DOCUMENT RESUME

ED 129 918 TM 005 792

AUTHOR McMorris, Robert F.; Leonard, Gregory
TITLE Item Response Changes and Cognitive Style.
PUB DATE Apr 76
NOTE 16p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (22nd, San Francisco, California, April 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Academic Achievement; *Cognitive Style; College Students; *Multiple Choice Tests; *Response Style (Tests); Scores; *Test Wiseness
IDENTIFIERS *Answer Changing (Tests)

ABSTRACT

According to conventional wisdom, a test taker should not change his/her first response to a multiple-choice, although empirical evidence has consistently supported such changes. Quizzes for masters level students in educational measurement and evaluation showed increments due to answer changing. Low anxious students tended to make more changes and to gain from the changes than did high anxious students. Relationships between test-taking behavior and other cognitive style, attitudinal and academic performance variables are also discussed. (Author/MV)

Item Response Changes and Cognitive Style

Robert F. McMorris & Gregory Leonard

SUNY at Albany

Should a test taker change his initial response to a multiple choice item? Are changed responses more likely to be correct than are initial responses? According to conventional wisdom, the first response is the best response. For example, Jacobs (1972) asked students to summarize their answer-changing experience. A plurality felt they had lost points, with about as many students having no opinion as estimating a gain. Also, a class of graduate students taught by one of the authors strongly advised against answer changing. Lynch and Smith (1975) quoted several published sources who advise against changing answers, and also surveyed students with similar results: over 75% of the students estimated that changing answers would tend to lower their scores.

In contrast, empirical evidence has consistently shown increments due to changing initial responses (e.g., Lynch & Smith, 1975; Matthews, 1929; Mehrens & Lehman, 1973, p. 317; Reiling & Taylor, 1972; Smith & Moore, 1976). Further, several researchers compared answer-changing behavior among subgroups of their total samples. For example, subgroups were defined according to total test score, course grades, GPA, or a standardized cognitive test. Generally the better students tended to gain more than did the poorer students (Archer & Pippert, 1962; Bath, 1967; Copeland, 1972; Mueller &

A paper presented at the convention of the National Council on Measurement in Education, San Francisco, April 1976.

Our appreciation to Dr. James Powers, Miss Leslie McMorris, Miss Lucy Weiner, and Mr. Robert Pfeiffer, who assisted in obtaining the sample, scoring the instruments, and analyzing the data.

Shwedel, 1975; Reile & Briggs, 1952), although no differences were found by Jacobs (1972) and Pascale (1974). Several of these researchers also investigated sex differences with very mixed results. Yet even though some groups gained more than others, the average for each of these subgroups was an increment.

Based on empirical evidence, then, the test taker should expect a gain in score if he reconsiders an item, concludes that his response should be changed, and then makes the change. Such a discrepancy in advice vs. evidence suggests a need for further investigation using a different approach. Perhaps if subgroups are defined quite differently, answer-changing behavior would differ. "It is possible, for instance, that certain personality variables such as impulsivity or anxiety . . . may correlate with the incidence and effectiveness of answer-changing behavior" (Mueller & Shwedel, 1975).

The following questions guided this investigation: 1) To what extent do test takers change answers? 2) Does changing answers tend to result in higher scores? Defining answer-changing behavior as number of changes and increment due to changing, 3) Does answer-changing behavior relate to cognitive styles and to attitudes? 4) Is answer-changing behavior related to academic performance?

Sample

Students from three universities enrolled in an educational measurement and evaluation course constitute the sample. The four classes numbered 43, 24, 17, and 50 students. Two of the classes were taught by one professor, two by another professor. The majority of the students were inservice or preservice teachers. Prior to all data being collected, the professors did not mention the appropriateness of changing answers, nor the purpose of the study.

Instrumentation

The dependent variables are the number of changed item responses and the increment in score from changing responses on classroom tests in the educational measurement and evaluation course. For two of the classes the quizzes totalled 54 and 75 multiple choice items, most of which had four alternatives. For the other two classes the quizzes totalled 132 and 135 items, most of which were multiple choice with a small percentage of true-false items. The data were collected unobtrusively, with no special directions given to the respondents. Two judges examined each answer sheet and agreed on the changes made by the respondent. When an item response was changed, the changes were recorded in one of the following categories: 1) wrong response changed to right response, 2) wrong response changed to another wrong response, and 3) right response changed to wrong response.

The independent variables include measures of four cognitive styles, plus self-report grades, attitudes towards the course and towards tests, course performance, and sex. The cognitive styles and the major measure used for each are:

impulsivity-reflectivity, measured by Barratt's Personal Evaluation (Barratt);

anxiety, measured by Spielberger, Gorsuch, and Lusheme's Self-Evaluation Questionnaire (1968);

preference for complexity, measured by Barron's (1963) Complexity Scale; and

field-independence, measured by the Hidden Figures Test (Educational Testing Service, 1962).

Several variables were assessed using a three-page questionnaire which

contained a variety of item types. Brief questions solicited both undergraduate and graduate GPAs and a report of previous course work in measurement. Likert items allowed specification of attitudes toward the course and toward the quizzes. A semantic differential item specified attitudes toward testing using both the evaluation and potency dimensions. Several rating scales were constructed to measure cognitive styles.

Course performance was assessed both as total score on the quizzes and as letter grade for the course. Because of the different numbers of items for each class, quiz score, number of changes, and increment from changing were converted to standard scores within class for computing total group correlations.

Results

All students changed at least one answer, with the mean number of changes ranging from 5.0 to 8.6 depending on the class (see Table 1). Although one person lost three points from answer changing, no one else's loss exceeded one point. Most people gained from changing: the percentage of students gaining ranged from 71% to 91% for each of the four classes, and the mean and median increment tended to be approximately three points for each class.

Performance on the cognitive style variables is described in Table 2. Inspection of this table suggests that Group 3's performance often differed from the others; however, it was the smallest sample.

The best predictor of gain from changing answers was the number of items on which answers were changed. As shown in Table 3, the correlation coefficients ranged from .76 to .80. (Specifically, the coefficient for Group 1 was .80, for Group 2 it was .78, Group 3 and Group 4 were each .76, and the total group was .78.)

The relationships of cognitive style measures with answer changing behaviors were less consistent than was the relationship between the two ways of

describing answer changing behaviors. For the first group those who changed and profited from changing tended to be low anxious, for Group 2 impulsive, for Group 3 anxious, impulsive, and orderly, for Group 4 non-compulsive, and for the total group impulsive. (While the coefficients used in support of the above statement tend to be just above or just below the value needed for significance at the .05 level, the extent of inconsistency within cells of the first two rows of Table 3, plus the number of near zero correlations suggests caution in inferring relationships between answer changing behavior and cognitive style constructs.)

Several cognitive style variables showed greater consistency with each other, however. For example, anxiety measured using the Spielberger instrument correlated appropriately with questionnaire items, e.g., positively with anxiety and negatively with happiness and extroversion. The field independents tended to consider themselves low anxious and happy. Those high in impulsivity tended to be self-rated as compulsive, extroverted, and casual. Frequently for these relationships within the cognitive style variables, as well as for other relationships, the inconsistent group was the third (and smallest) one.

Attitudes toward tests generally, toward the course, and toward the quizzes in the course tended to be minimally related to answer changing behavior. Semantic differential ratings of tests grouped under evaluation and potency dimensions correlated between -.16 and -.18 with answer changing behavior for the total group.

While females in the fourth group changed more answers and profited more from changes (r_s = .24 and .31), the correlations for the total group were only .06 and .14.

Academic performance in general and for this course in particular was also ineffective in estimating answer changing behavior, as may be seen in

Table 4. The general academic performance variables are based on self report; the extent of bias due to self report is likely trivial (see, e.g., McMorris & Ambrosino, 1973).

Discussion

The consistency of results among studies of answer changing remains unblemished: test takers change answers, and when they do, the tendency is to gain rather than either to lose or to remain at the same score.

The search for cognitive style, attitude, and academic performance variables to correlate with answer-changing behavior was attempted. Most relationships were quite low, especially when all four classes were combined. Judging from these data, impulsivity might be the most likely of the present cognitive style and attitude variables to show relationships with number of changes and increment from changing in a replication. As noted earlier, several researchers have found that academic performance is positively related to answer-changing behavior, although our no difference findings are not unique.

One reason for low relationships with answer changing is the limited reliability of the answer changing measures. For the fourth (and largest) class we intercorrelated the number of changes between pairs of the three quizzes, averaged using the z transformation, and corrected the reconverted average using the Spearman-Brown formula. We repeated the process for the increment. The reliability estimates were .76 and .21, respectively.

Why do test givers caution test takers not to change answers? One reason may be that changes from right to wrong are better remembered than are changes from wrong to right. Students going over a corrected test are more likely to be wondering why they got certain items wrong and not why they got other items right. They are thus more likely to notice right to

wrong changes and be convinced that changing responses tends to hurt their test scores (Lynch & Smith, 1975, p. 223). Perhaps we humans expect a gain and can't tolerate a loss, and recall the losses more vividly.

Another reason for the caution may have to do with how test givers conceptualize memory. For example, memory may consist of ideas or concepts that represent copies of sensory experiences. Remembering or recall may require the arousal of a sequence of neural connections that have stored these copies of "reality." Utilizing this theoretical perspective, an associative network is presumed to be responsible for relating stimuli and responses. In test taking situations, the tester assumes that the first presentation of the stimulus (the multiple choice question) will most likely evoke the response (correct answer option); later evaluation and reconsideration of the question might disrupt the original memory trace. Thus, the test taker is discouraged from altering his initial response.

A review of pertinent research findings indicates that changing answers generally augments the overall performance of the test taker. An explanation for this consistent empirical finding might be to consider memory as an active process that synthesizes stored information to create a new construction. Jenkins (1974) suggests that this theoretical position, which he refers to as contextualism, is more useful in conceptualizing human memory processes than is the traditional associationistic model. This theoretical perspective suggests that remembering is an active process of elaboration that takes time. The reevaluation of questions posed during an examination session could lead to an effective reconstruction of the concepts required in answering these questions. The additional experience gained through exposure to related topics during the examination session should provide information that will augment the development of pertinent concepts

to be used during the review of answer options. The test taker will use this additional information to choose the best option available--changing answers, based on a constructive memory process, would be expected to augment overall performance.

Such a view of memory is supported by Lynch and Smith (1975, p. 223).

> The authors hypothesize that students taking exams have their memories jogged by other items or other memories so that reconsideration after having answered other items makes students more able to reason the correct answer. Students going over a corrected test are more likely to be wondering why they got certain items wrong and not why they got other items right. They are thus more likely to notice right to wrong changes and be convinced that changing responses tends to hurt their test scores. Thus, many students do not change responses.

The consistency of the research findings suggests that countless numbers of test takers may be penalized by inappropriate advice. Yet, the performance of test takers suggests that this advice is either ignored or conditionally interpreted, i.e., individuals do change answers on multiple choice tests. But should our advice be to change answers? (Advice to change is typically given in the discussion or conclusion section of studies on answer-changing behavior.) Why should we question the application of such a consistent research finding? Two concerns of ours are specified below. The unobtrusive nature of the data collection procedure used here provides a naturalistic approach but does not permit unambiguous interpretation of a student's rationale for changing responses. Using a contrived experimental procedure, Jacobs (1972) has demonstrated results similar to ours, but even with his additional controls the reasons for the changes are unclear. For example, what percentage of changes are clerical corrections? What percentage are due to clues from other

items? What percentage are due to additional thinking about the item?

What are the effects of attempting to modify answer changing strategies? Instructions could be varied along a change--do-not-change continuum. This approach could be useful in evaluating the effects of a modified strategy in terms of both number of changes and increment in test performance. If test takers are encouraged to change a higher percentage of responses, would the increment from answer changing documented in the research literature remain an increment?

References

Archer, N., & Pippert, R. Don't change the answer! The Clearing House, 1962, 37, 39-41.

Barratt, E. S. Personal evaluation. Personal communication, April 1975.

Barron, F. Discovering the creative personality. In The behavioral sciences and education. Princeton, N.J.: College Entrance Examination Board, 1963.

Bath, J. A. Answer-changing behavior on objective examinations. Journal of Educational Research, 1967, 61, 105-107.

Copeland, D. A. Should chemistry students change answers on multiple-choice tests? Journal of Chemical Education, 1972, 49, 258.

Educational Testing Service. Hidden figures test. Princeton, N.J.: Author, 1962.

Jacobs, S. S. Answer changing on objective tests: Some implications for test validity. Educational and Psychological Measurement, 1972, 32, 1039-1044.

Jenkins, J. J. Remember that old theory of memory? Well, forget it! American Psychologist, 1974, 29, 785-795.

Lynch, D. O., & Smith, B. C. Item response changes: Effects on test scores. Measurement and Evaluation in Guidance, 1975, 7, 220-224.

Matthews, C. O. Erroneous first impressions on objective tests. Journal of Educational Psychology, 1929, 20, 280-286.

McMorris, R. F., & Ambrosino, R. J. Self-report predictors: A reminder. Journal of Educational Measurement, 1973, 10, 13-17.

Mehrens, W. A., & Lehmann, I. J. Measurement and evaluation in education and psychology. New York: Holt, Rhinehart, and Winston, 1973.

Mueller, D. J., & Shwedel, A. Some correlates of net gain resultant from answer changing on objective test items. Journal of Educational Measurement, 1975, 12, 251-254.

Pascale, P. J. Changing initial answers on multiple-choice achievement tests. Measurement and Evaluation in Guidance, 1974, 6, 236-238.

Reile, P. J., & Briggs, L. J. Should students change their initial answers on objective-type tests?: More evidence regarding an old problem. Journal of Educational Psychology, 1952, 43, 110-115.

Reiling, E., & Taylor, R. A new approach to the problem of changing initial responses to multiple choice questions. Journal of Educational Measurement, 1972, 9, 67-70.

Smith, A., & Moore, J. C. The effects of changing answers on scores of non-test-sophisticated examinees. Measurement and Evaluation in Guidance, 1976, 8, 252-254.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. Self-evaluation questionnaire. Palo Alto: Consulting Psychologists Press, 1968.

Table 1

Number of Answer Changes and Increment from Changing for Each of Four Groups

	Number of Answer Changes				Increment from Changing			
Statistic	1	2	3	4	1	2	3	4
Mean	5.0	8.6	5.6	6.2	3.0	2.8	1.9	3.4
Median	5	6.5	5	5	3	3	2	3
Mode(s)	1,6	8	-	3	3	3	2	5
Standard Dev.	2.8	7.1	3.2	5.1	2.1	2.6	1.9	2.6
Low-High	1-12	1-35	1-11	1-27	(-1)-8	(-1)-9	(-1)-5	(-3)-10
Percent of n Gaining					91	79	71	90
Percent of n Losing					2	12	12	6
Number of Items	74	135	54	132				
Number of Testees	43	24	17	50				

Table 2

Means and Standard Deviations for Cognitive Style Variables

	Means					Standard Deviations				
	Group					Group				
Variable[a]	1	2	3	4	Total	1	2	3	4	Total
Anxiety:										
State	37.7	34.6	36.6	36.1	36.4	9.5	8.2	6.1	11.4	9.7
Trait	38.8	36.9	42.8	36.7	38.0	8.5	8.5	9.5	9.6	9.1
Total	76.4	71.5	79.3	72.8	74.4	15.2	15.1	12.4	19.6	16.8
Anxious-Calm	3.4	3.7	4.1	3.4	3.5	1.6	1.6	1.7	1.8	1.7
Happy-Unhappy	5.5	5.7	5.0	5.3	5.4	1.1	1.0	.9	1.5	1.2
Field Independence	9.8	8.7	5.7	10.1	9.3	4.8	5.0	3.9	6.1	5.4
Preference for Complexity	23.4	23.2	21.1	23.3	23.1	5.9	5.2	4.6	6.2	5.7
Impulsivity	106.4	109.3	104.5	107.1	107.1	12.4	16.2	11.8	12.4	13.1
Compulsive-Noncompulsive	3.9	4.0	4.7	4.0	4.0	1.6	1.8	1.3	1.6	1.6
Extrovert-Introvert	4.5	4.7	4.1	4.6	4.5	1.5	1.7	1.4	1.5	1.5
Orderly-Casual	4.2	3.9	5.0	4.5	4.3	1.7	1.5	1.7	1.6	1.6
n	43	24	17	50	134					

[a]The bipolar adjectives were included on the questionnaire. Maximum agreement with the word on the left would yield a score of seven.

Correlations among Answer Changing Behaviors and Cognitive Style Variables
Four Groups and Total (Decimals Omitted)

Variable	Inc.	St.	Tr.	Tot.	A-C	H-U	F.I.	P.C.	Imp.	C-N	E-I	O-C
Number of Changes	80 78 76 76 78	-18 -14 43 19 04	-34 -13 27 23 00	-30 -15 42 22 02	-26 -05 55 22 06	16 21 -16 -21 -04	15 -23 -05 -07 -02	18 02 -19 -09 01	03 42 59 02 17	-20 13 -47 -22 -09	11 25 -06 -10 04	-05 -30 80 08 04
Increments from Changing		-26 -01 38 -07 -08	-44 -10 05 03 -12	-40 -06 22 -02 -11	-29 12 27 12 02	40 24 00 -20 03	06 -15 01 -05 -03	06 19 -19 07 06	-06 33 60 05 14	-25 21 35 -34 -14	23 10 -13 -19 00	-02 -36 67 16 06
Anxiety: State			42 64 22 74 59	86 90 66 94 90	48 64 13 48 47	-40 -50 -54 -37 -40	-29 -17 -55 -21 -22	-05 35 -41 15 09	14 -09 -11 04 02	-09 26 41 02 06	-10 -36 -31 -37 -28	-19 -27 48 -12 -11
Trait				82 91 88 92 88	36 66 60 60 54	-51 -68 -69 -60 -57	-08 -29 -41 -24 -24	-02 50 -15 12 10	16 -23 24 09 03	20 -12 -03 08 08	-25 -66 -01 -34 -36	-37 -09 51 -05 -07
Total					48 72 52 58 56	-51 -65 -79 -51 -54	-23 -26 -59 -24 -26	-04 47 -32 15 11	17 -18 13 06 03	06 08 18 05 07	-20 -56 -16 -38 -36	-32 -20 63 -10 -10
Anxious-Calm						-59 -67 -51 -52 -55	-25 -27 31 -21 -24	27 29 -08 21 19	-05 -25 06 10 01	02 16 26 18 15	03 -64 11 -24 -22	10 -17 44 13 10
Happy-Unhappy							28 34 65 25 30	-28 01 53 -21 -14	15 48 -09 05 15	02 19 -33 10 05	38 73 00 43 43	-06 -14 -31 04 -07
Field Independence								-25 -07 34 -13 -10	-01 -11 15 -11 -05	19 -21 14 16 05	16 36 26 10 19	-14 27 01 07 01
Preference for Complexity									11 09 -42 03 04	-04 33 -12 -05 01	04 -42 -22 -37 -23	-13 -18 -07 -20 -17
Impulsivity										47 33 -11 34 33	67 46 54 32 49	-55 -63 35 -25 -37
Compulsive-Noncompulsive											50 08 -23 35 27	-16 -49 49 -02 -10
Extrovert-Introvert												-20 -13 -07 -08 -14
Orderly-Casual												

Note. Coefficients for the four classes and the total group are given in order; ns = 43, 24, 17, 50 and 134, respectively. The bipolar adjectives were included on the questionnaire.

Table 4

Correlations of Answer Changing Behaviors with Academic Performance Four Groups and Total (Decimals Omitted)

Academic Performance	Answer Changing	
	Number of Changes	Increment
In General:[a]		
Undergraduate GPA	-06 -09 -36 -01 -07	-01 -03 -38 07 -01
Previous Testing Course (yes-no)	-20 -08 44 09 02	-12 -05 35 -09 -04
Number of Graduate Courses	27 -12 30 00 05	43 -07 30 15 17
Graduate GPA	13 02 05 20 11	13 13 -10 26 14
This Course:		
Total Quiz Score	21 -07 -09 -08 -02	16 03 05 08 08
Letter Grade	11 -27 -21 -08 -11	05 -12 -26 02 -03

Note. Coefficients for the four classes and the total group are given in order; ns = 43, 24, 17, 50 and 134, respectively.

[a]The "In General" variables were included on the questionnaire.