobs' commentary points the way toward further empirical research on the applicability of the covering generalization approach to evaluating inferences in natural language. There are however challenges in designing such research.

It is an easy task to secure inter-rater reliability in the identification and standardization of claim-reason complexes, as was done in the previous study (Hitchcock 2002). It is more difficult to determine what sort of inter-rater reliability can reasonably be expected in applying the covering generalization approach to a standardized inference. The covering generalization approach does not seek to reconstruct what covering generalization, if any, the arguer might have had in mind in making the inference. Nor does it attempt to determine what covering generalization the inference needs for the conclusion to follow. The task is one of evaluation, not of reconstruction by an analyst. The evaluative question is whether at least one covering generalization of the inference holds non-trivially, either always or for the most part or in the absence of exception-making circumstances. One can reasonably expect inter-rater reliability on the question whether a proposed inference-licensing principle is in fact a covering generalization of the inference. One should not expect, however, to get inter-rater reliability on which covering generalization of the inference is the most plausible among an indefinitely large number of such covering generalizations. Nor should one expect to get inter-rater reliability in judgments of acceptability of a given covering generalization. Such covering generalizations are typically substantive and not merely formal. Hence judgments of their acceptability will vary according to background knowledge, normative and evaluative assumptions, and skill in generating counter-examples. Perhaps the best that can be expected is that divergent judgments about whether and how an inference is valid will yield after back-and-forth discussion either to consensus or to clarity about the grounds of an unresolved disagreement. Independent third-party consideration of cases of the latter sort may help. In any event, it would be important to keep a record of the results of initial application of the approach by different evaluators, as well as of the content of any subsequent back-and-forth discussion.

As for projection of the results of the present study and of (Hitchcock 2002) to all natural-language inferences, the present sample comes from a much more limited universe than the previous sample of inferences in the English-language books in the library of a research-intensive university. Samplings of natural-language inference from

Hitchcock, D. (2009). Reply to my Commentator. In: J. Ritola (Ed.), *Argument Cultures: Proceedings of OSSA 09*, CD-ROM (pp. 1-2), Windsor, ON: OSSA.

other universes are possible. Judgments will need to be made as to whether the results of such further samplings are worth the time required to obtain them.

Link to commentary Link to paper

REFERENCE

Hitchcock, D. (2002). Sampling scholarly arguments: a test of a theory of good inference. In: H.V. Hansen, C.W. Tindale, J.A. Blair, R.H. Johnson and R.C. Pinto (Eds.), *Argumentation and Its Applications* (CD-ROM), Windsor, ON: OSSA.