DOCUMENT RESUME

ED 362 163 IR 016 319

AUTHOR Dempsey, John V.; Driscoll, Marcy P.

TITLE Error and Feedback: The Relationship between Content Analysis and Confidence of Response.

PUB DATE Jan 93

NOTE 12p.; In: Proceedings of Selected Research and Development Presentations at the Convention of the Association for Educational Communications and Technology Sponsored by the Research and Theory Division (15th, New Orleans, Louisiana, January 13-17, 1993); see IR 016 300.

PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Biology; Computer Assisted Instruction; *Confidence Testing; *Discrimination Learning; *Error Correction; *Feedback; Higher Education; Programed Instruction; *Time on Task; Undergraduate Students

ABSTRACT

This study examined the relationship between discrimination error (determined by content analysis and tryout data) and confidence of response (determined by self report). Subjects were 63 undergraduate students enrolled in a biology class for nonmajors who received classroom expository information and read a text on the topic before they completed a computer-based instructional module. Before subjects received any feedback on their responses to the module, they were queried about their confidence of response. Feedback was provided only to incorrect responses. The results indicated that students spent more feedback study time (i.e., elapsed time from when response-contingent feedback was first presented on the display screen until the learner pressed the appropriate key to view the next item) and required more question-based examples in studying content involving rules than concepts. As expected, students spend much more time studying feedback after fine discrimination errors than gross errors. Surprisingly, confidence of response was inversely correlated with feedback study time, as well as fine discrimination error and gross error. The negative relationship between fine discrimination errors and confidence of response could be explained by inconsistencies with the learners' self reports of their confidence of response and the relationship between high confidence errors and effort. (Contains 21 references.) (KRN)

Title:

Error and Feedback: The Relationship Between Content Analysis and Confidence of Response

Authors:

John V. Dempsey
Marcy P. Driscoll

Abstract

Our prior studies using science-related concepts and rules have indicated that learners spend twice as much time studying feedback after fine discrimination errors than they do after gross errors. Likewise, studies by Kulhavy and his associates suggest that learners expend longer feedback study times after errors for which they had a high confidence of response. The purpose of the present study was to see if there were a relationship between discrimination error (determined by content analysis and tryout data) and confidence of response (determined by self-report). Results indicated that, as in prior studies, the relationship between fine discrimination error and feedback study time was positive. The relationship between fine discrimination error and confidence of response, however, was negative. Possible explanations for these results are discussed.

Error and Feedback: The Relationship Between Content Analysis and Confidence of Response

In his classic review, Feedback and Written Instruction, Kulhavy (1977) proposes a model of learner expectancy. Kulhavy's model expresses the relationship between correct or wrong answer post-response feedback given to a learner and her self-reported confidence of response in making that reply in the first place. High confidence error feedback (wrong answer feedback for erroneous responses the learner had expected to be correct) are predicted to yield longer feedback study times than either low confidence error feedback or feedback after correct responses.

Kulhavy's model adheres to the first of Ammons' eight empirical generalizations which states, "The learner usually has a hypothesis about what he is to do and how he is to do it, and these interact with knowledge of performance" (1956, p. 281). Based on pilot data, Kulhavy, Yekovich, and Dyer theorized that learners create a hierarchy of confidence in the the correctness of their responses. Under these conditions, learners' reactions to error range from surprize when confidence is high to acceptance when confidence is low (1976, p. 522). These observations were validated by several experimental studies (Kulhavy et al, 1976, 1979; Kulhavy, White, Topp, Chan, & Adams, 1985; Lhyle & Kulhavy, 1987) and further, endorse the "common sense criteria" so often ignored in experimental research involving human learning.

From a differing perspective, we felt that an area of concern in the work of Kulhavy and his associates was the reliance on learner's self-reports of the confidence of their responses. In the studies conducted by Kulhavy and his associates and the present study, which emulated their procedure, learners stopped after each response and rated their confidence in each response. Although the notion of a learner-constructed response hierarchy made intuitive sense to us, we wondered how often learners accurately portray the response hierarchy with which the question was actually answered. Clearly, we felt, more sophisticated learners with greater strategic learning ability would have an advantage over those with less ability. Likewise, older learners would have advantages in understanding their response hierarchies over very young learners (Kulhavy, Stock, Hancock, Swindell, & Hammrich, 1990).

Additionally, during instruction, self-report measures are distracting. In essence, learners are asked two questions, one content-related and one not. Unless, as is possible, the self-report measures were used as part of a game format (see, for example, Scarth & Litchfield), their use in instructional situations would be impractical.

Work involving the use of rational sets of concepts and simple rules (Driscoll & Tessmer, 1985; Klausmeier & Feldman, 1975; Markle & Tieman, 1970) have established a method by which errors of fine and gross discrimination, two ends of the error continuum, may be predicted. This method was adapted to the computer (Dempsey, 1986; Driscoll and Dempsey, 1987) and refined (Dempsey, Driscoll, & Litchfield, in press) by comparing the predictions of fine and gross discrimination errors (i.e., content analysis) with actual on-task observation of errors made during instruction. Our prior experiments using rational sets of concepts have indicated that learners make a higher number of incorrect answers that are fine discrimination errors than gross discrimination errors. Consistent with this research, we expected learners to make more fine discrimination errors than gross errors in the present study.

An important indicator of how engaged learners are in the instruction is the amount of time they spend studying textual feedback given after incorrect responses. Kulhavy and his associates (Kulhavy, Yehovich, & Dyer, 1976, 1979; Kulhavy, White, Topp, Chan, & Adams, 1985) have asserted also that students' expectancy for success is related to the amount of time students spend studying feedback. Likewise, based on prior studies with science concepts (e.g., Litchfield, Driscoll, & Dempsey, 1990), we expected that students would spend more time studying fine discrimination errors (errors associated with concepts which had similar attributes to a correctly classified concept). A student making a fine discrimination error, we posited, would have a high expectancy for success. Fine discrimination errors were, after all, close-in nonexamples of correct concepts. Making incorrect responses that were "almost" correct should serve to increase attention and stimulate curiosity. Under these conditions, feedback study time would be extended.

Gross discrimination errors, on the other hand, were far-out nonexamples and suggest that learners have failed to comprehend the material. We expected that learners will spend less time studying corrective feedback for gross discrimination errors. Failing to understand a concept or rule, learners guess quickly and move on to areas they better understand. It may be supposed that a learner who makes a gross discrimination error has little expectancy for success in classifying that particular concept.

Naturally, because our work and that of Kulhavy and associates made predictions based on assumptions of learners' expectancy for success, we speculated that these approaches were linked in some way. The purpose of the present study, therefore, was to see if there were a relationship between discrimination error (determined by content analysis and tryout data) and confidence of response (determined by self-report).

Method

Subjects and Procedure.

The subjects in this study were 63 mostly freshman and sophomore university students enrolled in a biology class for nonmajors. The class, which fulfilled a basic studies requirement for undergraduates, had a traditionally high enrollment and unsatisfactory pass/fail ratio. Subjects comprised three laboratory classes chosen by the undergraduate Biology coordinator to participate in a pilot program which incorporated the use of adjunct computer-based instruction (CBI).

Prior to completing the computer-based instructional module, students read a 12-page chapter on the topic of substance abuse from a required text produced by the Biology Department. Two hours of classroom time were also devoted to expository information of the module topic. Students received credit for completing the CBI module at their convenience during a 10-day period. To complete the CBI module, students located an unoccupied computer terminal at one of several public access locations on campus and "sign-on" to the system. After typing in their names and social security numbers, subjects were given all additional instructions by the computer program.

Materials and Instruments.

The content of the instruction were selected rational sets of concepts and rules related to a newly-introduced, state-mandated substance abuse module. The rational sets of interest in this study were types of drugs (defined concepts), the effects of drugs on the nervous system (rules), and alcohol use and abuse (rules) and included 44 exemplars. An instructional design strategy, the rational set generator, was applied in the design and development of the instruction. The rational set generator is a matrix model that incorporates multiple examples of concepts and rules and provides for discrimination and generalization learning. Discrimination here refers to the ability to make distinctions between examples and nonexamples of concepts and rules. The interrogatory examples used in this study required that subjects classify particular concepts or rules after reading narrative anecdotes containing varying degrees of concept or rule attributes. The CBI rational set generator used an adaptive strategy which branched subjects to more difficult examples after correct classification and easier examples after incorrectly classifying or applying concepts or rules. Items answered correctly were discarded from the program.

Fine and gross discrimination errors were diagnosed using a two-step approach. First, content experts predicted the relative likelihood of making a discrimination error for each nonexample distractor by considering the content relationships among concepts or rules in a rational set. Distractors representing closely related nonexamples, for example, were more difficult to discriminate than less closely related nonexamples and would represent fine discrimination errors. Thus, nonexamples were rank ordered by their "rational" content relationships. Second, before analysis this predictive relationship was compared to actual student responses and, where necessary, items were adjusted to reflect discrimination error trends.

Before subjects received any feedback, they were queried about their confidence of response in a similar manner to that proposed by Kulhavy et al (1979). A five point scale (1= lowest confidence, 5=highest confident) composed of touch boxes, and a question asking the student how sure she was about her answer appeared at the bottom of the computer screen immediately after a content response was made. After indicating confidence of response, subjects received content response-contingent feedback. Simple confirmation was provided for correct answers. After incorrect responses, subjects were informed of the correct concept or rule in a standard feedback box which remained on the screen along with the interrogatory example until students chose to touch the screen or press the keyboard to continue on to the next example.

In the present study feedback study time was collected after incorrect responses only. Feedback study time was defined as the elapsed time from the moment when response-contingent feedback was first presented on the computer display screen until the learner pressed the appropriate key to view the next item.

Results

The results of the study indicated that, as may be expected, students spent more feedback study time and required more question-based examples in studying content involving rules than concepts. Otherwise, as Table 1 indicates, the patterns were quite similar across the three learning outcomes used in this study, i.e., drugs, coordinate (or rationally-related) defined concepts; drugs, coordinate rules; and alcohol, successive (or nonrelated) rules.

Insert Table 1 about here

Feedback study time was directly correlated with fine discrimination errors ($r = .456$) as shown in Table 2. As expected, students spent much more time studying feedback after fine discrimination errors than gross errors.

Suprisingly, confidence of response was inversely correlated with feedback study time ($r = -.469$) as well as fine discrimination error ($r = -.466$) and gross error ($r = -.479$).

Insert Table 2 about here

Discussion

Although these findings are far from conclusive, two possible explanations could explain the negative relationship between fine discrimination errors and confidence of response. These are: (1) inconsistencies with the learners' self reports of their confidence of response, and (2) the relationship between high confidence errors and effort.

In the Kulhavy studies (as well as the present study) learners stopped after each response and rated their confidence of response. One wonders how often learners accurately portray the response hierarchy with which the question was answered. We would suppose, for example, that there would be a great difference in the reliability of self-reported confidence measures among sophisticated learners versus those with less ability -- or older versus younger learners.

An initial investigation by Swindell, Greenway, and Peterson (1992) upholds our suppositions. In a study with 4th and 6th grade students, these researchers found that 6th graders were more reliable in estimating response confidence than were 4th graders. They also found that the response patterns of the 6th graders were similar to those of college students (Kulhavy, Stock, Hancock, Swindell, & Hammrich, 1990), but response patterns of the 4th grade students were distinctly different.

Other researchers have called into question the use of self-reported confidence measures. For example, Koriat, Lichtenstein, & Fischoff (1980) have found that people are often overconfident in evaluating the correctness of their knowledge. Their research supports the notion that learner's assessment of confidence is biased by attempts to justify one's chosen answer. In discussing self- reports, Borg and Gall (1983) observed, "people often bias the information they offer about themselves, and sometimes they cannot accurately recall events and aspects of their behavior in which the researcher is interested" (p. 465).

In addition, self report measures during instruction are distracting. Essentially, learners are asked two questions, one content related and one not. Thus, the practical value of self-report as an instructional or motivational design measurement tool is reduced.

Regarding our second speculation, the findings of this study, considered in respect to the existing text-based feedback literature, indicate a more complex relationship between the type of error made (determined via content analysis), expectancy (measured by confidence of response scales), and the amount of effort a learner makes (as measured by feedback study time) than had been suspected. This relationship is illustrated in Figure 1. In addition to other factors, we suspect that confidence of response measures are greatly influenced by specific learning outcomes, the difficulty of material to be learned, the learner's prior knowledge, and the relevance of the material to the learner.

Insert Figure 1 about here

One practical implication of the present study is for researchers to explore more sophisticated systematic explanations for the use of corrective feedback in interactive instruction. While the tendency is to look for simpler clarifications such as those proposed by Kulhavy (1977), the evidence of this and certain other studies suggest the relationship among error, expectancy, and feedback is a complex one. More recent work among several researchers (Dempsey, Driscoll, & Swindell, in press; Kulhavy & Stock, 1989; and Bangert-Drowns, Kulik, Kulik, & Morgan, 1991) have begun to address these concerns at least within the limited area of text-based feedback. What is needed are disciplined explorations of these and other models.

References

Ammons, R. B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. Journal of General Psychology, 54, 279-299.

Bangert-Drowns, R.L., Kulik, C.C., Kulik, J.A., and Morgan, M.T. (1991) The instructional effect of feedback in test-like events. Review of Educational Research, 61(2), 213-238.

Borg, W.R., & Gall, M.D. (1983). Educational research (4th ed.). New York: Longman.

Dempsey, J. (1986). Using the rational set generator with computer-based instruction for creating concept examples: A template for instructors. Educational Technology, 26 (4), 43-46.

Dempsey, J.V., Driscoll, M.P., & Litchfield, B.C. (in press). Feedback, retention, discrimination error, and feedback study time. Journal of Research on Computing in Education.

Dempsey, J. V., Driscoll, M. P. & Swindell, L. K. (in press) Text-Based Feedback. In J. V. Dempsey & G.C. Sales (Eds.). Interactive Instruction and Feedback. Englewood Cliffs, NJ: Educational Technology Publications.

Driscoll, M.P., & Dempsey, J.V. (1987, January). Developing and testing a method for enhancing concept learning in computer-based instruction. Paper presented at the annual meeting of the Association for Educational Communications and Technology. Atlanta, GA.

Driscoll, M.P., & Tessmer, M. (1985). The rational set generator: A method for creating concept examples for teaching and testing. Educational Technology, 25 (2), 29-32.

Klausmier, H.J., & Feldman, K.V. (1975). Effects of a definition and a varying number of examples and nonexamples on concept attainment. Journal of Educational Psychology, 67, 174-178.

Koriat, A., Lichtenstein, S. , & Fischhoff, B. (1980). Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory, 6, 107-118.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. Educational Psychology Review, 1, 279-308.

Kulhavy, R. W., Stock, W. A., Hancock, T. E., Swindell, L. K., & Hammrich, P. L. (1990). Written feedback: Response certitude and durability. Contemporary Educational Psychology, 15, 319-332.

Kulhavy, R.W. (1977). Feedback in written instruction. Review of Educational Research, 47, 211-232.

Kulhavy, R.W., White, M.T., Topp, B.W., Chan, A.L., & Adams, J. (1985). Feedback complexity and corrective efficiency. Contemporary Educational Psychology, 10, 285-291.

Kulhavy, R. W., Yekovich, F.R., & Dyer, J.W. (1976). Feedback and response confidence. Journal of Educational Psychology, 68, 522-528.

Kulhavy, R. W., Yekovich, F.R., & Dyer, J.W. (1979). Feedback and content review in programmed instruction. Contemporary Educational Psychology, 4, 91-98.

Lhyle, K.G., & Kulhavy, R. W. (1987). Feedback processing and error correction. Journal of Educational Psychology, 79, 320-322.

Litchfield, B.C., Driscoll, M.D., & Dempsey, J.V. (1990). Presentation sequence and example difficulty: Their effect on concept and rule learning in computer-based instruction. Journal of Computer-Based Instruction, 17 (1), 35-40.

Markle, S.M., & Tiemann, P.W. (1971). Conceptual learning and instructional design. Journal of Educational Technology, 1 (1), 27-33.

Scarth, G., & Litchfield, B.C. (1990). Ergomotion. (Level III Interactive Videodisc). Boston, MA: Houghton-Mifflin.

Swindell, L.K., Greenway, R., & Peterson, S.E. (1992). Children's use of response confidence in the processing of instructional feedback. Manuscript submitted for publication.

Table 1

Descriptive Statistics for number correct, attempts, feedback study time, fine & gross discrimination errors and overall confidence of response (n = 63).

VARIABLE*	MEAN	SD
B-correct	7.08	.83
C- correct	6.20	.99
D- correct	12.02	2.21
B-attempts	11.79	4.95
C-attempts	15.05	3.95
D-attempts	20.12	2.59
B-FB study time	131.46	211.72
C-FB study time	176.30	209.19
D-FB study time	170.73	178.37
B-fine errors	.73	.61
C-fine errors	2.11	1.47
D-fine errors	1.89	1.23
B-gross errors	.56	.98
C-gross errors	.76	1.10
D-gross errors	.65	1.01
Confidence of Response (all 3 matrices)	3.807	.396

Note: 44 items -- from three instructional matrices:
B matrix = drugs (16 items, defined concepts)
C matrix = drug rules (12 items, simple rules)
D matrix = alcohol (16 items, simple rules)

Table 2

Intercorrelations among the variables of feedback study time, confidence of response, fine discrimination errors, and gross discrimination errors (n = 63).

Variable	FB Study Time	Conf of Response	Fine Errors	Gross Errors
FB Study Time	1.00			
Conf of Response	-.469*	1.00		
Fine Errors	.456*	-.466*	1.00	
Gross Errors	.204	-.479*	.055	1.00

* $p < 0.001$

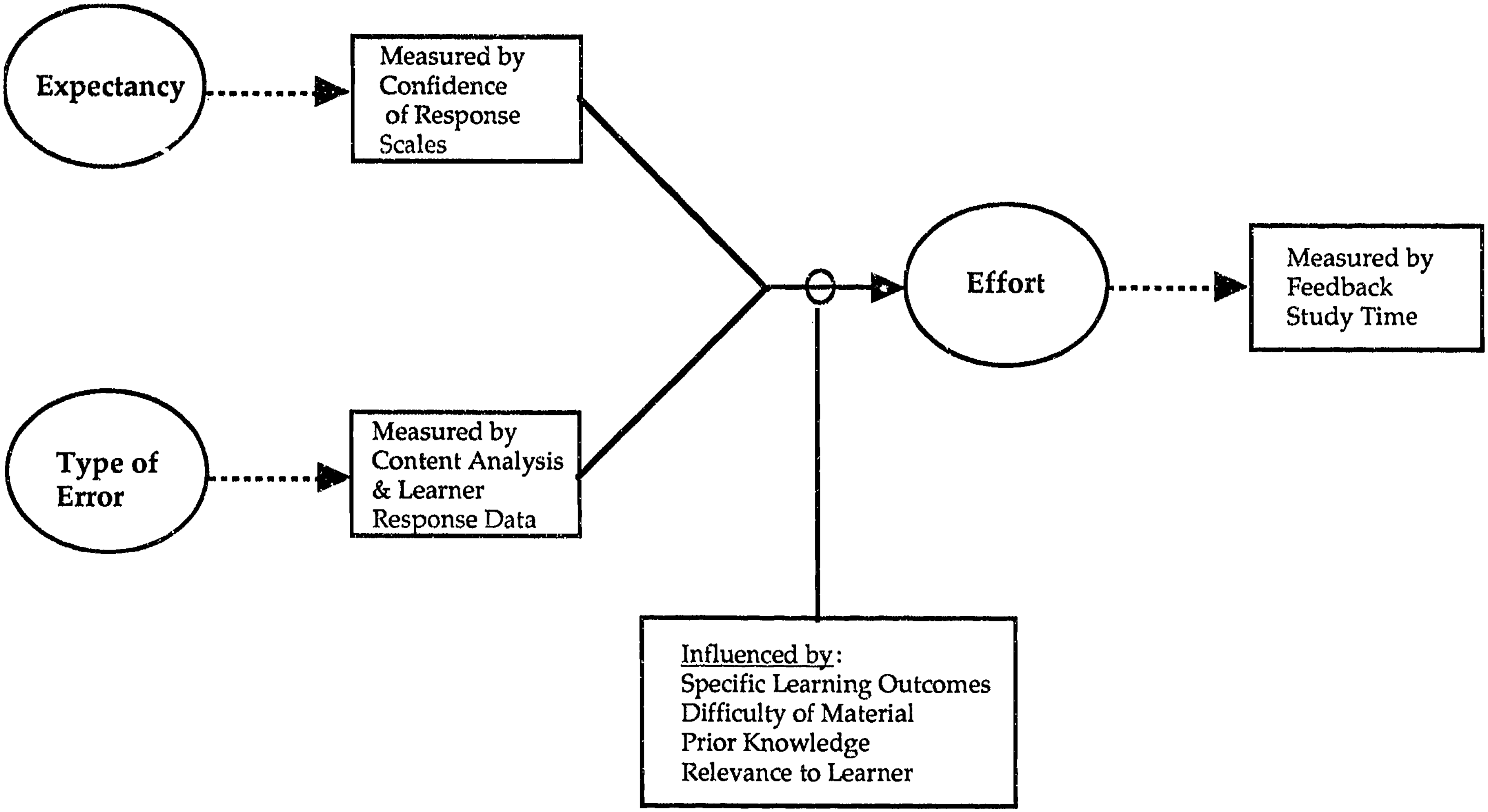

Figure 1. Possible relationship among expectancy, type of error, and effort.

Note. From J. Dempsey, M. Driscoll, & L. Swindell (in press) Text-Based Feedback. In J. V. Dempsey & G.C. Sales (Eds.). Interactive Instruction and Feedback. Englewood Cliffs, NJ: Educational Technology Publications.