

How many sentences are there in this extract? How many words? How many paragraphs? How many ideas? The difficulty of answering these questions illustrates a number of serious challenges in both qualitative and quantitative content analysis. The challenge of identifying the unit of analysis (or, as it is sometimes termed, unitizing) is critical to reliable content analysis (Rourke, Anderson, Garrison Archer, 2001).

The Sentence as Unit of Analysis

In text transcript analysis, a common method is to use a grammatically defined unit of analysis. Some e-researchers have used the sentence as the unit of analysis (Fahy, Crawford, Ally, Cookson, Keller, Prosser, 1999; Hillman, 1999), however, as the sample email illustrates, online text dialogue often follows its own rules. These rules create a relaxed grammar that falls somewhere between text and voice communication. This often makes the identification of sentences problematic. The sentence as a unit of analysis is also challenging in that the number of sentences can be very large, making for a time-consuming process. [n addition. it may also be difficult to identify a relevant variable in each and every sentence.

The Paragraph as Unit of Analysis

Other e-research content analysts (Hara, Bonk, Angeli, 2000) have chosen the paragraph as the unit of analysis. This unit has an advantage in that it is larger, requiring fewer decisions of the researchers. [n addition it should be "a distinct division of written or printed matter that begins on a new, usually indented, line, consists of one or more sentences, and typically deals with a single thought or topic or quotes one speaker's continuous words" (www.dictionary.com). However, our email example shows that often users do not write in clearly defined paragraphs, leaving an unfortunate amount of interpretation to the e-researcher. As the size of the unit expands. so does the likelihood that the unit will multiple variables. Conversely, one variable may span multiple paragraphs. Our experience does not support Hara, Bonk, and Angeli's optimism that "college-level students should be able to break down the messages into paragraphs" (p. 9). Further, once the syntactical criteria are lost, the definition of the unit as a paragraph becomes meaningless, and what the coders are identifying are, in fact, arbitrary blocks of text. Hara, Bonk. and Angeli's ad hoe coding protocol reveals these problems: "when two continuous paragraphs dealt with the same ideas, they were each counted as a separate unit. And when one paragraph contained two ideas. it was counted as two separate units" (p, 9), Thus, the selection Of the paragraph presents a very

problematic unit of analysis.

The Message as Unit of Analysis

The full email or computer conferencing message has also been the unit of analysis (Ahern, Peck, Laycock, 1992; Marttunen, 1997). This unit has important advantages. First, it is objectively identifiable. Unlike other units of analysis, multiple raters can agree perfectly on the total number of cases. Second, it produces a manageable set of cases.