DOCUMENT RESUME

ED 299 138 SE 049 663

AUTHOR Beins, Bernard C.
TITLE Student Knowledge of Statistics: To Know Is To Do.
PUB DATE 87
NOTE 10p.; Paper presented at the Annual Meeting of the American Psychological Association (Washington, DC, August 27-September 1, 1987).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *College Mathematics; Comprehension; Computation; Concept Formation; *Difficulty Level; Educational Research; Higher Education; *Knowledge Level; Mathematical Concepts; *Mathematics Achievement; Mathematics Education; *Mathematics Tests; Performance; *Statistics

ABSTRACT

Students in a college introductory statistics class were evaluated with conceptual and computational tests and the relationship between their levels of knowledge on the two forms of testing was assessed. There were significant correlations between their abilities to perform computations and to answer more conceptual questions on individual tests and across separate tests, for the final examination, and for the total number of points earned throughout the semester. The correlations indicate that the two styles of testing provide partially redundant, but not totally overlapping, information about student knowledge. Further, although student averages did not differ in the two types of tests, they seldom preferred only conceptual tests of their knowledge, judging the computational tests a better means of evaluation. Applications of these results extend to prediction of missing test scores and to the testing of students for whom English is a second language. (Author)

***********************************************************************
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
***********************************************************************

ED299138

U S DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Bernard C. Beins
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

# Student Knowledge of Statistics: To Know is to Do

Bernard C. Beins
Ithaca College
Ithaca, New York 14850

## Abstract

Students in an introductory statistics class were evaluated with conceptual and computational tests and the relationship between their levels of knowledge on the two forms of testing was assessed. There were significant correlations between their abilities to perform computations and to answer more conceptual questions on individual tests and across separate tests, for the final examination, and for the total number of points earned throughout the semester. The correlations indicate that the two styles of testing provide partially redundant, but not totally overlapping, information about student knowledge. Further, although student averages did not differ in the two types of tests, they seldom preferred only conceptual tests of their knowledge, judging the computational tests a better means of evaluation. Applications of these results extend to prediction of missing test scores and to the testing of students for whom English is a second language.

SE 049 663

BEST COPY AVAILABLE

Student Knowledge of Statistics:

To Know is to Do

Bernard C. Beins

Ithaca College

Statistics classes are seen as critical by teaching psychologists, as evidenced by the nearly universal requirement that our majors take such a course (Bartz. 1981). At the same time, students seem to show a marked aversion for the class, and instructors have attempted to ameliorate this situation through a number of techniques (e.g., Beins, 1985; Dillbeck, 1983; Hastings, 1982; Jacobs, 1980). In a continuing attempt to identify variables that will lead to a successful course, this paper will report on the relative efficacy of two testing formats in a statistics course, present an analysis of student responses to the tests and suggest some applications of this knowledge.

## Method

### Subjects

Thirty-one students enrolled in an introductory level statistics class provided the data for this study. The class consisted largely of students majoring in social and behavioral sciences or nursing.

### Procedure

Students in the class took four hourly tests and a cumulative final exam. Each test consisted of two portions--an initial, closed book segment involving multiple choice, sentence completion and fill-in items, and definitions. A computational segment involved the typical form of statistical tests, with students selecting statistical tests when given experimental descriptions, creating graphs and figures, and solving problems; students were permitted the use of books, notes and calculators without statistical capabilities (e.g., those that will compute sums of squares, means, variances, automatically) for the computational part. The tests were scheduled for a 50-minute class period, although the students could use the 10-minute interclass interval if they desired.

Near the end of the term, the students also responded in class to an anonymous questionnaire about their impressions of the structure of the course. The questions of relevance here were "Did the open book (closed book) portions of the tests reflect your knowledge of statistics?" Responses were on a seven-point scale. On the last day of classes, they were asked whether, if given a choice, they would prefer their grades to be based on the conceptual (closed-book) or computational (open-book) test, or on a combination of the two. I made it clear to them that their course grades would be based on scores from both kinds of tests. They were asked to write their names with their replies to this question.

## Results & Discussion

Relationship between conceptual and computational tests. There was a marked association between students' scores on the computational and conceptual segments of the tests. When the point totals for all quizzes were summed, the results revealed that students who did well in the computational portions also did well conceptually, $r(29) = 0.63$, $p < 0.001$. Likewise, conceptual and computational final exam grades are related, $r(29) = 0.49$, $p < 0.01$.

In general, the tests showed a considerable degree of intercorrelation. Pairing each possible set of scores on conceptual and computational segments for each test yielded 45 possible correlations. A z-test revealed that, of these, 23 were significant at or beyond the 0.05 alpha level with a two-tailed test; another two were significant with one tail. With respect to the conceptual-computational relationships for individual tests, only quiz 1 failed to show a significant correlation. The values appear in Table 1; the means and standard deviations for each test also appear in the table. The first test in a given class may be a time of adjustment for many students in which they acclimate to the nature of the test, so a low correlation is not totally surprising. The other quizzes seemed to indicate that when students grasped the concepts, their computations were also performed adequately.

although quiz 3, which produced the highest mean on both conceptual and computational portions, showed inconsistent correlations with performance on other tests. It also showed the highest means and relatively small dispersion, both of which may have led to the low correlation, significant only at $p = 0.08$.

---

Insert Table 1 About Here

---

Student responses to the tests. In addition to considerations of the relationship between conceptual and computational test modes, it might be useful to know whether student preferences for different kinds of tests are related to ultimate performance in the class. Consequently, I performed an Analysis of Variance to see whether those students who professed a preference for the open-book format ($n = 12$) wound up with higher average grades than the students favoring the combination format ($n = 9$). There was no difference in the class averages (in percentages) between students who wanted open-book tests alone ($M = 71.31$) and those who would have liked both ($M = 68.13$), $F(1,19) = 2.11$, $p= 0.159$. Likewise, the performance on the conceptual part ($M = 66.22$) and on the computational part ($M = 73.21$) were not significantly different, $F(1,19) = 2.134$, $p = 0.157$. The interaction was also non-significant, $F < 1$.

As a part of a final, anonymous questionnaire (filled out on a day when 28 of the 29 enrollees were present), students were queried about the degree to which they felt that the open- and closed-book segments tested their knowledge. Rating each format on a seven-point scale, the students judged the Open-Book portions (Mean = 2.71) as being better tests of their knowledge than Closed-Book segments (Mean = 3.75), $t(27) = 2.78$, $p< 0.01$. Their belief that the open-book, computational segments were better tests of their knowledge is reflected only in nonsignificantly higher mean scores on those segments, $t(28) = -1.76$, $p >0.05$. Scores on the open-book tests were only 3.16 percent higher than closed-book scores: 70.03 versus 66.77, respectively.



ERIC

(Obviously, the magnitude of this difference will vary according to the relative difficulty of the two segments; the important point here is that one's feelings of competence or one's reaction to the nature of the test may not adequately reflect actual knowledge or adequacy of the test instrument.) Thus, the students showed less satisfaction with the conceptual portion of the test even though their scores were not systematically lower than the computational part, and the two were actually related to one another. When I when through a hypothetical exercise of assigning grades based on the conceptual and computational scores separately, the difference in grade point averages for the class was relatively small: 2.07 for conceptual versus 2.24 for computational. This difference reflects a slightly lower percentage average on the concepts combined with my own standards as to the minimum score that is appropriate for a particular letter grade.

The results here suggest that a conceptual test of statistics might not be totally indispensable in assessing student knowledge in the class. The consistently high correlations between computational and conceptual information makes it seem plausible that students could be given a test whose format is consistent with their own desires. It should be noted that, simply because a student selects one format over the other, the resultant grade will not necessarily be higher than if a different structure were used. To illustrate this fact, I will point out that of two students who would stated a preference for only closed-book tests, one received a grade of A in the course, the other an F. At the same time, if an instructor wants as complete an assessment of the students' knowledge as possible, the two different formats can be used; the results, while correlated, are not perfectly overlapping and the two test types are not totally redundant.

The pattern of preferences by students suggests that they probably feel that closed-book, conceptual tests are harder, as reflected in their belief that open-book tests assess their knowledge better than closed-book tests. This inference is based on my assumption that so-called "easier" tests are viewed more positively because they reinforce the students' desires to feel competent in statistical procedures.

<u>Applications</u>. There is another reason to consider including both kinds of tests. The findings here provide a potentially useful means of estimating missing test scores. A simple linear regression technique could be employed instead of merely using the student's overall average as the estimate. According to the present data, the final exam score could serve as a predictor of a missing test score, as could the total point value accumulated over all tests or, if the computational and conceptual tests were given on separate days, one of these two could be used quite adequately to predict the other.

Another potential use for this estimation technique involves the testing of students whose native languages in not English. By the admission of one of the two students in this class for whom English is not the first language, the closed-book tests posed difficulty with respect to language considerations. When predictions of total conceptual scores for both students were made from the computational scores, the difference between actual and predicted scores was not significant in either case, $z = 0.36$ and $0.42$, both $p$'s $> 0.05$. Both of these students had at least adequate conversational skills in English, so it is not clear students with a poorer command of English would be able to understand the textbook and the problems, even if the skill level is potentially very high. One reason for optimism here is that neither the student with self-admitted language problems nor the other student deviated significantly from their predicted conceptual scores, based on the computational. It may be the case that when a student has a basic ability to comprehend a textbook, the linguistic factors involved in computation are not important. An instructor would need to decide whether a poor grade on the conceptual part of a test was due to language deficiencies or to lack of statistical knowledge, but if the low mark were due to linguistic factors, an estimate of the language-based portion of the test might prove satisfactory. The criterion for using the estimate rather than the actual score might rest on the instructor's subjective assessment of the student, or it might be based on systematically lower grades on language-dependent questions, or on any of a number of different factors. One caveat, however, is that a poor command of English may ultimately affect computations as well. Obviously, consideration of mitigating factors will be required in adoption of any of these suggestions.

References

Bartz, A. E. (1981) The statistics course as a departmental offering and major requirement. Teaching of Psychology, 8, 106.

Beins, B. C. (1985). Teaching the relevance of statistics through consumer-oriented research. Teaching of Psychology, 12, 168-169.

Dillbeck, M. C. (1983). Teaching statistics in terms of the knower. Teaching of Psychology, 10, 18-20.

Hastings, M. W. (1982). Statistics: Challenge for students and the professor. Teaching of Psychology, 9, 221-222.

Jacobs, K. W. (1980). Instructional techniques in the introductory statistics course: The first class meeting. Teaching of Psychology, 7, 241-242.



ERIC

Table 1

Intercorrelations among conceptual and computational segments of statistics tests and final exam. (Means and standard deviations are in margins.) [a]

| | | Computational | | | | | Conceptual | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Test # | 1 | 2 | 3 | 4 | Final | 1 | 2 | 3 | 4 | Final | Mean | S.D. |
| | 1 | -- | .20 | .53* | .15 | .28 | .12 | .40* | .26 | .19 | .20 | .69 | .20 |
| | 2 | | -- | .44* | .20 | .52* | .26 | .54* | .10 | .26 | .41* | .76 | .15 |
| Comput. | 3 | | | -- | .26 | .45* | .45* | .32 | .32 | .55* | .46* | .86 | .14 |
| | 4 | | | | -- | .45* | .44* | .48* | .22 | .37* | .42* | .58 | .22 |
| | Final | | | | | -- | .20 | .51* | .32 | .45* | .49* | .61 | .12 |
| | 1 | | | | | | -- | .39* | .16 | .62* | .24 | .80 | .20 |
| | 2 | | | | | | | -- | .22 | .51* | .54* | .73 | .20 |
| Concept | 3 | | | | | | | | -- | .42* | .13 | .85 | .15 |
| | 4 | | | | | | | | | -- | .31 | .73 | .19 |
| | Final | | | | | | | | | | -- | .62 | .19 |

*$p < 0.05$

[a]Test topics are as follows:

Test 1: Graphing, Correlation/Regression

Test 2: Normal distribution; Sampling distributions

Test 3: Statistical inference; one-sample z-test and t-test

Test 4: Two sample t-test; analysis of variance

Final: Cumulative test