

Exploring the Use of Factor Analysis to Understand the Nature of Verbal Protocols

Stacey A. Todaro (shaberkl@niu.edu)
Joseph P. Magliano (jmagliano@niu.edu)
Department of Psychology, Northern Illinois University
DeKalb, IL 60115

Recently, there have been a number of studies that use verbal protocols as a method for understanding processes that give rise to comprehension. In most verbal protocol procedures, students read sentences in a text and at specific points in the text are required to report their understanding. The thoughts that participants produce can include information from many different sources. Typically, the different sources of information are thought to reflect inferential processes (Trabasso & Magliano, 1996) and reading strategies (Magliano & Millis, 2003; McNamara, 2004) that give rise to comprehension. For example, when using a self-explanation strategy, it is commonly thought that the reader uses information from the current sentence, the prior discourse context, and general world knowledge. However, when making a bridging inference, it is assumed that the reader primarily uses information from the prior text. Traditionally, these inferential processes and strategies are identified based on theory and experimenter intuition. Therefore, the goal of the present study was to empirically assess the extent to which these different information sources map onto these strategies and inferential processes using principle components analysis.

Method

The study included 70 participants from Northern Illinois University who were enrolled in an introductory psychology course. Three scientific texts were used in the study. The topics of the texts were the origin of coal, the development of thunderstorms, and heart disease. Each text ranged between 20 and 34 sentences in length. For each of three texts, five sentences were selected as target sentences for which participants were prompted to type a verbal protocol. In typing a response, participants were instructed to report their understanding of the sentence in the context of what they have read thus far.

Protocol Analysis

The think-aloud protocols were first parsed into clauses using the criterion advocated by Trabasso and Magliano (1996), which involves identifying clauses based on the presence of main verbs. The information source for the verbs, nouns, adverbs, adjectives, and pronouns contained in a given clause were then identified. There were seven information sources from which the clause constituents could be derived: current sentence, prior text, relevant world knowledge, irrelevant world knowledge, evaluations, recollections, and metacognitive.

Results and Discussion

Principle components extraction with varimax rotation was conducted on the 7 information sources. The overall Kaiser-Meyer-Olkin measure of sampling adequacy was .66, indicating an adequate factorability of the data. Based on the eigenvalues and scree test results, a three-factor extraction was adopted with a .40 factor loading as the practical significance criterion. The varimax rotated component matrix showed that 7 items loaded on three factors with a range of loadings from .58 to .84. The total amount of variance explained by the three factors was 68%. Factor 1, labeled as *strategy-explaining*, consisted of three items associated with deep-level processing strategies (e.g., prior text and relevant world knowledge). Factor 2, labeled as *irrelevant-explaining*, consisted of two items that are typically not associated with successful comprehension (e.g., irrelevant elaborations and evaluation). Factor 3, labeled as *episodic understanding*, consisted of two items that reflected personal involvement with the task or situations involving concepts mentioned in the text. These data suggest that the information sources that comprise a verbal protocol are most closely associated with the strategy of self-explaining (McNamara, 2004). That is, when self-explaining, readers use information from the current sentence (e.g., paraphrase), prior text (e.g., bridging), and world knowledge (e.g., elaboration). It should be noted, however, that the sample size used in the study was somewhat small ($N = 70$) for conducting factor analysis. Nonetheless, it was sufficient for exploratory purposes. Future analyses would need to focus on using confirmatory factor analysis with a new data set using the three factors that emerged from the exploratory factor analysis. This would enable one to test how well the data fit the model extracted from the exploratory analysis. Additionally, given that skilled and less-skilled readers use different strategies during reading (e.g., Magliano & Millis, 2003), it would be beneficial to test whether different information sources load onto different factors as a function of reading skill.

References

McNamara, D. S. (2004). SERT: Self-Explanation Reading Training. *Discourses Processes*, 38, 1-30.

Magliano, J. P & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction*, 21, 251-284.

Trabasso, T. & Magliano, J.P. (1996). Conscious understanding during comprehension. *Discourse Processes*, 21, 255-287.