

lications for test validity. *Educational and Psychological Measurement*, 32, 1039-1044.

Johnston, J. J. (1975). Answer-changing behavior and grades. *Teaching of Psychology*, 5, 44-45.

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

Ramsey, P. H. (1951). Power of univariate pairwise multiple comparison procedures. *Psychological Bulletin*, 90, 352-366.

Skinner, N. F. (1953). Switching answers on multiple-choice questions: Shrewdness or shibboleth? *Teaching of Psychology*, 10, 220-222.

Note

Requests for reprints should be sent to Philip H. Ramsey, Department of Psychology, Queens College, Flushing, NY 11367.

Misconceptions Tests or Misconceived Tests?

Richard A. Griggs
Sarah E. Ransdell
University of Florida

It was found that taking a psychology course in high school did not improve the performance of college students in an introductory psychology class on a modified version of Vaughan's (1977) misconceptions test. However, college experience did lead to a significant, but not large, improvement. A comparison of our results with those from previous studies, along with an analysis of specific test items, indicated that perhaps the test itself is misconceived.

Since Vaughan's (1977) study on misconceptions about psychology among introductory psychology students using her Test of Common Beliefs (TCB), numerous articles on this topic have appeared, especially in this journal. There were earlier relevant papers (e.g., McKeachie, 1960), but Vaughan's article seems to have had a greater impact. Later articles have dealt with several topics, including the relationship between the amount of misconception and subsequent performance in the introductory psychology course (Gutman, 1979), the amount of misconception as related to the number of accumulated college credits (Gardner & Dalsing, 1986), criticism of specific TCB items (Brown, 1984; Ruble, 1986), and faculty evaluations of items on the TCB (Gardner & Hund, 1983). Our article touches on all of these topics but focuses mainly on: (a) the effect of having taken a high school psychology course on the performance of college students on a misconceptions test and (b) a critical assessment of the test items that have been labeled *misconceptions*.

More and more high schools have added psychology to their curriculum, and enrollment in these courses has increased (Breland, 1978). For example, about 50% of our introductory psychology students at the University of Florida have had a high school psychology course. Two studies (Carstens & Beck, 1986; Hedges & Thomas, 1980) found that students who completed a high school psychology course began their college introductory course with a greater knowledge of psychology than students who had

not taken a high school psychology course. In addition, studies from McKeachie (1960) to Gardner and Dalsing (1986) have usually reported statistically significant (although not large) reductions in the amount of misconception as a function of taking a college introductory psychology course. Thus, a high school course might have a similar effect; students who have had a high school psychology course might do better on the misconceptions test than students who have not.

If there is such an effect, the overall amount of misconception observed today might be less than the 39.5% overall initial level reported by Vaughan (1977). Consistent with this hypothesis, Gardner and Dalsing (1986) reported a 22.9% overall misconception level, which would represent a substantial decrease from Vaughan's 39.5%. However, in analyzing Gardner and Dalsing's raw data, we discovered that they reported their overall level incorrectly.¹ In their study, the mean number of true responses across 60 test items was 22.9. Thus, they observed an overall level of almost 38%, only slightly less than the 39.5% observed by Vaughan (1977).

Gardner and Dalsing (1986) used a 60-item version of the TCB and included a third response category in addition to true and false responses. They deleted items that they considered unclear or that previous studies had shown few students judged to be true, and they allowed students to respond with a question mark if they had no opinion on an item. For reliability assessment of the Gardner and Dalsing results, we used their version of the TCB.

¹This mistake has been verified (R. M. Gardner, personal communication, September 26, 1986). It should be pointed out that the results in Figures 1 and 2 in Gardner and Dalsing (1986) are also incorrectly given as percentages of items marked true. However, this is only an error of labeling, and all of their results and conclusions pertaining to these two sets of data remain the same as originally reported.

Table 1. Intercorrelations Between Sex of Subjects, Number of College Credits, High School Psychology, Course Grade in Total Points, and Misconceptions Test Score

Variable	SEX	CREDITS	HSPSY	GRADE	MISC SCORE
SEX	—	-.10	-.17**	-.01	-.07
CREDITS		—	.23**	.16*	-.25**
HSPSY			—	-.02	.01
GRADE				—	-.18**
MISC SCORE					—

Note. The negative correlations are due to the coding scheme used. Male students were coded as 0; female students as 1. Students who had taken a high school course in psychology were coded as 2; students who had not taken such a course as 3. The misconceptions test score is the number of items marked true (errors).

* $p < .05$. ** $p < .01$.

If the corrected Gardner and Dalsing results are reliable, then the overall level of misconception has remained fairly constant over the decade since Vaughan's initial study. We decided to see if the particular test items labeled *misconceptions* have also remained relatively constant across studies. Because only Vaughan (1977) and Lamal (1979) reported test items for which at least 50% of their subjects responded true (Vaughan's definition of a misconception), our comparison involved these two studies, our study, and the study by Gardner and Dalsing, who provided us with their data. Thus, our analysis involved two studies from the late 1970s that essentially employed the original TCB and two recent studies that used a shorter version of the TCB with a third response category added. The results of this analysis, along with the subsequent critical assessment of the individual test items involved, are presented and used to argue that the misconception test itself may have been misconceived.

Method

The subjects were 273 undergraduates at the University of Florida; 198 were entering freshmen who had no college credits, and 75 had earned nine or more college credits; 135 had taken a high school psychology course and 138 had not. The test was administered during the second class meeting in an introductory psychology class.

We used the 60-item modified form of Vaughan's (1977) TCB reported by Gardner and Dalsing (1986). As in Gardner and Dalsing's study, subjects were told that the items represented controversial statements about psychology and that we were interested in their opinions about them. Three responses were allowed: true, false, and ? (don't know/no opinion). Subjects also reported whether or not they had taken a psychology course in high school, their sex, and the number of college credits they had completed.

Results and Discussion

The number of students marking each statement true, false, and ?, as well as the number of true, false, and ? responses for each subject, were computed.

We computed correlations between the sex of the subjects (SEX), subjects' misconceptions test score (MISC SCORE) defined as the number of items marked true, completion or not of a high school psychology course (HSPSY),

final grade in the college course (GRADE), and the number of college credits earned (CREDITS). Final grade in the course was coded as the total number of exam points earned in the course (0-180-point scale). The intercorrelations matrix is given in Table 1.

The correlation between HSPSY and MISC SCORE was not significant. The mean number of misconceptions for subjects who had taken a high school psychology course was 23.99; for those who had not, 24.12. Thus, taking a high school course in psychology did not improve performance on the misconceptions test.

However, the correlations between HSPSY and SEX and HSPSY and CREDITS were significant. These results are consistent with recent descriptions of the high school course. More female than male students take high school psychology (Goldman, 1983). Fifty-nine percent of our subjects who had taken a high school course in psychology were women. The relationship between HSPSY and CREDITS indicates that more beginning students had taken such a course. This too agrees with the recent trend toward more high school psychology courses and a growing enrollment in them (Breland, 1978).

The significant correlation between CREDITS and MISC SCORE reflects the fact that beginning students did significantly worse on the misconceptions test. Subjects with no college credits had a significantly higher mean number of misconceptions (24.68) than did subjects with some college experience (22.39), $t(271) = 2.41, p < .05$. This result is consistent with Gardner and Dalsing's finding that as the number of college credits increases, the amount of misconception decreases. However, it is clear that this reduction is not of much practical significance because the omega-squared score for this significant effect shows that only 1.74% of the variance in the misconceptions scores is accounted for by the number of college credits.

The significant correlation between CREDITS and GRADES is not surprising. Students with some college experience did better in the course. The mean number of points earned in the course for beginning students was 141.4, and for students with some college experience, 146.6, $t(260) = 2.15, p < .05$.² But, again, the practical significance of this finding is limited. The omega-squared indi-

²Only 262 students received grades in the course. Thus, the number of degrees of freedom is lower in this case.

Table 2. Data for TCB Items With Proportion of True Responses .50 or Greater in at Least Two Studies

TCB Item	Proportion of True Responses			
	Vaughan (1977)	Lamal (1979)	Gardner and Dalsing (1986)	Present Study
1. To change people's behavior toward members of ethnic minority groups, we must first change their attitudes.	.92	.91	.84	.76
2. By feeling people's faces, blind people can visualize how they look in their minds.	.83	.78	.70	.83
3. Children memorize much more easily than adults.	.66	.68	.52	.66
4. Unlike man, the lower animals are motivated only by their bodily needs — hunger, thirst, sex, etc.	.69	.53	.59	.62
5. "The study of the mind" is the best brief definition of psychology today.	.57	.54	.62	.57
6. The more you memorize by rote (e.g., poems), the better you will become at memorizing.	.50	.62	.50	.53
7. The best way to ensure that a desired behavior will persist after training is completed is to reward the behavior every single time it occurs throughout training (rather than intermittently).	.77	.70	—	.68
8. Fortunately for babies, human beings have a strong maternal instinct.	.73	—	.62	.77
9. The ability of blind people to avoid obstacles is due to a special sense that develops in compensation for their absence of vision.	.65	—	.63	.68
10. By giving a young baby lots of extra stimulation (e.g., mobiles and musical toys), we can markedly increase his intelligence.	—	.65	.60	.61
11. Psychiatrists are defined as medical people who use psychoanalysis.	.67	.57	—	.52
12. Boys and girls exhibit no behavioral differences until environmental influences begin to produce such differences.	.61	—	.56	.55
13. The high correlation between cigarette smoking and lung cancer proves that smoking causes lung cancer.	—	.55	.56	.66
14. Genius is akin to insanity.	.53	.56	—	—
15. In love and friendship, more often than not, opposites attract one another.	—	—	.51	.51

Note. Proportions not given are less than .50.

cates that only 1.29% of the variance in the students' grades is accounted for by the number of college credits.

The last significant correlation, GRADES and MISC SCORES, is consistent with Gutman's (1979) finding that students who received A or B grades in an introductory psychology course had fewer misconceptions at the beginning of the term than students who did poorly in the course. However, our results indicate that this finding may only apply to experienced students. When the correlation between GRADES and MISC SCORE was computed separately for beginning students and students with some college experience, it was only significant for the experienced students ($r = -.35, p < .01$); for beginning students, $r = -.05$.

With respect to overall level of misconception, 40% of the responses were marked true, 48% were marked false, and only 12% were marked ?. This last percentage is almost identical to that observed by Gardner and Dalsing (1986).

In addition, the overall level of misconception is comparable to Vaughan's (1977) 39.5%, Lamal's (1979) 41% to 43%, and the 38% observed by Gardner and Dalsing (1986). Thus, it does not appear that the overall level of misconception has decreased since Vaughan's initial study using the complete TCB.

Next we consider our analysis of the test items that have been labeled *misconceptions*. Vaughan (1977) defined a misconception to be any test item marked true by at least 50% of the subjects. Using this criterion, Vaughan observed 23/80 misconceptions; Lamal, 22/76; Gardner and Dalsing, 16/60; and we found 23/60. There is consistency across these four studies. If we look at the 60 items these four studies have in common, only 27 different items constitute misconceptions: Nine items were misconceptions in all four studies, 8 appeared in three studies, 4 were misconceptions in two studies, and 6 in only one study.

Using the criterion that in order to be classified as a reliable misconception a test item has to have been so categorized in at least two of the four studies, we are left with a group of only 21 items. Six of these 21 items have already been criticized in published articles (Brown, 1984; Ruble, 1986). If these criticisms are accepted, we are left with only 15 items. These 15 items are given in Table 2, along with the percentage of subjects responding true in each of the four studies being examined.

The items are ordered first according to how often they were categorized as a misconception. Then, within each of these groups (i.e., the items classified as misconceptions in all four studies, those classified as such in three studies, etc.), the items are ordered according to the average amount of misconception, from highest to lowest. Thus, Item 1 is the strongest misconception, and Item 15 is the weakest.

The fact that there are only 15 reliable misconceptions across the studies reviewed should not be too surprising. Brown (1983) concluded that "widely shared misconceptions are rare indeed" (p. 209). Panek (1982) found that only 2 of 10 suspected misconceptions about the psychology of aging were believed by the majority of his undergraduate students. But what about the 15 items remaining in the present analysis? Are they all unquestionably misconceptions? It appears not.

Even psychologists do not rate several of these items as completely false. Gardner and Hund (1983) had psychologists and other scientists assign ratings to 20 items from the TCB on the following scale: *completely false* (1), *mostly false* (2), *partly false and partly true* (3), *mostly true* (4), and *completely true* (5). Ten (Items 1, 2, 3, 4, 6, 7, 8, 9, 12, and 14) of the 15 items in Table 2 were included. The ratings of psychologists for these 10 items ranged from 1.20 (for Item 7) to 2.59 (for Item 1) with a mean rating of 1.97 (mostly false on the 5-point scale). The means for six of the items were at least 2.00, and four were below 2.00. Thus, psychologists do not even agree that these items are completely false.

Further evidence of this lack of agreement is provided by Lamal (1979). He had seven introductory psychology instructors rate items from the TCB with respect to their agreement with each item's falsity. In addition, the instructors rated their concern with how their students answered each item. There were nine items that at least one instructor did not rate as false. Items 1, 6, and 9 were in this set of nine. In addition, for 7 (Items 1, 2, 3, 4, 6, 10, and 11) of the 11 items in Table 2 considered in Lamal's study, fewer than half of the instructors said they would be unhappy if their students missed the item. Thus, both the falsity and the importance of several of the items have been questioned by psychology instructors.

It is also the case that a college-level introductory psychology course does not change many of these misconceptions. Lamal (1979) provided some pretest (at the beginning of the term) and posttest (at the end of the term) results for 11 of the 15 items in Table 2. Performance on 6 of the 11 items (Items 1, 2, 3, 5, 6, and 14) did not change much. There was a substantial decrease (greater than 5%) for three items (Items 7, 11, and 13) and a substantial increase (greater than 5%) for two of the items (Items 4 and 10). This

lack of effect may be due partially to the fact that some of these misconceptions are not even addressed in introductory psychology texts. For example, Items 2 and 9 deal with blind people, a topic not covered in most current introductory texts.

Given that such problems exist even for these 15 test items, one begins to doubt the value of the misconceptions test. The number of misconceptions normally observed is already fairly low, and the problems described indicate that this low number is likely to be an overestimate. Many of these problems are probably due to the true-false test format, which is not a good way to assess students' knowledge. In his book on teaching tips for beginning college teachers, McKeachie (1986) did not advocate the use of true-false tests. He argued that "students can usually figure out reasons why any particular item can be either true or false" (pp. 90-91). Thus, it would appear that rather than assessing student misconceptions, the misconceptions test itself may have been misconceived. As Barnett (1986) argued, perhaps teachers of psychology should not be so concerned with assessing what beginning students know or misconceive to be true or false about psychology; but, rather, it is more important that we "help them to appreciate the complexities of human behavior that are rarely captured in simple true-false statements" (p. 64).

References

- Barnett, M. A. (1986). Commonsense and research findings in personality. *Teaching of Psychology, 13*, 62-64.
- Breland, N. S. (1978). A survey of precollege psychology in New Jersey. *American Psychologist, 33*, 959-961.
- Brown, L. T. (1983). Some more misconceptions about psychology among introductory psychology students. *Teaching of Psychology, 10*, 207-210.
- Brown, L. T. (1984). Misconceptions about psychology aren't always what they seem. *Teaching of Psychology, 11*, 75-78.
- Carstens, C. B., & Beck, H. P. (1986). The relationship of high school psychology and natural science courses to performance in a college introductory psychology class. *Teaching of Psychology, 13*, 116-118.
- Gardner, R. M., & Dalsing, S. (1986). Misconceptions about psychology among college students. *Teaching of Psychology, 13*, 32-34.
- Gardner, R. M., & Hund, R. M. (1983). Misconceptions of psychology among academicians. *Teaching of Psychology, 10*, 20-22.
- Goldman, J. J. (1983). Recent trends in secondary school psychology: The decade from Oberlin to the HBCP. *Teaching of Psychology, 10*, 228-229.
- Gutman, A. (1979). Misconceptions of psychology and performance in the introductory course. *Teaching of Psychology, 6*, 159-161.
- Hedges, B. W., & Thomas, J. H. (1980). The effect of high school psychology on pre-course knowledge, midterm grades, and final grades in introductory psychology. *Teaching of Psychology, 7*, 221-223.
- Lamal, P. A. (1979). College students' common beliefs about psychology. *Teaching of Psychology, 6*, 155-158.
- McKeachie, W. J. (1960). Changes in scores on the Northwestern misconceptions test in six elementary psychology courses. *Journal of Educational Psychology, 51*, 240-244.
- McKeachie, W. J. (1986). *Teaching tips: A guidebook for the beginning college teacher* (8th ed.). Lexington, MA: Heath.

- Panek, P. E. (1982). Do beginning psychology of aging students believe 10 common myths of aging? *Teaching of Psychology*, 9, 104-105.
- Ruble, R. (1986). Ambiguous psychological misconceptions. *Teaching of Psychology*, 13, 34-36.
- Vaughan, E. D. (1977). Misconceptions about psychology among introductory psychology students. *Teaching of Psychology*, 4, 138-141.

Notes

1. We thank Steve Lange for his assistance with the data analyses, and Rick Gardner and Peter Lamal for providing us with further information about their studies.
2. Requests for reprints should be sent to Richard A. Griggs, Department of Psychology, University of Florida, Gainesville, FL 32611.

Of Windmills and Rope Dancing: The Instructional Value of Narrative Structures.

L. Dodge Fernald
Harvard University

Rationale for the instructional value of narrative structures has been emerging steadily in cognitive psychology. The organizational framework that narrative structures offer may be useful for college-level instruction. Introductory textbook chapters were prepared in three modes: traditional, using successive examples to illustrate psychological concepts; intermediate, using a brief sequence of related events; and narrative, using an extended narrative structure. Undergraduates preferred the narrative mode and considered it more enjoyable and more useful for learning.

When Don Quixote tilts with long-armed giants, Sancho Panza says they are windmills. When his squire is blanket-tossed by khaves, the knight declines to intervene, saying he is prevented by the supernatural. These differences in perspective occur throughout *Don Quixote*, providing a framework for discussions of reality and morality in the Spanish Golden Age. Cervantes is not alone, however, in using the narrative for instructional purposes. Countless authors have used the narrative as a fundamental method for the transmission of knowledge in all societies, Eastern and Western, ancient and modern.

Background

In his early work, Hermann Ebbinghaus (1885) found that stanzas from Byron's *Don Juan* were learned more readily than nonsense syllables and that, in general, learning time varied with the nature of the material. Since that time, the rationale for the narrative as an instructional device has been emerging steadily in cognitive psychology, just as the mind has been emerging as a pattern-making, pattern-recognizing process (de Bono, 1984; Graesser, Haut-Smith, Cohen, & Pyles, 1980; Wessells, 1982). Efficient mastery of verbal material requires an overall form or framework, which sometimes can be provided by a narrative. The organization imposed by such structure may become a key factor in learning and retention.

The value of an organizing theme was demonstrated experimentally when subjects learned lists of words under two conditions. When these words were woven into a story, they were recalled six to seven times more frequently than when no narrative was used. The median recall scores for the narrative and control conditions were 93% and 13%, respectively, a difference attributed to the facilitating effect of thematic organization in learning (Bower & Clark, 1969).

The narrative has also been demonstrated to facilitate memory for prose passages. In one study (Graesser et al., 1980), prose passages differed in degree of familiarity, use of an outline, and mode of presentation (expository or narrative). The degree of familiarity and use of an outline had little or no effect on memory. However, the narrative mode of presentation was highly influential, predicting from 38% to 84% of the variance in recall. If the material has no obvious organization or theme, then learners sometimes impose structure upon it; this is called *subjective organization* (Tulving, 1966).

Mnemonists learn large amounts of material by using a narrative. The items to be remembered are inserted into a well-known story in some convenient manner, and then the narrator simply extracts them when retelling the story later.

Another advantage of the narrative is that it can markedly increase readers' motivation. Readers of *Gulliver's Travels*, for example, enjoy the hero's trip to Lilliput, where politicians display their readiness for office by dancing on a tightrope. Once installed, they debate whether to break an egg at the big or little end. With this satirical story format, Jonathan Swift greatly augmented interest in English politics. The earliest narrative poems, Homer's *Iliad* and *Odyssey*, present the Greek heroic tradition in an equally stimulating manner. And today, Trudeau's *Doonesbury* has become a highly popular, seminarrative commentary on contemporary American life. Thus, the narrative can be pedagogically useful in at least two ways: It can serve as a