

Short Communication

Test format effects, self concept and anxiety changing behaviour

E. R. I. Afolabi

Department of Educational Foundations and Counselling, Obafemi Awolowo University, Ile-Ife, Nigeria. E- mail: pojegade@yahoo.co.uk

Accepted 28 May, 2021

The study examined the effects of item format, self -concept and anxiety on response changing behaviour. Four hundred undergraduate students who offered a counseling psychology course in a Nigerian university participated in the study. Students' answers in multiple – choice and true – false formats of an achievement test were observed for response changes. Results indicated that more changes were made in true-false than in multiple-choice test items, and students having moderate trait anxiety made significantly more changes than those having low or high trait anxiety. Academic and general self concept was not found to have significant influence on response changing behaviour.

Key words: Test format, Self-concept, Trait anxiety, State anxiety, Response changing, test scores.

INTRODUCTION

Response changing behaviour is concerned with a testee's changing of his or her mind on an option he or she had already chosen as key (that is, the correct or best answer) in which case he or she thereafter deletes or cancels the initial response or answer and substitutes this with another response or alternative. This phenomenon is usually observed by the sighting of an erased initial response, a neat or rough cancelling of the response, or by the testee's indication of a changed response on an answer format formally provided by the tester.

Studies on response changing behaviour have been concerned mainly with the profitability or otherwise of students' changing their initial responses. Prominent among these studies are those of Stoffer et al. (1977), Schwarz et al. (1991), Prinsell et al. (1994) and Kruger et al. (2005) The general consensus among these authors and researchers is that students do profit by changing their initial responses, and that the gains made are not restricted to a few individuals but occur in the majority of respondents who changed answers. Also, there is consistency in the finding that the more changes that are made, the larger the gains. This finding is particularly true when the ratio of the number of response changes made to the number of items in the test is found.

Personality characteristics have been suggested as factors influencing the extent to which students change responses in objective tests. These characteristics include "willingness to take a stand" and impulsiveness (Kruger et al., 2005; Nunally, 1959). Mueller and Schwedel (1975) related anxiety to response changes and found

that the high anxious groups made the most number of changes and had the most gains. Morris and Leonard (1976) followed this result up and found contradictory relationships. In the latter study, it was found that those who profited from changes were low anxious, while in another group in the same study, those profiting from changes were high anxious. The subjects in each group ranged in number from 17 to 40. Prinsell et al. (1994) obtained similar results. In another study, Green (1981) used graduate students and a 25-item multiple-choice test in Basic Statistics and found results that tended to agree with those of Mueller and Schwedel (1975). Ratios of correct-to-total changes were however not significantly different among anxiety groups.

This study was an attempt to compare the relative vulnerability of test item format and some selected personality variables on response changing behaviour. This would aid our decision on what type of test format to employ that will have the most resistance to test contaminants and further provide empirical evidence to assist in determining where emphasis should be placed in reducing errors on final test scores. The following hypotheses were therefore postulated: i) Test item format will have no significant effects on response changing behaviour of students. ii) Trait and state anxiety will not have any significant effects on response changing behaviour. iii) General self concept and academic self concept will have no significant effects on response changing behaviour iv) The effects of item format on response changing behaviour of students will not be dependent on

Table 1. Mean and standard deviation of item response changes by self concept and anxiety groups.

	Multiple – Choice Items									True – False Items								
	Low			Moderate			High			Low			Moderate			High		
	N	X	S.D	N	X	S.D	N	X	S.D	N	X	S.D	N	X	S.D	N	x	S.D
GSC	30	2.73	1.01	66	2.11	1.09	44	2.07	1.26	26	3.50	1.90	92	4.28	2.70	50	4.26	2.08
ASC	40	2.37	1.48	58	2.22	1.26	42	2.09	1.32	49	3.84	1.84	64	4.08	2.60	58	4.50	2.35
TA	84	1.62	1.02	26	3.61	1.75	30	2.17	0.97	92	3.99	2.14	38	5.79	2.46	38	2.92	2.55
SA	86	2.10	1.38	30	2.33	1.21	24	2.54	1.36	85	3.82	2.19	55	3.87	2.08	28	5.71	2.77

Table 2. Two-Way ANOVA on response changes by Test Item Format (TIF) and GSC.

Source of variation	Df	SS	MS	F	P
Test item Format	1	282.64	282.64	67.29	P < 0.05
General Self Concept	2	5.81	2.91	0.69	P > 0.05
TIF x GSC	2	15.9	7.95	1.89	P > 0.05
Within Group	302	269.36	4.20		

the students' level of anxiety and self concept.

METHODS

Four hundred part four (final year) students of Obafemi Awolowo University offering a Counselling Psychology course were used as subjects. It was a purposive sample. The study made use of a 40-item four-alternative Multiple - Choice Achievement Test (MAT) developed by the researcher following general technical considerations of test construction and validation. The content validity of the MAT was ascertained while the criterion-referenced validity (using GPA as criterion) of the MAT yielded an index of 0.72. Its difficulty indices (P) range from 0.061 to 0.990 with mean difficulty index of 0.485, $n = 40$, $S = 0.275$. The indices of discrimination range from 0.04 to 0.64 with a mean value of 0.310, $n = 40$, $S = 0.157$. The Kuder-Richardson internal consistency reliability was 0.72 and the test-retest reliability (over four weeks) was 0.67. The MAT was carefully transformed into a true-false (TF) format, both versions having the same content. In this way, the TF equivalent of the MAT had 160 items and an internal consistency reliability of 0.79, while its test-retest reliability was 0.71. Its mean P was 0.540 and the mean discrimination index was 0.234, with a modal value of 0.33 and a standard deviation of 0.166.

The Tennessee Self Concept Scale (TSCS) (Fitts, 1964) was used to measure General Self Concept (GSC). Vacchiano and Strauss (1968), Gable et al. (1971), Melanie et al (1978), and Dibuj-Ojerinde (1984) have generally supported both the construct and the content validity of TSCS. Yarworth and Carthier (1978) and Congdon (1968) have also obtained reliability coefficients of 0.91 and 0.88 for it. The Academic Self Concept Scale (ASCS) adapted by Boyinbode (1984) from Shaw and Wright (1967) was used to measure students' academic self concept. Its test-retest reliability was 0.76. In the case of anxiety, general anxiety was measured by the Taylor Manifest Anxiety Scale (MAS) developed by Taylor (1955), while state anxiety was measured by Sarason Test Anxiety Inventory (TAI) developed by Sarason (1972). Vabasa (1974) has demonstrated the usefulness of both MAS and TAI for African students, and both have been widely used in Nigeria in reputable studies (Adeola, 1984; Morakinyo, 1984; Adedipe, 1984).

The personality instruments were first administered on the subjects, while the two test item formats, MAT and TF, were administered randomly on the subjects thereafter. The tests were not speeded and did not contain instructions against guessing. The number-right-score was employed in scoring the MAT and the TAF while the TSCS and the ASCS were scored using the conventional Likert procedure. Each test paper was examined for evidence of response changes as indicated by erasures or when ink was used, the blotting or crossing out of responses. Responses were counted as changed only if the original response had been crossed out or erased and a new response made. Data were analyzed using Analysis of Variance (ANOVA). The results are as presented in Table 1.

One hundred and forty or 70% of the 200 subjects who took the MC test changed their initial responses while the corresponding number of response changes from the 200 subjects who took the TF test was 168 or 84%. When item response changes were counted on the answer sheets of subjects, there were a total of 312 response changes made on the MC items. The item response changes ranged from one to nine per subject with a mean number of response changes of 2.22 and a standard deviation of 1.43. In TF items, response changes per subject ranged from 1 to 16 with a mean response change of 4.19 and a standard deviation of 2.44. Altogether, 698 response changes were made on the TF items.

In both MC and TF items, the tendency of subjects to change their responses decreased with their level of state anxiety (see Table 1). Generally, students in moderate and low anxious groups tended to change responses more than those in the high anxious group. The result of ANOVA analysis on response changes is presented in Tables 2 and 3.

From the two tables, only the TIF effect was significant. $F(1,306) = 67.29$, $P < 0.05$; and $F(1,302) = 64.19$, $P < 0.05$. The main effects of GSC, ASC, and their interaction effects with TIF (that is, TIF x GSC, and TIF x ASC) were not significant. That is, while the type of test format that students took had significant effect on response changing behaviour, neither the difference in their general or academic perception of self nor the interaction between test format and self concept had significant effect on response changing behaviour. The results of ANOVA on response changes by anxiety are indicated in Tables 4 and 5.

Table 3. Two-Way ANOVA on response changes by (TIF) and ASC.

Source of variation	df	SS	MS	F	P
Test item Format	1	271.55	271.55	64.19	P < 0.05
Academic Self Concept	2	12.6	6.30	1.49	P > 0.05
Interaction	2	6.85	3.42	0.81	P > 0.05
Within Group	302	1277.81	4.23		

Table 4. Effects of test item format and trait anxiety on response changes.

Source of variation	Df	SS	MS	F	P
Test Item Format	1	296.42	296.42	56.35	P < 0.05
Trait Anxiety	2	31.77	15.88	3.02	P < 0.05
Interaction	2	10.58	5.29	1.11	P > 0.05
Within Group	302	1589.28	5.29		

Table 5. Effects of test item format and state anxiety on response changes.

Source of variation	Df	SS	MS	F	P
Test item Format	1	278.09	228.09	65.28	P < 0.05
State Anxiety	2	2.40	1.20	0.28	P > 0.05
Interaction	2	4.92	2.46	0.58	P > 0.05
Within Group	302	1281.88	4.26		

Results in Tables 4 and 5 showed that there were significant TIF effects on response changes: (F (1,302) = 56.35, P<0.05 for trait anxiety, and F (1,302) = 65.28, P<0.05 for state anxiety. The trait anxiety effect was significant F (2,302) = 3.02, P<0.05). While the state anxiety effect and the interaction effects of test format with both trait and state anxiety were not significant.

Based on all the results presented above, the hypotheses postulated for this study can be resolved as follows: i) The test item format of an achievement test have a significant effect on response changing behaviour. ii) True-False items have significantly more impact on response changing behaviour than multiple test items. iii) The trait anxiety of a testee has significant effects on his response changing behaviour but the state anxiety does not. iv) Neither the general self concept nor the academic self concept of a testee has significant effect on his response changing behaviour. v) The effect of item format on a testees response changing behaviour is independent of his/her self-concept or anxiety.

DISCUSSION

The results of the series of ANOVA on the data for the study clearly show that test item format is a significant factor affecting response changing behaviour. A comparison of the mean response changes obviously suggest that more items were changed in the TF (X = 4.19) than in the MC test items (X = 2.23). Also, 70% of the MC test takers changed their initial responses, the corresponding figure for TF items was 84%. That TF items are more often changed than MC items might not be unconnected with the nature of the two item formats. The TF item, being conventionally a proposition whose truth or falsity is to

be indicated by the testee, is likely to be more vulnerable to hasty decision-making than the MC items. Infact, Gronlund (1985) is of the opinion that TF items are more susceptible to guessing than MC items. This author is also of the opinion that cheating at examinations might be related to the tendency of students to change their initial answers in TF than in MC items. Further, since in this study, the transformation of a single MC item into four TF items would necessarily involve similarities in the resultant item propositions, students are likely to have cause to change their responses in the items previously encountered if later items are discovered to be related and are understood.

It was only the TA groups that differed significantly in the response changes made among the personality groups under consideration. In the MC items, the mean response changes for the low, moderate, and high TA groups were 1.82, 3.61, and 2.17 respectively, while corresponding mean values in the TF items were 3.99, 5.79, and 2.92. Here, then, the moderate TA group in both MC and TF test items made significantly more changes than the low or high TA groups. While the low TA group made less change than the high TA group in MC items, more changes were made by the low TA than the high TA group in the TF items.

Gilmer (1978) has suggested that practically all students are test anxious but that when this anxiety interacts with academic or scholastic ability, it is mostly the anxiety of the moderate ability group that becomes debilitating.

Galassi et al. (1981) and Samson and Stroops (1979) have also found that high anxious students display a greater tendency of negative thoughts in evaluative situations than their low anxious peers. Part of this negative pre-occupation, they opined, often include doubts on the correctness or appropriateness of their item responses, and consequent possible failure. In the same vein, Boyinbode (1988) argued that moderately test anxious subjects are likely to be really more adversely affected by worry than their low or high test-anxious counterparts. This, according to the author, is because high test-anxious students most often belong to the high ability group with strong needs to achieve. But their anxiety, like those of the low test-anxious, low ability groups, hardly brings about impairment to their academic performance.

Conclusion

The tendency of students to change their initial answers in objective tests occurs among majority of test-takers, and most response changers often profit from the behaviour. The reason why students change their answers has been shown to be related to the students' sex, ability and the level of difficulty of the test items. The present study explored the influence of test format, anxiety and self concept on response changing behaviour. The findings of the study tend to suggest that: (i) subjects tend to change their initial responses in true- false more than in multiple-choice tests; (ii) there is no significant difference in the frequency of response changes made in true-false or multiple- choice tests by subjects who are low, moderate or high in self concept, academic self concept, and state anxiety; and (iii) subjects who are moderate in trait or general anxiety tend to make more response changes than those who are low or high in trait anxiety;

Since response changing is capable of distorting a student's true knowledge and ability, teachers should take cognisance of the item format employed in testing, and the need to create a conducive psychological climate amongst students both before and during testing. The desired end is to control as much as possible for the factors that could contaminate test score.

REFERENCES

- Adedipe VO (1984). Differential Effectiveness of Model Reinforcement and Desensitization Techniques in Improving Academic Performance. Unpublished Doctoral Dissertation. University of Ibadan.
- Adeola AO (1984). Comparative Effectiveness of Relaxation and Systematic Desensitization in the Management of Speech Anxiety. Unpublished Doctoral Dissertation. University of Ibadan.
- Boyinbode RI (1984). Effects of discrepancies between teachers' grades and students expectation on students attitude. Unpublished Masters Thesis. University of Ife.
- Boyinbode RI (1988). Personality correlates of item response changing behaviour in two Objective Test Formats. Unpublished Doctoral Dissertation, International Centre for Educational Evaluation, University of Ibadan.
- Congdon CS (1968). Self Theory and Chorpromazine Treatment. Unpublished Doctoral Dissertation. Vanderbilt University.
- Dibu-Ojerinde OO (1984). Factor Analytic Study of Tennessee Self Concept Scale. Unpublished Masters Thesis, University of Ife.
- Fitts WH (1964). The Tennessee Self Concept Scale, Nashville, Tennessee; Counsellor Recording and Tests.
- Gable RK Lassale AJ and Cooke, KE (1971). Dimensionality of self-perception. Tennessee Self Concept Scale. Perc. Motor Skills, 32: 840-846.
- Kruger J, Wirtz D, Miller DT (2005). Counter Factual thinking and the first instinct fallacy. J. Per. Soc. Psy. 88(5): 725
- Lynch DU, Smith BC (1972). To change or not to change item response when taking tests: Emirical evidence for test takers. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois, April.
- McMorris RF, DeMers LP, Schwarz (1987). Attitudes, behaviours, and reasons for changing responses following answer-changing instruction. J. Ed. Meas. 24(2): 131-143.
- Melenis M, William BH, and Myrod HD (1978). The factorial validity of three frequently employed self-report measures of self concept. J. Ed. Psy. Meas. 38: 547-563.
- Morakinyo A. (1984). Relative Efficacy of Systematic Desensitization, Self-Statement Monitoring, and Flooding on Subjects' Test Anxiety. Unpublished Doctoral Dissertation, University of Ibadan.
- Mueller, DJ and Schwedel A. (1975). Some correlates of net gain resultant from answer changing in objective achievement test items. J. Ed. Meas. 12: 252-254.
- Mueller, DJ and Wasser V. (1977). Implication of changing answers on objective test items. J. Edu. Meas. 24: 9-14.
- Prsinsell, CP, Ramsey PH, Ramsey PP (1994). Score gains, altitudes, and behaviour changes due to answer changing instruction. J. Ed. Meas. 31(4): 327-337
- Reilling E, Taylor RA (1972). A new approach to the problem of changing initial responses to multiple-choice questions. J. Ed. Meas. 9: 63-70.
- Roille PJ, Briggs CT (1952). Should students change their initial answers on objective tests? More evidence regarding an old problem. J. Ed. Psy. 43: 110-115.
- Sarason IG (1972). The Test Anxiety Scale: Concept and Research in Washington D.C. Hemisphere Publishing Company.
- Schwarz SP, McMorris RF, DeMers LP (1991). Reasons for changing answers: An evaluation using personal interviews. J. Ed. Meas. 28(2): 163-171
- Shaw MC and Wright JM (1967). Scales for the Measurement of Attitudes. New York: McGraw-Hill.
- Smith A, More JC (1976). The effects of changing answers on scores of non-sophisticated test takers. Meas. Ed. Guid. 8: 252-254.
- Stoffer GR Davis KE, Brow JB (1977). The consequences of changing initial answers on objective tests: A stable effect or a stable misconception. J. Ed. Res. 70: 272-277.
- Taylor JA (1975). A Personality Scale for Manifest Anxiety. J. Abn. Soc. Psy. 48: 285-290.
- Vabasa KT (1974). The Relationship between Personality Factors (Anxiety and Extraversion) and Academic Performance among some Nigerian Secondary School Students. Unpublished Masters Thesis, University of Ibadan.
- Vacchiano, RB and Strauss, PS (1968). The Construct Validity of the Tennessee Self Concept Scale. J. Ed. Psy. 24: 323-326.
- Yarworth JS, Carthier WJ (1978). Relationship of students' self concept and selected personal variables of participants in school activities. J. Edu. Psy. 70: 335-344.