DOCUMENT RESUME

ED 397 717 HE 029 317

AUTHOR Duckwall, Julia M.; Wilson, Steven J.

TITLE Can Core Components in Mathematics Courses Replace Comprehensive Common Course Finals? AIR 1996 Annual Forum Paper.

PUB DATE 6 May 96

NOTE 20p.; Paper presented at the Annual Forum of the Association for Institutional Research (36th, Albuquerque, NM, May 5-8, 1996).

PUB TYPE Reports - Evaluative/Feasibility (142) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Academic Standards; Case Studies; *College Mathematics; Community Colleges; Comparative Analysis; *Course Content; Curriculum Development; Grade Inflation; Grades (Scholastic); Higher Education; *Institutional Research; Mathematics Education; Mathematics Instruction; *Mathematics Tests; Program Descriptions; Scores; Student Evaluation; Teaching Styles; Test Bias; Testing

IDENTIFIERS *AIR Forum; *Final Examinations; Johnson County Community College KS

ABSTRACT

In an effort to balance common academic outcomes, as measured by common course finals, with faculty's individual teaching styles, Johnson County Community College (Kansas) mathematics faculty developed "core components" for 10 of the mathematics courses offered in Spring 1995. The core components were designed by faculty teaching each course, and were deemed essential elements of the course. Additionally, all faculty teaching each course agree to grade the core components identically and include assessment of them as part of course final examinations. The students' core component scores were then compared with their final grades. Results indicated a positive correlation between student scores on core components and course grades as a whole, suggesting that this approach could help in balancing common academic outcomes with instructional freedom without weakening academic standards. However, several other issues arose that require further study, including use of review questions on material covered in prerequisite courses, lower correlations in some courses, concern about the core components' ability to meet final exam objectives, possible bias in scoring, and potential for grade inflation. (Author/MSE)

# Can Core Components in Mathematics Courses Replace Comprehensive Common Course Finals?

by

Julia M. Duckwall and Steven J. Wilson

Johnson County Community College

12345 College Boulevard

Overland Park, Kansas 66210-1299

913-469-8500



Paper presented at the 36th Annual Forum of the Association for Institutional Research, May 6, 1996.

This paper was presented at the Thirty-Sixth Annual Forum of the Association for Institutional Research held in Albuquerque, New Mexico, May 5-8, 1996. This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of Forum Papers.

Jean Endo
Editor
AIR Forum Publications

18


## Abstract

In an effort to balance common academic outcomes (measured by common course finals) with academic freedom of instruction, community college mathematics faculty developed "core components" for ten of their mathematics courses in Spring 1995. These core components were designed by the faculty teaching each course and deemed essential elements of the course. Additionally, all faculty teaching each course agreed to grade the core components identically and include assessment of them as a part of course final examinations. The students' core component scores were then compared with their final grades. Results indicate that the students' scores on the core components did correlate with their math course grades for the courses as a whole.


## Introduction

Today's society demands accountability from all public entities, and educational institutions are no exception. Accountability demands have not only changed the way accreditation bodies evaluate schools, but have caused legislative bodies to become more actively involved in the outcomes assessment process. In order to meet these demands, some commonality of assessment that is able to be generalized across different sections of the same course is needed. At the college level, though, a long history of academic freedom persists, and college instructors expect to be able to individualize their teaching process and assessment methods. This study was undertaken to try and resolve these conflicting viewpoints--could faculty retain control over their own courses while meeting the demand of commonality required for accountability?

## Background

Prior to 1992, faculty in the mathematics department at Johnson County Community College used common course final exams. These two-hour finals were prepared and periodically revised by faculty in the department who traditionally had the responsibility for curricular issues. Originally, the purpose of these common finals was to evaluate student learning, as defined by the department faculty as a whole, as well as to encourage instructors to cover all the required course content.

Current movements in education have greatly affected the nature of mathematics teaching. Calls for reform were issued by the National Council of Teachers of Mathematics (Commission on Standards for School Mathematics, 1989), and the American Mathematical Association of Two-Year Colleges (Cohen, 1995). Both documents advocate increasing emphasis in such areas as problem-solving, modeling, conceptual understanding, and the ability to communicate mathematical ideas. Technology is also changing the nature of mathematics teaching. Within the last decade, the advent of the graphing calculator has enabled students to avoid some of the tedious nature of graphing, and move toward understanding the implications of the results. These changes, however, have not been welcomed by all members of the mathematical education community. In fact, the faculty of the mathematics department at Johnson County Community College form a wide cross-section of that community, with diverse opinions on these issues. Since departmental agreement on many aspects of the teaching of mathematics was not possible, instructors began to desire more academic freedom to pursue what they considered to be more effective teaching and assessment strategies.

Due to these concerns, instructors received permission to produce alternative forms of course final exams, beginning in the spring of 1992. These alternative exams had to cover the defined course content in at least as comprehensive a fashion as the departmental final exams, and required departmental approval before their use. Some alternative final exams have been developed.

In the fall of 1994, administrators at Johnson County Community College voiced concern about the outcomes assessment process across the college at the departmental level. In the mathematics department, the most obvious component of the current assessment strategy (and the one that was most generalizable across different sections of the same course) was the common course final exam. Although every instructor gave the departmental final exam in courses where alternatives did not exist, grading was done independently, by whatever means the instructor deemed appropriate. Therefore, although the same exams may have been given, the scores could not be used as a consistent outcomes assessment measure across the entire course. When an alternative existed, of course, no common outcomes measure could be provided by final exams. The Final Exam Committee was established and charged by the mathematics program director to examine the current process, obtain information from departmental faculty, and facilitate discussion of the issues with the hope that faculty could reach consensus on how to meet the request for a consistent outcomes assessment measure. Essentially, if the committee reached its objective, the resulting final exam would then be simultaneously providing several measures, including:

- student academic achievement (as defined by the department as a whole)
- student academic achievement (as defined by Carl Perkins funding requirements)
- student academic achievement (as defined by the individual instructor)
- instructor accountability (as defined by administrators)

A survey of all faculty in the mathematics department (which achieved a response rate of 96% for full-time and 64% for part-time faculty) produced some interesting results. Basically, it was determined that there were many philosophical differences about what constitutes good mathematics teaching and testing. Rarely did even 70% of the faculty have the same opinion about any particular final exam issue.

The committee, after considerable deliberation, proposed a solution, to try to satisfy both the accountability and academic freedom issues; specifically, a two-part final exam. The first part, called the "core," would provide a completely consistent measure of academic achievement of the basic concepts across all sections of the course by requiring that it be included on all exams, without alteration, and that it be graded identically by all instructors. It was proposed that these core component items would comprise no more than half of the points that could be awarded on the final. The second part, called the "free portion," would allow for differences in individual instructor's professional judgements. The format of the free portion, the questions, the manner of administration, the grading, the use of calculator, formulas, and/or notes, would be completely at the discretion of each instructor.

To evaluate whether this solution was even feasible, subcommittees of faculty, involving all full-time and interested part-time mathematics instructors at the college, were established for the ten mathematics courses that offered at least four sections of each course in Spring 1995 (Table 1). First, those faculty teaching each course reviewed the established course outcomes, leading to revised outcomes in four courses. Second, they reviewed items on the established final exams to determine whether previously written questions could be used for the core items. Third, new core exam items or modified available exam items were written so that each core question was a basic question of moderate difficulty for the course, met one major outcome in the course outline, and was unbiased with regard to calculator use. Also, the committee insured that every major outcome of the course was addressed by at least one core question. Finally, to be included in the core, it was decided that at least 80% of the faculty had to agree that the component was an essential element of the course. Time was set aside during regularly scheduled in-service sessions for the ten committees to meet, allowing for almost all faculty to be involved in this process for the courses in which they expressed interest.

Table 1. Subcommittee Results for Each of the Ten Mathematics Courses Selected for Study

| Math Course | Were the outcomes revised? | Number of new/changed questions | Number of questions on core | Minimum % faculty approving |
|---|---|---|---|---|
| Fundamentals of Math | No | 1 | 11 | 100% |
| Introduction to Algebra | No | 2 | 10 | 91% |
| Intermediate Algebra | Yes | 7 | 10 | 90% |
| Business Math | No | 2 | 9 | 100% |
| College Algebra | Yes | 5 | 10 | 83% |
| Trigonometry | No | 0 | 5 | 100% |
| Pre-Calculus | No | 2 | 9 | 100% |
| Statistics | No | 3 | 7 | 100% |
| Calculus I | Yes | 5 | 10 | 100% |
| Analytic Geometry-Calculus I | Yes | 5 | 11 | 83% |

Course grades across various instructors are recognized as unsuitable common measures of achievement. Milton, Pollio, and Eison (1986) provide evidence that grades carry different meanings to different faculty (not to mention other groups). Olson's (1989) study in the Dallas school district is a typical example of the poor correlation between teacher assessments and standardized tests. In contrast, the faculty in this project were not interested in what makes faculty assessment strategies different from others, but what may be common. With this approach, course grades become the measure of each individual instructor's assessment of student performance.

## Purpose

The purpose of this study was twofold:

1. to determine if consensus among faculty with diverse opinions was possible in constructing a set of final exam questions (called core components) that measure essential elements in a series of mathematics courses and in the assessment of student knowledge, and
2. to determine whether success on these core components, required as part of each course's final examination, was related to each instructor's assessment of a student's performance (as measured by course grades).

## Methodology

The core component items, constructed by faculty to measure essential elements in a series of mathematics courses, were included as part of the Spring 1995 final exams in the ten selected courses. The core component items were graded identically by the instructors in the agreed upon manner. The scores were reported separately along with the course grade for each student who took the final examinations. Scores and grades were then grouped by course. No identifiers enabling links to individuals were used for any participating faculty or students.

To determine if the core components were related to the course grades, Pearson correlations between the core component scores and the final grade were computed for each course and for each instructor. To determine if certain course examinations were much more difficult than other course examinations, the mean core component scores were compared to the total points possible and compared across courses. To determine if the data appeared linear (or at least showed a monotonic trend), cross tabulation tables were created. To determine whether students were obtaining high grades without the core component knowledge or were receiving low grades in spite of their knowledge, cross tabulation tables were also examined. To find if any individual instructors' assessments (i.e., course grades) were unrelated to the core components, individual low correlations and high p-values were examined and compared to the overall correlation for the course.

## Results

The average grades and standard deviations for the courses were very similar across sections and most of the course means for the core components fell within two standard deviations of the total number of points possible (see Table 2). The courses for which this was not true were College Algebra and Calculus I, where the mean scores on the core components were more than two standard deviations below the total points possible.

Since the total points possible on each core varied from 5 to 55, comparison of mean core scores across courses was accomplished with percentages (Table 2). The highest average score percentage was in Business Mathematics with 76.1% while the lowest, and the only average score of less than half, was in College Algebra with 38.8%.

Table 2. Means and Standard Deviations of Core Components and Grades by Course

| Math Course | Number of Students | Total Points | Mean Core | SD Core | Mean Grade | SD Grade | Core % |
|---|---|---|---|---|---|---|---|
| Fundamentals of Math | 223 | 10 | 7.43 | 2.13 | 2.75 | 1.27 | 74.3 |
| Introduction to Algebra | 494 | 10 | 6.58 | 2.53 | 2.45 | 1.31 | 65.8 |
| Intermediate Algebra | 579 | 10 | 6.08 | 2.34 | 2.41 | 1.29 | 60.8 |
| Business Math | 146 | 9 | 6.85 | 1.90 | 2.76 | 1.30 | 76.1 |
| College Algebra | 581 | 17 | 6.59 | 4.38 | 2.39 | 1.21 | 38.8 |
| Trigonometry | 80 | 5 | 2.61 | 1.33 | 2.30 | 1.44 | 52.2 |
| Pre-Calculus | 58 | 9 | 4.50 | 2.26 | 2.81 | 1.18 | 50.0 |
| Statistics | 176 | 16 | 11.89 | 3.15 | 2.95 | 1.08 | 74.3 |
| Calculus I | 78 | 18 | 9.22 | 4.28 | 2.59 | 1.14 | 51.2 |
| Analytic Geometry-Calculus I | 79 | 55 | 33.71 | 12.28 | 2.68 | 1.19 | 61.3 |

Table 3 indicates the Pearson correlations between the core components and the course grades by course. All correlations were positive, most were relatively high, and all were statistically significant at p=0.005 or better. Only in trigonometry, with a maximum of 5 possible points for the core components, was the correlation of the core with the course grade less than 0.60.

However, at the individual instructor level, Intermediate Algebra had five instructors (out of 21 reporting) with correlations below 0.50, including one with a correlation of 0.15; Introduction to

Algebra had two instructors (out of 26 reporting) with correlations below 0.50; and Business Math (nine instructors), Trigonometry (three instructors), Pre-Calculus (five instructors) each had one instructor with a correlation below 0.50. In every case, these correlations below 0.50 were not significant at p=0.005.

Table 3. Pearson Correlations of Core Components with Course Grades by Course

| Math Course | Number of Students | Total Points | Pearson Correlations | *p* value |
|---|---|---|---|---|
| Fundamentals of Math | 223 | 10 | 0.79 | <.001 |
| Introduction to Algebra | 494 | 10 | 0.74 | <.001 |
| Intermediate Algebra | 579 | 10 | 0.64 | <.001 |
| Business Math | 146 | 9 | 0.63 | <.001 |
| College Algebra | 581 | 17 | 0.70 | <.001 |
| Trigonometry | 80 | 5 | 0.58 | <.001 |
| Pre-Calculus | 58 | 9 | 0.61 | <.001 |
| Statistics | 176 | 16 | 0.64 | <.001 |
| Calculus I | 78 | 18 | 0.71 | <.001 |
| Analytic Geometry-Calculus I | 79 | 55 | 0.76 | <.001 |

## Discussion

Overall, it appears that the process of constructing and using core component items on final examinations can, in fact, help produce a balance between common academic outcomes and faculty's academic freedom of instruction. Additionally, the use of fewer common examination items may not necessarily lead to the weakening of standards. During this process, however, several other issues and concerns were raised.

Some difficulties occurred during the development of the core components that resulted in immediate curricular changes. For example, during the college algebra discussions, the only question appearing on the old departmental final exam regarding systems of equations was actually covered in a prerequisite course. Substantial discussion was held as to whether it was appropriate to include what was essentially a "review" question as part of the core, whether the "new" material in the chapter was sufficiently appropriate and significant to test on the final exam, or whether the chapter was even necessary as part of the course. The result was to include the "new" material, and several instructors commented that they would need to spend more time on this topic.

Since there was contention among faculty about what constitutes an appropriate pedagogical approach to teaching in College Algebra, faculty expected that its correlation would be among the lowest. Conversely, because of the agreement among Business Math faculty, it was expected that the correlation for this course would be fairly high. In both cases, faculty expectations were not validated by the data. Correlations for College Algebra, both by course and instructor, were in the middle of the group. The course correlation for Business Math was in the lower half of the ten courses, with several instructor correlations failing to reach significance.

Several faculty expressed concern about the ability of the design of the core components to meet the four simultaneous objectives of the final exam. Ideally, to obtain ideal measures for all four objectives, one might use four separate exams. However, class time and student patience are finite, and this design seems to meet the basic requirements of all four objectives. The core provides a common measure of student achievement for the department and for Carl Perkins funding, as well as an administrative control, while the free portion provides individual instructor measures of student achievement.

Since a portion of the course grade was computed by using the core questions (although the grading procedure on those same questions may have differed), the analysis may be biased toward producing higher correlations. One way to approach this problem is to mathematically remove the effect of the core questions on the course grade (see technical note). An examination of all spring 1995 mathematics course syllabi provided values for the percentage of the course grade determined by the core score (denoted as $\alpha$) for each section of the ten courses (Table 4). In no case did any correlation coefficient drop below 0.50, nor was the significance of any of the correlations affected.

Table 4. Original and Unbiased Correlation Coefficients by Course

| Math Course | Original Correlation Coefficient | Unbiased Correlation Coefficients ... using... | |
|---|---|---|---|
| | | average value of $\alpha$ | maximum value of $\alpha$ |
| Fundamentals of Math | 0.79 | 0.77 | 0.76 |
| Introduction to Algebra | 0.74 | 0.71 | 0.70 |
| Intermediate Algebra | 0.64 | 0.61 | 0.60 |
| Business Math | 0.63 | 0.60 | 0.60 |
| College Algebra | 0.70 | 0.64 | 0.60 |
| Trigonometry | 0.58 | 0.54 | 0.53 |
| Pre-Calculus | 0.61 | 0.57 | 0.55 |
| Statistics | 0.64 | 0.59 | 0.59 |
| Calculus I | 0.71 | 0.67 | 0.66 |
| Analytic Geom.-Calculus I | 0.76 | 0.71 | 0 71 |

The mathematical approach was successful in removing one of the two possible sources of instructor-induced bias in our study, namely that due to the final exam being a percentage of the final grade in each section. The bias caused by contingency clauses, through which instructors replace low or missing scores with the final exam score, could not be accounted for here, since the replacement could potentially be different for every student enrolled in the ten sections, and individual student data was not collected. Based on examination of the course syllabi, replacement policies ran from a single test to 100% of a student's grade being replaced with the final exam score.

An additional question concerning grade inflation was raised. In order to try to identify whether grade inflation was present in any of the courses, mean grades were compared to the core scores and regression equations were calculated (see Table 5). The regression equations indicate that, on the average, students who earned no points from the core items on either College Algebra and Pre-Calculus final examinations would still pass the course with D grades.

## Implications

Currently, the project information and results are being shared and discussed with the mathematics faculty. Process problems, indicated by lack of agreement between instructors (in Intermediate Algebra, for example) should be addressed and solutions implemented by the mathematics faculty, prior to the implementation of required core final examination items. Course problems, indicated by low core component means (in College Algebra, for example), will also need to be addressed by the instructors in those courses.

Table 5. Regression Equations and Means of Core Components and Grades by Math Course

| Math Course | n | Total Points | Mean Core | Mean Grade | Regression Equation |
|---|---|---|---|---|---|
| Fundamentals of Math | 223 | 10 | 7.43 | 2.75 | 0.47x - 0.74 |
| Introduction to Algebra | 494 | 10 | 6.58 | 2.45 | 0.38x - 0.08 |
| Intermediate Algebra | 579 | 10 | 6.08 | 2.41 | 0.36x + 0.24 |
| Business Math | 146 | 9 | 6.85 | 2.76 | 0.43x - 0.20 |
| College Algebra | 581 | 17 | 6.59 | 2.39 | 0.19x + 1.11 |
| Trigonometry | 80 | 5 | 2.61 | 2.30 | 0.63x + 0.66 |
| Pre-Calculus | 58 | 9 | 4.5 | 2.81 | 0.32x + 1.39 |
| Statistics | 176 | 16 | 11.89 | 2.95 | 0.22x + 0.36 |
| Calculus I | 78 | 18 | 9.22 | 2.59 | 0.19x + 0.85 |
| Analytic Geometry-Calculus I | 79 | 55 | 33.71 | 2.68 | 0.07x + 0.20 |

In terms of statistical implications, if the project is replicated, the data should be collected in a manner which will allow the core component scores and the final course grades to be independent. It would also be advisable to ensure that the impact of the core final exam items be the same on all students' course grades. Contingency clauses, allowing students to accumulate enough points during the term so that the final exam is unnecessary, also affect the relationship between the core scores and the final course grade and need to be taken into account.

One of the benefits of this study to the institutional researcher is that it affords the possibility to conduct outcomes assessment and research across courses while maintaining the academic freedom of the individual faculty member. The relationship between final grades and the common course final also provides for the possibility of examining other curricular differences across sections of a course, e.g., whether the use of scientific calculators makes a difference in the success of students or whether a three-credit-hour course offered in a five-contact-hour format for underprepared students is beneficial.

In terms of broad educational issues, the results of this research suggests that a balance is possible between academic freedom and the commonality necessary for assessment--it is not an either/or situation. Although agreement did not occur in every instance, this project also indicates that faculty consensus is possible on these issues.

## Technical Note

We can mathematically adjust the correlation coefficient for the bias caused by the effect of the core questions on the course grade. Let us assign the variable x to the core score data, and the variable z to the course grade data. From this data, we originally computed a correlation coefficient $r_z$ and a regression equation $z = a_z + b_z x$. Now z is not independent of x, so the coefficient $r_z$ is biased. Let us assign the variable y to the pre-core grade of the student, and assume that the relationship between the three variables is given by $z = \alpha x + \beta y$. The (x,y) data will not contain the bias we introduced earlier. We need to find an expression for the unbiased correlation coefficient $r_z$ in terms of the (x,z) data. Let the coefficient $\alpha$ be the percentage of the course grade which was determined by the core (adjusted for the different scales of measurement used in the x and z data) and $\beta$ be the percentage of the course grade determined by pre-core student work. Now we use definitions and properties of statistics (from Weiss, pp. 289 and 787), and employ some algebra. From the definitions of the correlation coefficient $r_y = \dfrac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$ and the regression coefficient $b_y = \dfrac{S_{xy}}{S_{xx}}$, we

can obtain the relation $r_y = \dfrac{b_y}{b_z}\dfrac{s_z}{s_y} r_z$. Using the definitions of $S_{xy}$, and $S_{xx}$ in the definition of $b_y$,

we can substitute the equation $z = \alpha x + \beta y$ and obtain the relation $b_z = \alpha + \beta\, b_y$, which leads to the ratio $\dfrac{b_y}{b_z} = \dfrac{1}{\beta}\left(1 - \dfrac{\alpha}{b_z}\right)$. Finally, we can show that the standard deviation properties

$\sigma_{x+y} = \sqrt{\sigma_x^2 + \sigma_y^2}$ and $\sigma_{cw} = |c|\sigma_2 w$ also hold true for the sample formulas as well as the population formulas, and then show that $\frac{s_y}{s_z} = \frac{1}{\beta}\sqrt{1 - \alpha^2 \frac{r_z^{2^2}}{b_z^2}}$. This leads to our final result, the unbiased correlation coefficient $r_y = \frac{r_z(b_z - \alpha)}{\sqrt{b_z^2 - \alpha^2 r_z^2}}$. It is probably noteworthy that this formula does not directly depend on the sample size n.

## References

Cohen, D. (Ed.). (1995). *Crossroads in Mathematics: Standards for Introductory College Mathematics Before Calculus*. Memphis, TN: American Mathematical Association of Two-Year Colleges.

Commission on Standards for School Mathematics. (1989). *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Milton, O., Pollio, H.R., and Eison, J.A. (1986). *Making Sense of College Grades*. San Francisco, CA: Jossey-Bass.

Olson, George H. (1989). *On the Validity of Performance Grades: The Relationship between Teacher-Assigned Grades and Standard Measures of Subject Matter Acquisition*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco.

Weiss, Neil A. (1995). *Introductory Statistics* (4th edition). Reading, MA: Addison Wesley Publishing Company, pp. 289, 787.