Aspects of Emic and Etic Measurement: Lessons From Mary Poppins

Christopher J. Lee

An understanding of human individuality is fundamental to the practice of occupational therapy. However, this understanding cannot progress without appropriate methods of measuring the individual ways in which people construe experience. In this article, an illustrative passage from one of the Mary Poppins stories serves to introduce a general discussion of three important aspects of emic and etic methods of measuring human individuality.

In the story Mary Poppins Opens The Door (Travers, 1943), Mary Poppins resumes her position as nanny for the Banks family and sets out to measure how the children have grown during her absence (pp. 21–22). She instructs Michael to stand straight as she extends a measuring tape from his head to his feet. Mary Poppins examines the tape and exclaims, “You’ve grown Worse and Worse.” When Michael protests that “tape measures don’t tell words,” Mary Poppins simply shows him that printed on the tape at his height are the words worse and worse. She measures Jane with the same tape. “Wilful, Lazy, Selfish,” it reads. After obtaining descriptions of the children, Mary Poppins obtains a measurement of herself. It reads, “Practically Perfect,” and she comments, “No more than I expected.”

In this passage, the incongruence evident between the measurement instrument and the character descriptions is a source of humor because we recognize that measurement instruments with the simplicity of a measuring tape do not afford descriptions of the complexity of human individuality. Measurement is a process of relating constructs to empirical indicators, and in the passage, Mary Poppins relates constructs such as “worse and worse” and “practically perfect” to empirical units on the measuring tape.

There are two general approaches to the measurement process, and Mary Poppins’s measuring tape provides an apt counterpoint to discussion of several important aspects of these approaches. An emic approach to the measurement process begins with empirical indicators and aims to discover the constructs that facilitate our understanding of these indicators. In contrast, an etic approach to the
measurement process begins with formal constructs and aims to select empirical indicators to represent these constructs. Mary Poppins may have adopted an emic approach by beginning with the empirical units of the measuring tape and subsequently inferring constructs such as “willful, lazy, selfish” or “practically perfect.” Or, she may have adopted an etic approach by beginning with the intent to measure these constructs and subsequently selecting the empirical units of a measuring tape as her indicator of these constructs. In either approach, the integrity of the measurement process relies on the selection of an appropriate range of indicators, the relevant association of indicators and constructs, and a recognition that this process provides a bounded description of the phenomenon being measured.

Selection of an Appropriate Range of Indicators

In the emic literature, research is seen as involving representation of “the multiple constructions that various respondents have made” (Lincoln & Guba, 1985, p. 212). The indicators of these multiple constructions frequently are words, phrases, or other meaningful units contained in respondents’ narratives. An emic researcher aims to describe “the range of experience rather than the average experience” (Krefting, 1991, p. 216), which is achieved through prolonged engagement. In the emic literature, prolonged engagement refers to spending sufficient time with respondents to become familiar with the variety of ways in which the respondents construe an experience (Gliner, 1994; Krefting, 1991; Lincoln & Guba, 1985).

Prolonged engagement provides an appropriate basis for determining the relationship between the empirical units and individual constructions of an experience. In the case of respondents’ narratives, words are approximate indicators of personal constructions (Kelly, 1955), and prolonged engagement provides a sufficient basis for deciphering the association among the words, phrases, or other lexical units that comprise individual narratives and the individual constructions that each person uses to understand an experience.

Personal constructions evolve as experience is gained, and persons construe events differently as a result of having encountered different experiences (Kelly, 1955). For instance, two persons might prepare the very same meal, encountering identical obstacles, yet construe the task differently by virtue of having experienced different obstacles in preparing meals on previous occasions and having had different life histories.

Contrast is an important component of the meaning of an experience (Kelly, 1955). One person might consider a task to be difficult relative to other tasks that are easy, whereas another person might consider a task to be difficult relative to other tasks that are impossible. Although both persons use the word difficult to describe a task, the meaning of a task in the life of each is quite different. One must know more than the dictionary meaning of a word to understand its personal significance, and prolonged engagement provides an appropriate basis for interpreting the personal meaning of experiences.

The selection of an adequate range of indicators is also important in etic measurement, where the issue of content representativeness (Messick, 1989, 1995) is concerned with incorporating a sufficient number of items, or questions, to represent the various facets of an experience. The questions asked of respondents define the boundaries of their responses. Consequently, a standardized assessment must use a sufficient range of questions to obtain responses that are representative of the variety of ways in which different persons could construe an experience. The selection of an adequate range of indicators enables persons to obtain the same assessment for different reasons. For example, two persons could exhibit the same level of “personal independence” (as represented by the same total score) yet have made different responses to the individual questions that sum to the score on a standardized assessment of personal independence.

In sum, an important component of both emic and etic measurement is an appropriate range of empirical units that serve as indicators of personal constructions of experience. The measurement of a complex construct requires a large selection of indicators, and even the simplest of constructs “can rarely, if ever, be represented by a single item” (Pedhazur & Schmelkin, 1991, p. 63). The integrity of Mary Poppins’s assessments can be questioned on the grounds that a single indicator—a single point on the measuring tape—is a poor representation of the constructs involved in her measurements.

Relevant Association of Indicators and Constructs

Although a legitimate basis for criticizing Mary Poppins’s assessment is its unreasonably narrow selection of indicators; this issue is not the source of Michael’s complaint. Michael protests that “tape measures don’t tell words.” The basis of his complaint is her association of indicators and constructs. The association of indicators and constructs is a matter of establishing relevance. One suspects that a more relevant way of construing Michael’s position on the measuring tape might have been “taller and taller” rather than “worse and worse.” It is unlikely that Michael would have complained that “tape measures don’t tell words” were...
“taller and taller” the words offered as an interpretation.

The issue of relevance is concerned with the evidential basis of an interpretation, and a compelling argument that the available evidence supports a particular interpretation is required to substantiate an investigator’s association of indicators and constructs (Messick, 1989, 1995). One method of addressing this issue is to have colleagues review the evidence supporting an interpretation as part of the measurement process. In an emic approach, the measurement process begins with empirical indicators and develops constructs that facilitate an understanding of these indicators. An apt review of this process is conducted by impartial colleagues for the purpose of exploring alternative ways of construing the indicators, a method termed peer debriefing (Lincoln & Guba, 1985) or peer examination (Krefting, 1991). In an etic approach, the measurement process begins with formal constructs and selects empirical indicators to represent the definition of these constructs. An apt review of this process is conducted by colleagues with expert knowledge of the focal constructs (Messick, 1989, 1995) for the purpose of evaluating alternative selections of indicators.

**Bounded Description**

An experience is open to a large number of alternative interpretations given that each individual construes phenomena in ways that reflect his or her personal history (Kelly, 1955). Consequently, the assumption should be that the respondents’ personal constructions cannot be applied uniformly to other persons. This limitation has been noted in both the emic literature (Lincoln & Guba, 1985) and the etic literature (Messick, 1989, 1995).

The idea that respondents’ personal constructions should not be applied uniformly to other persons underlies the issue of transferability in the emic literature (Lincoln & Guba, 1985) and the issue of generalizability in the etic literature (Messick, 1989, 1995). In the conduct of emic research, the categories and themes that comprise a researcher’s reconstruction of the meaning of an experience are afforded by the researcher’s prolonged engagement with a specific set of individuals. The person wanting to apply the interpretations from emic research is responsible for considering the extent to which other persons resemble the original respondents (Lincoln & Guba, 1985). Similarly, in the design of etic measures, the items that comprise a standardized assessment are developed in reference to a specific sample of individuals. Those administering a standardized assessment are responsible for assuring comparability with an appropriate standardization sample when interpreting respondents’ scores. Thus, in either approach to the measurement process, it is feasible to anticipate the meaning of an experience for a limited range of persons whose constructions resemble those of the reference sample.

Perhaps to check the integrity of her measurement process, Mary Poppins obtains a measurement of herself. As she expected, it read “Practically Perfect.” However, measurement methods have a limited range of convenience. A measurement method that may be apt for Mary Poppins may not be equally apt for Michael and Jane.

**Conclusion**

In providing evidence of the integrity of her interpretation of the empirical units of the measuring tape, Mary Poppins must do more than point to words printed on it. In considering the merits of various methods of measuring human individuality, it is important to establish evidence of an adequate range of indicators and a relevant association of indicators and constructs and to recognize that such evidence places bounds on the interpretation of measurement outcomes. Without satisfying these conditions, measurement in occupational therapy cannot aspire to being client centered.

**References**


