Promoting Gender Equity by Changing Math Tests – Does it Really Work?

Dorit Neria, Miriam Amit
Ben Gurion University of the Negev
Center for Science and Technology Education

Abstract

The gender gap, favoring boys, in mathematics education is commonly perceived the world over. One proposed intervention, designed to allow girls to use their abilities and to diminish this gap, is to change the structure of math tests. Since, in general, girls have better verbal skills than boys, a written verbal element which might raise the overall test scores of the girls could be included. We analyzed the answers of 81 girls and 83 boys from the 9th grade, to three test tasks that provided an opportunity to demonstrate mathematical justifications in verbal modes. Results did not indicate gender differences in achievement, nor in girls' preference to justify by verbal modes. These findings do not support the assumption that the addition of verbal components to math tests will raise girls' scores, thus helping to eradicate the perceived gender gap in mathematics. Several factors, such as the influence of classroom practice and pressure of time in exams, remain to be studied.

In the last two decades, there has been a trend to make mathematics accessible to all, since mathematics is considered as a social gate-keeper for advanced education in science and technological fields, and the participation of individuals in these fields enables them to improve their economic status (Galbraith, 1993). In their book "*Radical Equations – Math Literacy and Civil Rights*", Moses and Cobb (2001) even compare the call for proper mathematics education today, to the struggle for equal rights in the US in the 1960s. Moses and Cobb refer mainly to ethnic minorities, but there is a large "minority" which is yet to enjoy full equality of incorporation in highly mathematical areas – females, who are under-represented in these fields (Leder, 1992; OECD, 2001).

Theoretical background

Gender differences in achievement in mathematics in favor of boys have been found in standardized tests and are most prominent at the very high levels of achievement (Leder, 1992; Mullis et al., 2000; OECD, 2001, 2004).

These differences are likely to be both content and ability dependant. While males outperform females in scientific and mathematical tasks, females outperform males in tasks involving verbal abilities (i.e. Fennema, et al., 1998; Leder, 1992; Mullis et al., 2000; Nowell & Hedges, 1998; OCED, 2004; O'Neill & McPeek, 1993; Ryan, 2001).

Gender differences in test results are also related to the nature of the test items. Females perform better than males on constructed-response items, open-ended questions, and items resembling textbook or homework problems, while males perform better on highly "objective" tests that include multiple-choice items and items that are non-conventional (i.e. Bolger & Kellaghan, 1990; Lane, Wang & Magone, 1996; O'Neill & McPeek, 1993; Pomplun & Capps, 1999).

The differences mentioned above have led policy makers and test designers to add gender oriented elements to tests, specifically – a written verbal component, in order to diminish the differences (Willingham & Cole, 1997).

The aims of this study were to examine the relationships between mathematical communication and gender, in the context of a standardized regional test. We hypothesized that females, once given the opportunity, would have a greater tendency than males to use written verbal answers as a means of communication, and that the quality of their mathematical communication would be higher.

Methodology

The purpose of the study: to examine gender differences in achievement, in representation mode and in the quality of mathematical communication and justification, in a problem solving context.

Settings and instruments: The population comprised 164 ninth grade students (83 male and 81 female) who participated in a regional mathematics test, from five multi-ability schools. Three problems requiring a fully justified answer were selected from this exam for the purposes of the study. Problem 1 dealt with a non-routine optimization problem, requiring the students to choose between two telephone companies and justify their preference in any mode, including formulae and graphs.

Problem 2 dealt with a textbook-like rate of change problem: water drained from a pool at a given constant rate. The students were asked whether the pool would be completely empty after one hour and to justify their conclusions.

Problem 3 was a multiple-choice non routine problem that dealt with the relation between the area and perimeter of a rectangle.

Method of analysis: Each answer was analyzed qualitatively according to three criteria: *correctness:* correct, wrong, or no answer; *representation mode:* verbal, numerical, algebraic, or graphic; and the *quality of justification***:** good, medium, or poor (Cai, Jakabcsin, & Lane 1996).

The results of the qualitative analysis were quantified and statistical methods were applied in order to check gender differences.

Results

Correctness of answers: No significant gender differences were found. **Representation mode**: No significant gender differences were found for Problems 1 and 2. For problem 3 statistical tests could not be applies due to low cell counts, although a similar distribution of representation modes was found.

Quality of justification: Significant gender differences were found in problem 2 favoring girls whose justifications were of good quality. No significant differences were found in Problems 1 and 3.

Discussion

Our assumption that implementing a verbal component in a regional mathematics test will be in favor of girls was not fully fulfilled.

In contrast to our hypothesis, the study did not find gender differences in the ***representation modes***, nor even a preference for verbal modes in the girls' answers, rather than graphic or symbolic. Why, then, was the girls' verbal advantage not reflected in the mathematics test?

One explanation is that mathematics is perceived as a merely numerical domain dealing with symbols and numbers, and that this approach is likely to find expression in the teachers' classroom practices (Amit & Hilman, 1999). Such a perception might prevent girls from applying their verbal ability by using verbal modes of representation in a mathematical context.

Another explanation is that girls do not fully make use of their verbal potential, because of mathematical anxiety and stress during math tests. It is known that females suffer from "mathematical anxiety" much more than males (Leder, 1992). This phenomenon might prevent females from utilizing their full potential when tackling problems and specifically, might prevent them from utilizing their verbal potential.

There is a possibility that girls' verbal advantage is not reflected in mathematics tests because of its context and structure. First, the female advantage in verbal ability has been characterized in research work that deal with the Humanities (History, Philosophy, Literature, etc.) and in test items that are specifically defined as testing verbal abilities. In our study, the context was a scientific one, a context that has been found to be problematic for females (O'Neill & McPeek, 1993).

In ***quality of justification*** gender differences were found in one of the three tasks that are the focus of this study. Only in Problem 2 did the girls surpass the boys in the quality of justification. This result might be related to the fully open-ended structure of this task, unlike the two others where there was a multiple-choice element. In this problem (Problem 2), a right solution implies a fully elaborated explanation. The literature suggests that girls tend to elaborate their work, more than boys, to explain their processes and to write explicit conclusions (Lane et al., 1996, Pomplun & Capps, 1999), and these are exactly the criteria for high quality justifications (Cai, Jakabcsin & Lane, 1996).

Additional questions and closing remarks

The current study seems to have raised several questions: Does the fact that verbal answers take more time to formulate and write down than do multiple choice answers inhibit the girls from expressing their answers fully? How influential are classroom practices on students' tendency to use verbal (or other) abilities? This leads us to the gender question: How far do current examination styles reflect the *real* mathematical abilities of students?

Today, there is a tendency to try to help girls close the gap in mathematical achievement by incorporating verbal elements in exams due to their advantage in verbal ability (Willingham & Cole, 1997). Our study provides no support for this approach. On the contrary, we suggest that the advantage that females have in verbal ability does not apply in a typical mathematics test situation. Therefore, before policy makers and test designers further adopt such a solution, especially in high stake assessment, a number of cultural and educational factors need to be addressed and additional questions need to be answered.

References

Amit, M., & Hillman, S. (1999). In B. Jaworski, T. Wood, and S. Dawson (Eds.), *Mathematics Teacher Education: Critical International Perspectives.* Philadelphia, PA. Falmer Press.

Bolger N., & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement, 27,* 165-174.

Cai, J., Jakabcsin, M.S., & Lane, S. (1996). Assessing student's mathematical communication. *School Science and Mathematics, 96,* 238-246.

Fennema, E., Carpenter, T.P., Jacobs, V.R., Franke, M.L., & Levi, L.W. (1998). New perspectives on gender differences in mathematics: A reprise. *Educational Researcher, 27,* 19-21.

Galbraith, P. (1993). Paradigms, problems, and assessment: Some ideological implications. In M. Niss (Ed.), *Investigations into Assessment in Mathematics Education* (p. 73-86). Kluwer: Dordrecht.

Lane, S., Wang, N., & Magone, M. (1996). Gender-related differential item functioning on a middle-school mathematics performance assessment. *Educational Measurement: Issues and Practices, 15,* 21-27.

Leder, G. C. (1992). Mathematics and gender: changing perspectives. In D.A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 597-622). New York: Macmillan.

Moses, R.P., & Cobb, C.E. (2001). *Radical Equations - Math Literacy and Civil Rights.* Boston: Beacon Press.

Mullis, V.S., Martin, M.O, Fierros, E.E., Goldberg, A.L. & Stemler, S.E. (2000). *Gender Differences in Achievement: IEA's Third International Mathematics and Science Study (TIMSS).* Retrieved September 24, 2002, from TIMSS1999 Web site: http://isc.bc.edu/timss1999.html

Nowell, A., & Hedges L.V. (1998). Trends in gender differences in academic achievement from 1960 to 1994: An analysis of differences in mean, variance and extreme scores. *Sex Roles, 39,* 21-43.

OCED (2001). *Knowledge and skills for life. First results from the OECD programme for international student assessment (PISA) 2000.* Retrieved March 31, 2005, from OCED Web site: http://www.pisa.oecd.org

OCED (2004). *Problem solving for tomorrow's world - first measures of cross curricular competencies from PISA 2003.* Retrieved March 31, 2005, from OCED Web site: http://www.pisa.oecd.org

O'Neill, K.A., & McPeek, W.M. (1993). Item and test characteristics that are associated with differential item functioning. In P.W. Holland and H. Wainer (Eds.), *Differential Item Functioning* (pp. 255-276). Hillsdale, NJ: Lawrence Erlbaum.

Pomplun, M., & Capps, L. (1999). Gender differences for constructed-response mathematics Items. *Educational and Psychological Measurement, 59,* 597-614.

Ryan, K.E. (2001). An examination of item context effects, DIF, and gender DIF. *Applied Measurement in Education, 14,* 73-90.

Willingham, W.W., & Cole, N. (1997). *Gender and Fair Assessment.* Hillsdale, NJ: Lawrence Erlbaum Associates.