CompLEAS: Computerized Scoring of an Emotional Awareness Test

Duncan E. Leaf, Kimberley A. Barchard, Wanwalai Charoenchote, Rhone D'Errico
University of Nevada, Las Vegas

ABSTRACT

The Levels of Emotional Awareness Scale (LEAS) is an open-ended test used in emotional intelligence research. Because scoring the LEAS is a time consuming process for humans, we developed a scoring program called CompLEAS. The essential function of CompLEAS is to search for valued words and phrases within a body of text. After finding the valuable items, the program calculates scores for the entire text using a number of Scoring Methods. A study to examine the effectiveness of CompLEAS Scoring Methods found that their scores were highly correlated with human-generated LEAS scores.

PROBLEM AND MOTIVATION

Open-ended Tests vs. Closed-ended Tests – An open-ended test asks a question or describes a topic to which the test-taker must respond. Examples include short-answer and essay tests. Open-ended tests are used in various kinds of social science research, including sociology, marketing, and psychology. In closed-ended tests, a respondent may choose only from a fixed set of responses (e.g. true/false and multiple-choice tests). Closed-ended responses can easily be assigned a score or value. On the other hand, a single open-ended question can result in unlimited responses. It is more challenging to quantify an open-ended response. This study examined the effectiveness of computerized administration and scoring of an open-ended psychological test called the Levels of Emotional Awareness Scale.

Levels of Emotional Awareness Scale –The Levels of Emotional Awareness Scale (LEAS; Lane, Quinlan, Schwartz, Walker, & Zeitlin, 1990) is a twenty-item, open-ended test designed by Richard Lane (psychiatrist) and Gary Schwartz (psychologist). Unlike other measures of emotional intelligence, scores on this test are based on the structure of the response rather than the specific content. Expert researchers have found the LEAS to be a "reliable, distinctive, and useful measure" in the relatively new and rapidly expanding field of emotional intelligence (Ciarrochi, Caputi, & Mayer, 2003, p. 1489).

Previous research shows that scores on the LEAS are related to neurological events associated with emotion processing. The anterior cingulated cortex (ACC) is a brain structure with functions known to involve "attention and response selection" (Lane, Reiman, Axelrod, Yun, Holmes, & Schwartz, 1998, p.525). Lane et al. found increased blood flow to the ACC during "film- and recall-induced emotion" (Lane et al., 1998, p.525). This implies that induced emotions are associated with increased attention and decisions regarding appropriate responses. Furthermore, subjects' LEAS scores correlated highly with cerebral blood flow during the emotionally evocative tasks. Thus it appears that people who are paying more attention to the emotionally evocative events know more about emotions and conversely people who are more emotionally intelligent are more attuned to emotional stimuli. A different study (Lane, Kivley, Du Bois, Shamasundara & Schwartz 1995) found a strong positive correlation between LEAS

scores and the degree to which the brain's right hemisphere dominates in the perception of facial emotion. This is taken to imply that "as right hemisphere dominance in the perception of facial emotion increases, the ability to perceive complexity during the processing of emotional information increases" (Lane et al., 1998, p. 525). Thus, two different studies have shown that LEAS scores have meaningful relationships with different aspects of neurological function.

Despite the uniqueness and utility of the LEAS, this measure is not widely used because scoring the LEAS is time-consuming. This section will describe the scoring of the LEAS. The LEAS was based on the theory of emotional development advanced by Lane and Schwartz (1987). They describe emotional awareness and expression as a developmental process, consisting of five levels of increasing complexity and abstraction (Lane & Schwartz, 1987). To provide respondents with an adequate opportunity to demonstrate their emotional awareness, Lane and his colleagues created twenty emotionally evocative scenarios involving the self and one other person. For each, the respondent answers the questions "How would you feel?" and "How would the other person feel?" Most LEAS responses are between 10 and 40 words.

Each item is scored using a three-step process (Lane, 1991). First, scorers assign a numerical rating to each emotion word/phrase given. The LEAS manual requires scorers to make *subjective* decisions about context and synonyms when assigning values to words. Often a LEAS scorer will have to decide whether or not two words are synonyms and how the context of a word changes its interpretation. For example, the word "hurt" can describe an emotion or a physical sensation. The value depends on interpretation. Human scorers must consider the context in which "hurt" was used before assigning it a value.

From the scores for each word or phrase, scores for Self and Other responses are derived. Finally, from the Self and Other scores, a Total score is calculated. This process is repeated for all twenty scenarios. Thus, the scorer must read the answer given, sift through the response for the words and phrases to be scored, assign scores to each, and then calculate scores for Self, Other, and Total. This process is typically laborious and time-consuming. In addition, training research assistants to score the LEAS often takes weeks or even months. These two facts likely have been the primary cause for the lack of wider use of this measure. A reliable computerized scoring program for the LEAS would eliminate the need to train expert scorers, would greatly decrease the time and resources necessary to score the LEAS items, and could significantly increase the number of researchers using this valuable measure.

Over the last year, we have developed a program called CompLEAS to solve this problem and we have conducted an empirical study to examine its effectiveness.

BACKGROUND AND RELATED WORK

James Pennebaker and his colleagues were faced with a similar problem in analyzing self-disclosures, and developed a program called Linguistic Inquiry and Word Count (LIWC; Pennebaker, Francis & Booth, 2001) to count the frequency of specified words in a body of text. The LIWC counts the number of words that fall into each of 72 linguistic and psychological categories. The LIWC has been used to examine the relationship between word choice and depression, honesty (Pennebaker & Lee, 2002), health, social integration (Pennebaker & Graybeal, 2001), and prayer (VandeCreek,

Janus, Pennebaker, & Binau, 2002). The primary limitation of the LIWC is that it can only be used to count the number of words that fall into each category. CompLEAS is capable of word-counting but it also goes beyond the functionality of the LIWC by assigning values to words and phrases, and using these values to calculate total scores for each body of text.

Two commercially available software packages to analyze open-ended responses are StatPac's *Verbatim Blaster* and *TextSmart* by SPSS. Both of these packages are similar to the LIWC in that they categorize words. However, they are limited because they do not perform any calculations within and between these categories. Because of this, they cannot be used on their own to produce numerical scores for open-ended tests.

Research in computerized essay scoring has used *k*-nearest neighbor classification to compare a to-be-scored essay with other similar essays (Larkey, 1998). This approach could be useful if a sufficiently large body of already-scored LEAS responses were on record. However, at this time, no such database exists.

APPROACH AND UNIQUENESS

A human LEAS scorer looks at a participant's responses and matches their words to words in the Glossary of the LEAS Scoring Manual (Lane, 1991). CompLEAS attempts to simulate the process employed by human scorers. The glossary used by CompLEAS is based upon – but is not identical to – the Glossary in the LEAS scoring manual. The CompLEAS Glossary consists of about 1,000 words and phrases, each of which is assigned a point value. It is stored in a list of trees to enable fast searching since the first few words of many phrases often overlap. CompLEAS uses this Wordlist to find the valuable words and phrases within a response. The point values of these valuable words and phrases are then used to calculate scores via four different Scoring Methods. The Scoring Methods were designed to approximate the way in which humans assign values to words and create total scores. These methods are described in Table 1.

Sixty-seven participants completed the LEAS in a web form. Each item had two text input fields. One was for responding about Self. The second field was for responding about Other. Responses were hand-scored by research assistants trained in LEAS scoring. These human-generated scores were correlated with scores generated by each of the CompLEAS Scoring Methods.

RESULTS AND CONTRIBUTIONS

This study examined the correlations between LEAS scores generated by the CompLEAS program and those generated by hand-scoring. Four different scoring methods are used within the CompLEAS program. All four methods produced scores that correlated highly with the human-generated scores. See Table 2. While all of the scoring methods had strong correlations, the 3345 method had the highest *(r = .85, p < .01)*.

Table 1
The Four CompLEAS Scoring Methods

AllSum	Total scores are calculated as the sum of all values found in both fields. Self and Other scores are calculated the same way, using the first and second fields respectively, but these scores were not used in this analysis.
Highest-4	Total scores are the sum of the four highest values in both fields. Self and Other scores are calculated the same way, using the first and second fields, respectively, but were not used in this analysis.
334	Total scores are calculated as the highest value in either field, unless there are two different valuables that both have a value of 3: in this case, the score is 4. For example, the words "happy" and "guilty" both have the value 3. If the first box contains "happy" and the second box contains "guilty", the Total score will be 4. However, if a response contains the word "happy" twice, this is not sufficient for a score of 4. The 334 method was designed to mimic the scoring rule that gives a score of 4 when two or more level 3 emotions are present and distinguishable from each other. Self and Other scores are calculated using the same method, but restricting the analysis to those valuables found in the first and second fields, respectively. However, separate Self and Other scores were not used in this analysis.
3345	Total scores are calculated using the Self and Other scores from the 334 method, described above. The Total score for the 3345 method is the maximum of the Self and Other scores for the 334 method, unless both Self and Other are 4: in this case, the Total score is 5. The 3345 method is conceptually the closest to the hand scoring method. The 3345 method uses the same Self and Other scores as the 334 method, but these scores were not used in this analysis.

Table 2
Correlations Between CompLEAS-Generated and Human-Generated LEAS Scores

CompLEAS Method	Correlation
AllSum	.72*
Highest-4	.76*
334	.78*
3345	.85*

* $p < .01$

There are some ways in which the CompLEAS program could be further improved. First, we could incorporate spell-checking into CompLEAS. Currently, spelling errors result in lower total scores. Second, we could expand the glossary, to include emotion words and phrases that occur less frequently. Third, we could incorporate *k*-nearest-neighbors classification and comparison. To do this, we would need to collect a sufficiently large number of already-scored and already-typed LEAS to provide a useful database. If such improvements are made, future revisions of CompLEAS could have even higher correlations with hand-scoring.

The present study has demonstrated that CompLEAS is an efficient and effective method of scoring the LEAS. It could also be used in any other context in which numerical scores need to be assigned to a body of text, based upon scores assigned to the individual words or phrases used. Therefore, it may be useful for scoring many kinds of open-ended tests.

REFERENCES

Ciarrochi, J., Caputi, P., & Mayer, J.D. (2003). The distinctiveness and utility of a measure of trait emotional awareness. *Personality and Individual Differences, 34,* 1477-1490.

Lane, R.D. (1991). *LEAS Scoring Manual and Glossary.* Unpublished manual for the Levels of Emotional Awareness Test. (Available from Richard D. Lane, General Clinical Research Center, University of Arizona, PO Box 245002, Tucson, AZ 85724-5002).

Lane, R.D, Kivley, L.S., Du Bois, M.A., Shamasundara, P., & Schwartz, G.E. (1995). Levels of Emotional Awareness and the degree of Right Hemispheric Dominance in the Perception of Facial Emotion. *Neuropsychologia,* 33, 525-538.

Lane, R.D., Quinlan, D.M., Schwartz, G.E., Walker, P.A., & Zeitlan, S.B. (1990). The Levels of Emotional Awareness Scale: A cognitive-developmental measure of emotion. *Journal of Personality Assessment, 55,* 124-134.

Lane, R.D., Reiman, E. M., Axelrod, B., Yun, L.S., Holmes, A., and Schwartz, G.E. (1998). Neural correlates of Levels of Emotional Awareness: Evidence of an interaction between emotion and attention in the anterior cingulate cortex. *Journal of Cognitive Neuroscience,* 10, 525-535.

Lane, R.D. & Schwartz, G.E. (1987). Levels of Emotional Awareness: A cognitive developmental theory and its application to psychopathology. *American Journal of Psychiatry, 144,* 133-143.

Larkey, L.S. (1998). Automatic essay grading using text categorization techniques. *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 90-95). New York: ACM Press.

Pennebaker, J.W., Francis, M.E., & Booth, R.J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC 2001*. Mahwah, NJ: Erlbaum Publishers.

Pennebaker, J.W., & Graybeal, A. (2001). Patterns of natural language use: Disclosure, personality, and social integration. *Current Directions in Psychological Science, 10,* 90 – 93.

Pennebaker, J.W. & Lee, C.H. (2002). The power of words in social, clinical, and personality psychology. *The Korean Journal of Thinking & Problem Solving, 12,* 35-43.

VandeCreek, L., Janus, M., Pennebaker, J.W., & Binau, B. (2002). Praying about difficult experiences as self-disclosure to God. *The International Journal for the Psychology of Religion, 12,* 29-39.