# Analyzing Collaboration in Online Communications

**Albert L. Ingram**
*Kent State University, USA*

**Lesley G. Hathorn**
*Kent State University, USA*

## INTRODUCTION

Elsewhere (Hathorn & Ingram, 2002b; Ingram & Hathorn, 2004; Ingram & Hathorn, 2005), the authors have argued that collaboration in a small group can be seen to consist of three major elements: interdependence, synthesis of information, and independence. These three factors can be found in the transcripts of online discussions by looking at such things as roughly equal participation among the group members, the interactions that occur, and the product that the group constructs. Here we demonstrate how these concepts can be applied to a specific set of online discussions to determine the amount of collaboration that has taken place in different groups. Specifically, we present a coding scheme that can be used to analyze online collaborative discussions. This scheme has potential utility for education and research.

## BACKGROUND

We can collect transcripts of messages from a threaded Web discussion and use content analysis to determine the extent of collaboration in the discussion (Henri, 1992; Mason, 1992). Content analysis (Silverman, 1993) involves identifying categories of statements and counting the number of items in the text that appear in the categories. Analysis schemes used for text, such as newspaper articles, letters, or written statements, may be inappropriate for Computer-Mediated Communications (CMC), even though it is also text-based communication (Henri, 1992). Communication patterns in CMC lie between spoken conversation and written discourse because the discussion follows neither a logical nor a predictable pattern (Mason, 1992). New coding schemes specific to online collaborative discussions are needed. Where a purely qualitative study might refer to examples of the discussion to justify an argument, content analysis allows all the data to be analyzed as a whole. The analysis is more credible because there is less reliance on the researchers' subjective impressions (Silverman, 1993).

Rourke, Anderson, Garrison, and Archer (2001) discussed the need to have a coding scheme designed to examine the variables of interest. Their paper, however, provided only an abstract theoretical model of content analysis, not a practical coding scheme. Most current research studies on collaboration or CMC still rely heavily on participation as the behavior that indicates the construct of interest, for example, demographics (McLean & Morrison, 2000), satisfaction (Ochoa & Gottschall, 2004) or extending the classroom learning situation (Nicholson & Bond, 2003).

In contrast, we argue that researchers interested in collaboration must use a scheme specific to that construct. There are three drawbacks of many current coding schemes for collaboration. First, the coding scheme may not be designed to measure collaboration but some other construct such as social presence (Rourke, Anderson, Garrison, & Archer, 1999). Second, the measurement model may be based on questionnaires that measure perceived degree of communication or interactivity; for example, Richardson and Swan (2003) use this approach. This method allows for quick and easy data collection, but it does not measure the actual discussion. It only measures opinions and attitudes toward the discussion and the outcomes. Using content analysis avoids the biases that occur when relying on self-report measures.

A third problem is that some coding schemes are based on face-to-face (F2F) collaboration rather than online collaboration. Models that compare F2F and

CMC (for example, Bennett & Dunne, 1991; Hawkes & Romiszowski, 2001) have shortcomings in analyzing collaboration in CMC because the coding schemes are created for F2F collaboration, meaning that online collaboration is measured in terms of the qualities of F2F interaction it possesses or lacks, and not as a distinct process in its own right.

Bennett and Dunne (1991) developed a model to study F2F interaction in groups. The categories they proposed for direct discussion of a topic were not entirely appropriate for CMC. Models for F2F interaction emphasize the social and nonverbal aspects of communication. In F2F collaborative groups, members may not even be aware of how their contributions compare to those of other members of the group. F2F conversation is often in fragments with incomplete sentences and with one idea being completed by different people (Wild & Braid, 1997). Coding CMC conversations is less vulnerable to the vagaries of interpretation. Verbal social communication can be coded separately outside the issue of collaboration.

Henri (1992) proposed a model for studying collaborative learning in a CMC environment, but it was of limited use to us because it emphasized learning rather than collaboration. Instead, we adapted Henri's definition of independent statements to mean the presentation of information without leading to further discussion. We added categories that distinguished between statements that simply agree with prior statements and those that add information. We distinguish between interaction that attempts to take the discussion further and responses that indicate little more than participation. This distinction becomes important in explaining why groups that appear to be discussing topics in a highly interactive manner may actually produce few new ideas.

## Coding Procedure

Each message is divided into statements. A statement is a complete sentence or a complete idea within a sentence. Complex sentences often contain more than one statement. In CMC discussion, one message can reply to one or many messages, as well as discuss various other topics that may or may not be related. Dividing messages into statements allows one to identify the ideas discussed as well as the true structure of the discussion threads. The statements are coded according to the specific characteristics that determine collaboration:

- **Interdependence:** Requires that each member actively contribute to the discussion. First, it requires roughly equal participation by the group members, measured by counting the number of messages and statements submitted by each discussant. Once equal participation has been established, individuals' contributions to problem solution are measured by the interaction in the group on the problem. Interdependence is also measured by comments about the problem and by other comments that create a social context.
- **Synthesis:** Requires that new information be created, which can be measured in two ways. First, there is the interaction pattern that occurs when a participant contributes a statement, another extends the idea, and a subsequent message synthesizes the information. A synthesizing thread requires at least three messages from two group members. We also assess synthesis by examining the relationship between the original comments and the final product. Does the final product meld contributions from individual group members, does it consist primarily of the work of one discussant, and can we identify individual work in the product?
- **Independence:** The ability of the group to work without the instructor. It is measured by analyzing the instructor's influence on participation and interaction. A discussion in which few threads occur without appeals from the group to the instructor for help and advice is not independent.

Tables 1 and 2 show detailed schematics of the coding categories, adapted from Bennett and Dunne (1991), Henri (1992), and Mason (1992). Statements are coded three times, once for each of the three characteristics: participation, interaction, and patterns of discussion. In the first step, participation is coded. In the next, coding for interaction is refined. The last step involves coding the patterns of discussion. Table 1 shows the coding of participation and then interaction. For further coding, we selected only on-task direct discussion of the scenario. Table 2