

# HOLISTIC AND DISCOURSE FEATURE MEASUREMENT IN WRITING RESEARCH

Speaker: Edward M. White, *California State University, San Bernardino*  
Introducer/Recorder: Rosemary Patton, *San Francisco State University*

Edward White opened his presentation with remarks about the purpose and scope of his project on Research in the Effective Teaching of Writing and the tentative nature of the conclusions at this date. He invited questions and throughout the session received a variety of provocative inquiries and challenges which led to his elaboration of the testing model and discussions of potentially controversial or problematic features of the research. He then introduced Charles Moore, CSU Sacramento, a member of the research team, who participated with White in the presentation. Working with White and Moore on the research team are Kim Flachman, CSU Bakersfield; William Stryker, CSU Northridge; David Rankin, CSU Dominguez Hills; and Linda G. Polin, Pepperdine University.

In December, 1980, the National Institute of Education funded White's proposal to analyze postsecondary writing programs in an attempt to discover if decisions made above the individual faculty level had any effects on either faculty or student outcomes. White claimed that, so far, the answer seems to be "yes". If such writing program organization does make a difference, what are the differences? What type of writing program structure is effective? This part of the research, which White and his colleagues are currently writing up, covers student outcomes only.

Before discussing the project and his conclusions, White distributed a summary of this phase of their research, including the methods used, the essay topic, three scoring guides for the readers, frequency distribution tables for the three different scoring criteria, three charts showing correlations between each of the criteria, and two sample student essays with the scores each received. This phase of the research was designed to evaluate one set of essays and to establish criteria for judging the writers' competency.

Using a sampling device, the researchers selected a sample of 3,422 45-minute essays and established their criteria based on what freshman composition courses aim to teach:

1. simple description,
2. movement from simple description (the concrete) to abstraction,
3. analysis, evaluation, comparison,
4. standard surface structures of college prose.

They were careful to explore these criteria from many perspectives, considering cognitive, heuristic, invention, and development concerns. Next, these writing samples were scored three times, using 1) a holistic scale, 2) a discourse-feature scale measuring "development and focus," and 3) a second discourse-feature scale measuring "correctness and efficiency." For each scale, White described the elaborate scoring guides with their criteria for each grade—the 6-point scale now familiar for holistic readings across the nation.

Criteria for *development* were derived from Francis Christensen's model of paragraph structure (a limita-

tion later questioned during discussion). *Focus* was defined as the features of the connective tissue—transitional markers, pronouns, audience awareness in the better writers. *Correctness* referred to usage and mechanics, and efficiency reflected rhetorical characteristics of sentences—predication, modification, variety of structures—and diction.

Graduate students were trained to mark specific *discourse blocks* (DB), one per essay. They were instructed to find a proposition (a "contract statement") which led to the most successfully developed passage in the paper, and to mark where this particular block began and ended (usually about 100-150 words). On the assumption that this would be the best block for both skills categories, these were the passages used to determine scores for both development and focus and correctness and efficiency. The isolated passages were presented deliberately to avoid possible second or third holistic readings. When asked about the reliability of these "markers," White commented that in the future they would use experienced composition instructors to select the blocks because the graduate students were not consistently reliable. Someone asked, "If the evaluators chose the 'best' block in each paper, wouldn't the scale scores naturally be higher?" White pointed out that they wanted to find out how successfully a student could develop a topic and thus intended to "reward students for what they do well."

White acknowledged the complexity of the holistic scoring guide and thus the problem of establishing what would constitute a gain on a second test when there were so many variables for each score. He was then asked, "since the holistic scores tended toward the mean (a finding illustrated on the chart), how could discriminations be made with such "bunching" near the middle?" White did not answer this question, but he explained the value of including the two discourse-feature categories to supplement the traditional holistic approach. He also pointed out that student writers often start out weak but get stronger in both coherence and surface features. The carefully selected passages could reflect this characteristic of student prose.

One finding surprised the researchers. They assumed that development and focus were high level skills, but discovered that the scores actually shifted up on this scale, and provided the top average scores in all categories.

White commented that these data could have implications for teaching basic writing. It is generally agreed that correctness and efficiency are taught more consistently than any other features of writing, yet his data suggest that freshman writers demonstrate greater proficiency in development and focus. Even low-level students demonstrated some ability in these areas. Thus, since this study suggests correctness may be the last feature to develop, one might assume that instruction in correctness could come later rather than early in the instructional sequence.

White concluded that, on the basis of a 45-minute writing sample after eight weeks, it might be possible to arrive at some judgments on what kinds of criteria could be useful for arranging college composition programs. He emphasized the tentative nature of all their conclusions and said that he and his team are still evaluating the data. NIE will receive their study in June and it will be published in the fall.