

## [Assessing quality of metaphor interpretation by right hemisphere damaged patients](#)

By: Hiram Brownell, [Kristine Lundgren](#), Carol Cayer-Meade, Michelle Nichols, Kathryn Caddick, Jacque Spitzer

Brownell, H., Lundgren, K., Cayer-Meade, C., Nichols, M., Caddick, K., & Spitzer, J. (2007). Assessing Quality of Metaphor Interpretation by Right Hemisphere Damaged Patients. *Brain and Language*, 103(1-2), 197-198. doi: 10.1016/j.bandl.2007.07.113

Made available courtesy of Elsevier: <http://dx.doi.org/10.1016/j.bandl.2007.07.113>

\*\*\*© Elsevier. Reprinted with permission. No further reproduction is authorized without written permission from Elsevier. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document. \*\*\*

**This is the author's version of a work that was accepted for publication in *Brain and Language*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Brain and Language*, Volume 103, Issue 1/2, (2007) DOI: 10.1016/j.bandl.2007.07.113**

### **Abstract:**

Patients with right hemisphere damage (RHD) can display various cognitive-linguistic deficits including disrupted comprehension of figurative language such as metaphor (Myers, 1999 and Tompkins, 1995). Most studies documenting figurative language comprehension difficulties have used multiple choice tasks (Van Lancker & Kempler, 1987). Multiple choice formats are generally easier than open-ended verbal explanation formats and, therefore, may be less sensitive and, accordingly, less adequate for documenting clinically relevant changes in the quality of comprehension over time. However, scoring a patient's extended verbal responses can be subjective. Also, there is uncertainty whether the resulting scores provide only ordinal information or approach interval level measurement, which makes statistical analysis easier. Although scales have been developed for scoring proverb interpretation (Gorham, 1956 and Nippold et al., 1997), we are not aware of a scoring system for representing the quality of metaphor interpretations.

**Keywords:** Cognitive-linguistic deficits | Metaphor | Figurative language

### **Article:**

### **Introduction**

Patients with right hemisphere damage (RHD) can display various cognitive-linguistic deficits including disrupted comprehension of figurative language such as metaphor (Myers, 1999 and Tompkins, 1995). Most studies documenting figurative language comprehension difficulties have used multiple choice tasks (Van Lancker & Kempler, 1987). Multiple choice formats are generally easier than open-ended verbal explanation formats and, therefore, may be less sensitive and, accordingly, less adequate for documenting clinically relevant changes in the quality of comprehension over time. However, scoring a patient's extended verbal responses can be subjective. Also, there is uncertainty whether the resulting scores provide only ordinal information or approach interval level measurement, which makes statistical analysis easier. Although scales have been developed for scoring proverb interpretation (Gorham, 1956 and Nippold et al., 1997), we are not aware of a scoring system for representing the quality of metaphor interpretations.

## Methods

We examined the scale properties of a new assessment tool based on open-ended verbal explanation, the Appropriateness of Metaphor Interpretation Scale (AMIS). The AMIS is a seven category scoring system with the "6" category being the highest quality response, and the "0" category being the lowest. (See Table 1. Category 6 was very rarely used and was excluded from analysis.) As part of a larger study on metaphor training for RHD patients (e.g., Lundgren, Brownell, Cayer-Meade, & Roy, 2006), RHD patients were asked to interpret over 250 metaphors throughout a 15-week period that included baseline, training, and post-training sessions. Responses were audio recorded and transcribed verbatim.

**Table 1.** Appropriateness of metaphor interpretation scale

AMIS category	Interval scale value	Definition of category
6	N/A	Complete and appropriate, characterized by descriptive, rich language
5	2.7	Complete and appropriate (non-concrete) but basic/simple
4	1.9	Complete and appropriate but delayed (longer than 5 seconds required for the initiation of a response). May contain self corrections, false starts but eventually gets to the correct response. May include some tangential comments and/or personalization
3	1.3	Close substitution using appropriate alternative metaphor to close substitution/with some clear elements of abstraction but not the complete, correct/appropriate interpretation

2	1.4	Literal associate, may be partially correct with some partial or inappropriate non-concrete extension
1	0.7	Literal associate without any non-concrete extension
0	0.0	No response, “I do not know” response, off topic comments

The quality of over 50% of the metaphor interpretations was rated by three trained individuals (2 of whom were naïve with respect to the training condition) using the AMIS. Reliability, computed by dividing the number of agreements by the sum of agreements plus disagreements and, then, multiplying by 100, reached 90%.

The primary goal of the present study is scale development. Six judges participated: 3 had no prior exposure to the AMIS, and 3 were highly trained.

Three metaphors (“Fear is a trap”, “The child is a pill”, “Work is a headache”) were examined because they elicited responses representing 1 to 3 instances of each of six AMIS categories. For all three metaphors, each response was paired with every other response from that metaphor, making a total 147 pairs that were presented in a random order. For each pair of responses, a judge decided which was better, without being told or reminded of the original AMIS assignment. Instructions mentioned the following criteria for responses (without weighting): prompt, appropriately abstract, accurate, and complete.

## Results

Analysis examined whether responses, which were classified according to their original AMIS value, were consistently judged “better” than other responses to the same metaphor. Data from all six judges were evaluated for inter-judge consistency. Their responses for all 3 metaphors were then combined in a single analysis. Thurstone’s law of comparative judgment (Case V) was used to determine whether the AMIS categories approximate a true interval scale, (see Duffy & Dale, 1977, for an example; Ghiselli et al., 1981 and Guilford, 1936). Using assumptions related to the binomial approximation to the normal curve, the analysis yields an average standard score value for each original AMIS category. A constant equal to the lowest average standard score is then added to each category’s value to bring the lowest scale value to 0.0. Subsequent analyses will be based on a larger number of naïve judges, as well as more metaphors.

Table 1 provides preliminary interval-level scale equivalents for the AMIS categories. These results, which will be adjusted on the basis of additional data, suggest combining categories 2 and 3, and the usefulness of other categories, except category 6.

## Implications

We have made considerable progress developing a useful scale for metaphor interpretation that, with modification, does a good job approximating interval level measurement and provides a relatively fine-grained assessment tool. Most important, the procedures used are straightforward and can be applied widely to scale development for behavioral assessment.

## References

- Duffy, J.R. & Dale, B.J. (1977). The PICA Scoring Scale: Do its statistical shortcomings cause clinical problems? In *Clinical Aphasiology Conference Proceedings* (pp. 290–296). Minneapolis: BRK Publishers. Clinical Aphasiology Conference (May, 1977, Amelia Island, FL: May 17-20, 1977)
- Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: W.H. Freeman.
- Guilford, J. P. (1936). *Psychometric methods*. New York: McGraw-Hill.
- Gorham, D. R. (1956). A proverb test for clinical and experimental use. *Psychological Reports*, 2, 1–12.
- Lundgren, K., Brownell, H., Cayer-Meade, C., & Roy, S. (November, 2006). Remediation of metaphor comprehension deficits: A pilot study. Poster presentation at the annual convention of the American Speech-Language Hearing Association, Miami, FL
- Myers, P. S. (1999). *Right hemisphere damage: Disorders of communication and cognition*. San Diego, CA: Singular Publishing Group.
- Nippold, M., Uhden, L. D., & Schwarz, I. E. (1997). Proverb explanation through the lifespan: A developmental study of adolescents and adults. *Journal of Speech, Language, and Hearing Research*, 40, 245–253.
- Tompkins, C. A. (1995). *Right hemisphere communication disorders*. San Diego, CA: Singular Publishing Group.
- Van Lancker, D. R., & Kempler, D. (1987). Comprehension of familiar phrases by left but not by right hemisphere damaged patients. *Brain and Language*, 32, 265–277.