משרד החינוך
Ministry of Education
وزارة التربية والتعليم

RAMA
National Authority for Measurement and Evaluation in Education

Gender Gaps in Mathematics and Language in Israel–

What Can Be Learned From the Israeli Case?

Joel Rapp

May 2015

בית אבגד, ז'בוטינסקי 5, רמת גן, 5252006 • טלפון: 03-5205555, פקס: 03-5205509
דוא"ל: rama@education.gov.il, אתר: http://rama.education.gov.il

I would like to thank my colleagues at the National Authority for Measurement and Evaluation in Education (RAMA) for their assistance in the following areas:

Reading and comments:

Einat Notea-Koren, Director of National Assessments
Inbal Ron-Kaplan, Director of International Assessment
Iman Awadie, Director of Assessment in Arabic
Aner Rogel, Associate Director of National Assessments

Writing, data collection and diagrams:

Tal Freeman, RAMA Consultant

Editing (in the original Hebrew report):

Nili Eden, RAMA Consultant

Layout and graphics:

Liron shitrit, Keren Dangor

Translation (from Hebrew):
Julie Rosenzweig, Sagir International Translations LTD

Abstract

There is a widespread assumption that boys are better than girls at mathematics, and that girls' language skills surpass those of boys. Gender gaps in school achievements are thought to have great theoretical and practical importance, given that scholastic performance in general, and math achievements in particular, have implications for later success in life. The recent literature, however, casts doubt on the very existence of a gender gap in math at school, and emphasizes the great heterogeneity that characterizes these gaps (e.g. reading). By contrast, girls' superiority to boys in native language skills and reading proficiencies has been consistently documented. While this language advantage of girls is usually attributed to innate biological factors, scholars are divided regarding the origins of the math gender gap in favor of boys (if, at all, such a gap exists). Many researchers believe that the frequently-documented superiority of boys in this domain stems from socio-environmental influences and from socialization processes that perpetuate and maintain this prevalent stereotype.

The current study presents the overall picture of the gender gap in mathematics and in native-language/reading domains in Israel schools over a period of several years, documented in large-scale assessment systems that are administered in the Israeli education system (the national GEMS exams and the international PISA, TIMSS and PIRLS studies). The picture is presented for the schools of the two main language sectors in Israel – Arabic- and Hebrew-speakers. The gender gap in Israel is also compared to that of other countries and to international norms. The study analyzes the gender gaps in the two school domains both separately and jointly.

In both language sectors, native language skills (reading) were found to be the subject area in which girls have the greatest advantage over boys, while for math a different picture emerges: in Hebrew-speaking schools boys tend to perform better than girls, while in Arabic-speaking schools girls tended to outperform boys. The size and direction of the gaps in the Hebrew-language schools resemble those of Western countries, while the situation in the Arabic-language schools is comparable to that of Arab countries. However, a simultaneous analysis of the gender gaps in both disciplines pointed to several features that are common to both Israeli sectors, as well as to most of the countries that participate in the international studies: math is almost always the subject in which the "situation" of boys is better than that of girls, i.e., the gender gap in favor of boys is the largest (or the smallest, if the gap happens to be in favor of girls), compared with the other tested school subjects. Another finding is that a correlation exists between the size of the gender gaps in math and in native-language skills (reading), and that there is a correlation between achievements in the two disciplines. When accounting for the influence of the level of reading proficiency on math performance and compares the math achievements of boys and girls whose reading skills are the same, boys always outperform girls. The consistency of this finding stands in contrast to the great heterogeneity of gender-gap sizes overall, which supporters of the social-environmental approach view use to maintain the idea that boys have no innate advantage in mathematics.

Table of Contents

List of Figures

Gender Gaps in Mathematics and Language in Israel – What Can Be Learned From the Israeli Case?

Introduction

In recent decades, researchers in the fields of psychology, psychobiology, psychometrics, education and economics, and the public in general have displayed a growing interest in the question of whether differences exist between boys and girls in the disciplines of mathematics and language. If such disparities do exist, one may then ask: what causes them and what can be done to narrow them? This study examines Israeli data and the degree to which they correspond to the overall picture and trends that emerge from the research literature; it also attempts to generalize from the Israeli case to international findings and to the issues that concern professionals in the field.

Math achievement disparities between boys and girls

A common convention within the general public and in education systems around the world is that boys are achieving better than girls in mathematics (e.g., Else-Quest, Hyde & Linn, 2010). One way of investigating the issue is to examine the results of large-scale tests administered in education systems – tests results that are supposed to represent the entire student population. Most of the research literature has focused on test results administered in the United States (whether at the nationwide level, e.g., the NAEP exams, or at the individual-state level), or on tests administered as part of international studies in which numerous countries participate (a representative sample within each country).

The question of gender gaps in school achievement in mathematics is of great theoretical and practical importance. Not only do disparities in mathematical ability between school-aged boys and girls reflect the current gap in terms of school performance, but they may also predict girls' and boys' occupational development later in life. Students' performance in math may affect their future selection of study discipline at university and, consequently, the field in which they will work as adults (Hyde, Fenemma & Lamon, 1990). These selections have economic consequences, as they eventually become visible in income gaps between men and women and affect women's chances of occupying key social and economic positions in developed countries. As early as the 1970s Lucy Sells described mathematics as a critical filter that prevents women from the more prestigious and better-paid jobs (Sells, 1973). It has also become clear that women in developed countries still choose STEM (Science, Technology, Engineering and

Mathematics) professions to a lesser degree than men (del Pero & Bytchkova, 2013). It also turns out that the proportion of women who perform at a superior level in math and who succeed in quantitative fields is still much lower than the proportion of men (Lindberg, Hyde, Peterson & Linn, 2010). Thus, the existence of gender gaps in math achievement has an impact in terms of equal opportunity between men and women and, consequently, is of importance to those who shape and formulate education policy around the world. It also has bearing on the way in which a given country's education system must act if it wants to advance social and economic gender equality.

One of the main points of controversy regarding the math gender gap is whether boys are really better than girls in this domain, and if so, what is the origin of this disparity. The dispute basically revolves around the two classic approaches of "nature" versus "nurture" (Eagly & Wood, 2013): the former holds that boys are naturally better than girls at math, for reasons that are innate-biological, and that their advantage in this domain is already manifested by their performance on large-scale standardized tests administered at schools. This approach relies on the documentation of lifelong disparities in favor of boys. For example, Fryer & Levitt (2010) demonstrated the existence in the United States of a boy-girl gap in math averaging one-fifth of a standard deviation in favor of boys as early as the end of Grade 5. There is also a gender gap in quantitative SAT scores (Brody & Mills, 2005). Additional, repeatedly-demonstrated evidence includes boys' larger share of scores at the upper end of the distribution of performance math and in science on large-scale tests the larger percentage of boys in math competitions (in the United States and around the world); the larger percentage of men among outstanding performers in math-intensive fields at all ages (see Ceci &Williams, 2010 and Ellison & Swason, 2010). The percentage of men studying STEM fields is higher, and men have a greater tendency to choose scientific-quantitative occupations (Pero & Bytchkova, 2013 and others). Studies also point to gender differences on the affective plan: girls suffer more math anxiety than do boys (see results of the 2012 PISA study[1] in OECD 2013; Goetz, Bieg, Lüdtke, Pekrun & Hall, 2013; Else-Quest et al., 2010; Birenbaum & Nasser, 1994); boys, compared with girls, have more confidence in their mathematical ability (Preckel, Goetz, Pekrun & Kleine, 2008); and there is an assembly of other affective-motivational findings which, though they do not prove the existence of a biologically-based advantage for boys, are nevertheless consistent with this approach.

It has been argued repeatedly, with regard to intra-gender performance differences, that there is a recurring pattern across different fields within

[1]PISA (the Program for International Student Assessment) is an international education survey that is administered once every three years by the Organization for Economic Co-operation and Development (OECD). The study assesses students' competencies at age 15 in reading, mathematics and science.

mathematics. For example, girls consistently perform better in arithmetic than in geometry (Guiso, Monte, Sapienza & Zingales, 2008), while boys tend to be better in spatial perception tasks (Gallagher, De Lisi, Holst, McGillicuddy-De Lisi, Morely & Calahan, 2000). Boys' relative success in math is often attributed to superior spatial abilities, whether rooted in evolutionary development or in their greater tendency to engage in activities and games that involve movement in space (see review by Geary, 1996; 2010, covering biological-evolutionary models explaining boys' superiority to girls in math generally, and in space relations particularly; also Berenbaum, Martin, Hanish, Briggs & Fabes, 2008). One way or another, this fixed pattern suggests that boys' advantage is based on gender differences in different parts of the brain (Baron-Cohen, 2003; Kimura, 1999). Nevertheless, attempts to identify such differences in spatial perception and cognitive development in terms of the functioning of brain regions linked to mathematical ability have been unsuccessful (Wilder & Spelke, 2005; Powell, 1989). Ultimately, it is hard to assess the degree to which gender gaps in math have a biological basis, since experience also affects brain structure and cognitive functions (Halpern, Benbow, Geary, Gur, Hyde & Gernsbacher, 2007).

A different approach, the environmental approach, holds that boys and girls are born with identical intellectual-mathematical potential (Spelke, 2005), and that any existing disparities are the result of sociocultural influences in the form of education, attitudes, expectations and messages conveyed by society in general and especially by parents and teachers. Recent studies indicate that the expectations a given society transmits to girls, and the extent to which girls are encouraged to select and excel at mathematical domains, correspond to the country's/culture's degree of gender equity. The struggle to gender equality and the investment in closing gaps between girls and boys have an impact and ultimately determine whether women will succeed in quantitative fields generally, and in math particularly (see Else-Quest et al., 2010 and Nosek, et al., 2009). Mosconi's research (e.g., Mosconi, 2001) shows how instructional practices applied within the classroom can contribute to the development of gender differences in math and embed them permanently in students' consciousness and in reality. Mosconi observed teachers in the classroom and demonstrated differences in the way both male and female teachers addressed, questioned and provided feedback to boys and girls. According to Mosconi, these differences stem from deeply-rooted stereotypes that affect the learning and achievements of boys and girls in the school retaining and preserving the gaps between them in math – in the manner of self-fulfilling prophesies. Not only that, but those who support the nurture approach argue that the stereotype according to which boys are better than girls at math has no actual empirical support (e.g., Hyde, Lindberg, Linn, Ellis & Williams, 2008). For the past several decades the research literature has been presenting data from large-scale American and international standardized tests which show, on average, no math disparities between boys and girls, with a trend toward the closing of the gaps that do exist. Meta-analyses of hundred large-scale studies and tests (e.g. the American NAEP) spanning the period 1973-1990 (Hyde et al., 1990) and 242 studies and tests from the period 1990-2000 (Lindberg et al., 2010) reveal the mean gap across all of the studies to be insignificant – .05 standard deviation in favor of girls.[2] Moreover, studies have documented a trend toward reduced gender gaps over the years in the United States (Hyde et al., 2008), in terms of mean achievements and a number of parameters relating to the percentage of high performers. It was found, for example, that the ratio of female to male test takers earning scores of over 700 on the math section of the SAT rose from 1:13 in 1983 to 1:3 (Brody & Mills, 2005), and there has also been a steady increase in the share of women among those completing doctoral degrees in mathematics in the United States (Burrelli, 2008).

This body of research notwithstanding, the stereotype of boys outperforming girls seems to be resistant (Else-Quest et al., 2010). People still continue to see boys as better than girls at math, and the matter continues to preoccupy both researchers and the mass media. Not only that: because the issue is linked to the broader and more politically sensitive topic of gender equality, it generates considerable interest. The approach that favors biological-innate models is regarded as unenlightened and inequitable toward women, as it appears to justify the gender inequality that prevails in Western society. The issue's sensitivity can be demonstrated by the resignation of Harvard University's then-president, Lawrence Summers in 2005, following a speech in which he conveyed his belief in biological reasons for the gender gap favoring males in quantitative-scientific disciplines.[3]

Heterogeneity of the gender gap

---

[2]"Gender gap" in these studies is defined as "the difference between the mean achievements of boys and girls in terms of standard deviations."

[3]In a speech, Summers suggested that women are underrepresented in key positions at science and engineering institutions due to innate differences. He explained afterward that he had been referring to the idea that there are "differences in variability" between male and female populations. But the mere reference to the idea that there are biological reasons for a given situation of women's underrepresentation in prestigious senior posts at academic institutions was felt to be unenlightened. Summers' statements generated fierce public debate, in the wake of which he was forced to resign. For more information on the affair, see:
http://www.washingtonpost.com/wpdyn/content/article/2006/02/21/AR2006022101842.html
The biological model to which Summers referred is the "greater male variability" hypothesis first raised by Ellis in 1894. For more on this hypothesis, see Machin & Pekkarinen, 2008 and, in the context of scholastic achievements in Israel, see in Rapp, Notea-Koren, Ron-Kaplan, Gelbart, Awadyeh, Rogel, 2013.

One finding that supporters of the environmental approach use to negate the biological basis for the math gender gap is the great variability of the gap magnitude. This has been noted in the various relevant studies, across age groups and time periods (there is, again, evidence from the United States that boys' advantage over girls has been dwindling in recent years); not only that, but the math gender-gap picture across different countries is diverse and inconsistent. For example, Kane & Mertz (2012), using TIMSS data for 2003 and 2007[4]and data from PISA 2009, examined math-achievement gender disparities in international studies conducted in 86 countries, for different age and school grade levels. They argued that, if boys' math advantage is indeed biological/innate, one would expect to find consistent and similarly-sized disparities favoring boys on the math tests of the various assessments and across the different countries and cultures. In their view, the fact that no such consistency was found, and that the size of the different countries' gender gaps in math varied greatly, suggests that the disparities are more culture-dependent than biology-driven.

Social models of math achievement

In order to establish that the socio-environmental explanation is viable and show how environment produces disparities, we must do more than merely note the existence of gender-gap size variability among countries and cultures. We must also successfully explain this variance in terms of the countries' socio-cultural characteristics. A review of the literature in this area highlights several attempts to link social variables with gender-gap size. The "gap due to inequality" hypothesis, sometimes referred to as the "gender stratification" hypothesis, links equality of opportunity between men and women in society as a whole with girls' math achievements. According to this theoretical model, the math disparity to girls' disadvantage should be smaller in countries where women are equitably integrated in the society and the economy. Equal opportunity between the sexes indirectly affects girls' math achievements – due, apparently, to societal attitudes and expectations transmitted to girls in more equitable societies. For example, Guiso et al. (2008) demonstrated a negative correlation between boys' superior performance on math achievement tests and the "Gender Gap Index (GGI)[5] – the

---

[4]TIMSS (Trends in Mathematics and Science Study) is an international educational study conducted every four years by the International Association for the Evaluation of Educational Achievement (IEA). It assesses the mathematics and science knowledge of students in grades 4 and 8.

[5]Index from Hausmann, Tyson & Zahidi (2006). The Global Gender Gap Report is an international report on disparities between women and men (The Global Gender Gap Report, World Economic Forum). The GGI reflects the level of opportunity available to women in the participating countries on the economic, political, educational and health/survival planes. It was developed in response to a need for a consistent, stable and understandable measure of gender equality that would facilitate cross-national

more equitable the society, the smaller the gap. Kane & Mertz (2012), who assessed the validity of the numerous theoretical models that seek to explain gender-gap size in terms of socio-environmental variables, did succeed in reproducing this finding, but only in part. They found that sometimes there is no such correlation, and that sometimes the correlation is in the opposite direction (as in the Grade 4 TIMSS, 2003). This correlation reversal was consistent with an earlier study by Fryer & Levitt (2010), who found a reverse correlation (from negative to positive) between gender-equality parameters and math gender gap size. The direction reversal happens when countries where women's social status is low (low GGI) are included, and where there is no math gap in favor of boys or a gap exists in favor of girls. These countries are, for the most part, Muslim and/or Arab.

In their article Kane & Mertz (2012) showed that in Arab and Muslim countries the size of the gender gap (in favor of girls) is due mainly to the boys' poor performance, that is, the reason for the gap is that the boys tend to be weak in math, rather than the girls tend to be good at math. This raises the question of what might explain boys' weakness in math in these countries. Kane & Mertz succeeded in ruling out a number of variables that other researchers had proposed as explanations. Among the variables they disqualified are country economic status, the share of Muslims in the total population (an explanation that assumes the existence of some feature of Muslim culture that indirectly results in boys' low achievements), and single-sex education, i.e., separate schools (or classes) for boys and girls, as is common in many Arab countries.

Boy-girl differences in terms of language

At a time when considerable scholarly attention was being devoted to boy-girl differences in math and when gender gaps in that domain were generating public interest and debate, research on gender gaps in language (reading, writing, reading comprehension and other language skills) received much less attention and had little resonance. This relative lack of interest may have been due to the fact that gender gaps in this domain are consistently in favor of girls (as discussed below), and that in the relevant psychological-educational literature there is a consensus that girls are better than boys on all language aspects. When disparities favor girls they are, of course, less likely to evoke sensitive issues of gender discrimination or gender equality. It may also be that good language skills are not regarded as crucial to success in adult life. According to a review by Niederle &

---

comparison and monitoring of trends over time within the individual countries. It reflects a policymaking concern for just and equitable division of resources between the sexes. The higher the index score, the greater the degree of equality of opportunity. Other measures of gender equality exist, but their results on all parameters are similar.

Vesterlund (2010), while math test scores are good indicators of future income, this is not the case for language test scores.

Unlike the math gender gap, which varies in size and direction across different studies, tests, age groups, countries, etc., and which some studies have found to be decreasing over time, the body of literature dealing with gender gaps in language skills points to a consistent and stable disparity in favor of girls. This has been found in different languages for which documentation exists, and in different age groups (Cole, 1997). Girls' superiority in language skills is reflected in all of the studies reviewed, including the American NAEP, as well as the international PISA and PIRLS studies.[6]In both of these assessments, girls' reading literacy is superior to that of boys in the vast majority of countries, if not in all of them.

It is conjectured that, compared with boys, girls' verbal skills emerge at earlier life stages, and that their advantage in this area starts from infancy. According to the research literature, girls learn to read more quickly in elementary school. A few studies note at that around age 10 the boys "catch up" to the girls in reading and that there is gender parity at this age. However, at age 11 girls once again show superior language skills. Girls do better at both simple and complex verbal tasks, with a mean disparity in their favor of a quarter of a standard deviation. Girls' advantage increases with age during the high school years and the years following (Maccoby & Jacklin, 1974). A meta-analysis by Leitz (2006) of the gender gap in reading achievements that encompassed the findings of 139 studies spanning the period 1970-2002 (including PISA, PIRLS and NAEP) showed that, on average, girls' achievements at high school age are higher than those of boys by two-tenths of a standard deviation. The American NAEP and the international PISA assessments found even larger gaps. Similarly, a meta-analysis that covered 165 assessments of verbal ability in the US and Canada showed that in 75% of the studies, girls' achievements were superior to those of boys (Hyde & Linn, 1988). Overall, the findings indicate that girls tend to be better than boys in all aspects of language and in the language domain as a whole, as it is studied and assessed in school.

Nevertheless, Hyde & Linn (1988) found that the gap in favor of girls declines over the years. They felt that this finding was consistent with the trend toward reduced gender gaps in math, and posited that, since gender roles have become more flexible, boys are now able to engage in activities that were once regarded as feminine, and that these activities have improved their verbal skills.

---

[6]PIRLS (Progress in International Reading Literacy Study) is an international study that assesses 4th graders' reading comprehension and reading competencies. The study is conducted on behalf of the IEA, which also conducts the TIMSS study.

Explanations of the language gender gap in favor of girls

Because there is virtually no doubt regarding the existence of a gender gap favoring girls in language, the small body of research in this area has dealt with questions such as: In which types of verbal ability is the gap most pronounced? At what ages do the differences appear, and do they disappear at specific ages or developmental stages? Unlike the explanations offered for the math gender gap – an issue mired in the nature-nurture controversy – most attempts to explain the language gender gap have viewed it as a biologically-driven phenomenon. Some explanations have pointed to brain lateralization patterns or larger language centers in female brains (see Harasty, Double, Halliday, Kril & McRitchie, 1977); others have called attention to higher protein levels in the brain regions linked to language (Bowers, Perez-Pouchoulen, Edwards & McCarthy, 2013), etc. However, some researchers, among them Fryer & Levitt (2010), maintain that the attempts to explain disparities in the language domain as functions of biological development have been unsuccessful. We should also note that, although environmental explanations of girls' superior language skills are less prevalent, the literature does furnish a few developmental-environmental explanations which, as with math, view societal expectations and supportive home environments as factors that contribute to the early development of language literacy.

Connection between language and math gender gaps

It is interesting to note that gender gaps in language and math have generally been investigated separately (and in many cases by different groups of researchers). Also, reports on the outcomes of large-scale standardized tests (American and Israeli, among others) and international comparative studies present the findings for each discipline separately (achievements or gaps between different groups), whether in separate reports or in separate sections of the same report, etc. It is clear that neither the research literature nor the standardized test reports pay sufficient attention to the relationship between gender gaps in the different school domains. The fact that the gender gaps have been studied separately is surprising given that both math and language skills are studied by the same students in the same school systems (i.e., in similar educational contexts), and especially considering that in many cases achievements in both subjects are measured via the same testing system.

The tendency to look at the two disciplines separately is also surprising given the existence of a significant positive statistical correlation between achievements in them. According to a review by Chen (2010), there is considerable evidence in the literature that math performance levels are related to language achievements – across grade levels, population groups (boys and girls, differing socioeconomic statuses, etc.). Chen cites Aiken (1971; 1972) and Secada (1992) to the effect that the correlations between math and language achievements range from 0.4 in some studies to 0.86 in others. Positive but lower correlations (from 0.2 to 0.5) are reported by Secada. The differences between the correlations found by the various studies are attributed to differences in the way the studies measure language skills (Chen, 2010). It is thus legitimate to argue that the scholastic performance gaps in the two disciplines should also be looked at jointly. That is, we should be asking: What is the nature of the relationships between the disparate findings for math and language skills and what could be learned from them about the boy-girl gaps in each area? For example, the effect of language achievements on math performance needs to be controlled when investigating the disparities between different population groups (boys-girls, cultural affiliations, etc.).

In recent years studies have been published on achievement gender gaps in a variety of school domains, including mathematics and language. A few of these studies have even inspected the relationships between the math and language gaps and illuminated important aspects of the issue: Voyer & Voyer (2014) performed a meta-analysis of gender differences in a range of disciplines that drew on hundreds of studies in which such disparities in schoolwide scores were reported (i.e., not scores on large-scale standardized tests). Most of the studies are based on data collected from American schools. One of Voyer & Voyer's main findings that transcends age levels, countries, population sectors, etc. is that girls' grades are higher than those of boys. This was true in all school domains studied, including in math.[7] The picture presented by Voyer & Voyer stands in contrast to that of the large-scale tests, on which boys generally outperform girls in math. However, when we look closely at Voyer & Voyer's findings it turns out that girls' greatest advantage lies in the language sphere (native or foreign), followed by the social sciences, then the sciences (physics, chemistry, etc.), and finally mathematics. Voyer & Voyer did not discuss this important rank between school disciplines or try to explain it, as it was not the focal point of their study.

Guiso et al. (2008), whose study focused on the relationship between math gender gaps (in PISA 2003) and cross-national measures of socioeconomic gender equality, noted that there is a correlation between gender gaps in language and in math. They reported that a Pearson's coefficient of 0.59 exists between the gender gaps in these two proficiency areas, across the 40 countries included in their study. In those countries where girls exhibited the greatest superiority to boys in reading, they displayed the smallest degree of inferiority to boys in math (sometimes the gap was in their favor). By contrast, in more "egalitarian" countries where the boy-girl gap in math had "disappeared," that was also accompanied by improved female achievements in reading and by larger girl-to-

[7]For each study they calculated the gap in therms of standard deviation (i.e. the effect size - the statistic d).

boy gaps in that domain. The authors also noted that in all of the countries boys' math skills were higher than their reading skills, and that girls' reading achievements were higher than their achievements in math. Because the authors focused in their study on the relationship between the equality indicators and the size of the gender gaps in math, they did not discuss these important findings or did not try to explain them.

A study by Stoet & Geary (2013) also belongs to this new wave of research that examines gender gaps in language and math simultaneously. They expanded on Guiso et al.'s finding of a correlation between gender gaps in language and math in PISA 2003, affirming the correlation's existence on the basis of four consecutive PISA test cycles over the course of an entire decade: 2000, 2003, 2006 and 2009 (see Figure 1). They also noted the rank that emerges in each of the PISA test years and in (nearly) all participating countries: girls' achievements were higher, on average, than boys' in reading, while in math the reverse was usually true – boys outperformed girls.[8]The correlation documented between the gap sizes in both PISA proficiency areas means that the gaps change in similar directions: the smaller the math advantage for boys, the larger the girls' language advantage. The size of each of these gaps is linked, as shown by Guiso et al. and other researchers (Kane & Mertz, 2012), to the given country's economic level and to its degree of gender equality; however, Stoet & Geary maintain that the direction of this relationship is contrary to that found by earlier investigators with regard to math disparities. Stoet & Geary stressed that one should differ between the overall improvement in girls' achievements (in all subjects) that is linked to the given country's economic and equality levels, and the reduction in gender-gaps. The authors assert that a rise in a country's economic level and/or degree of equality is linked to an overall improvement in school-wide achievements in all subjects (for both genders). But that improvement is actually liable to widen the math gender gap in favor of boys and to reduce girls' advantage over boys in language. They show that this situation characterizes countries that rank high in terms of economic level and degree of gender equality – countries that also exhibit higher scholastic achievements, e.g. the OECD states. It turns out that in these countries gender gaps in math are actually larger than in developing countries. Stoet & Geary thus refute the conclusion reached by their predecessors, namely that a rise in economic level and/or adoption of gender-equality policies necessarily leads to a narrowing of math gender gaps in favor of boys.

Additionally, Stoet & Geary demonstrate that a similar trend exists within countries, that is, between groups of students with differing achievement levels. For example, low-performing students display larger reading gaps in favor of girls

[8]The picture is not entirely reversed, as the gap favoring boys in math was not found in all countries or in all test years, and was smaller than that found in favor of girls in reading.

and negligible math gaps in favor of boys. The higher the achievement level, the smaller the language gap, but the larger the math gap. Ultimately, high-performing students exhibit the smallest language gap in favor of girls, and the largest math gap in favor of boys. There is also a larger share of boys who are outstanding at math (two and half times as many boys as girls) and a much larger (five-fold) proportion of boys who have trouble with reading. Similar findings were also obtained from the PISA 2012 assessment (OECD, 2013). The assessment found that the percentage of "top performer" boys at the higher math proficiency levels (levels 5 and 6) is greater than that of girls; yet at the lower levels (below Level 2) boys and girls are similarly represented or the gender gaps are small.[9] Stoet & Geary call upon the participating countries to pay attention to the differences they found between the two subject domains and to allocate their educational resources so as to advance the desired outcome: an overall improvement in achievements, a reduction of gender gaps in the upper proficiency range or specifically in the low-performing population. The authors conclude that further study is needed in order to fully understand the correlation between the two subject domains, reading and math, and the relationship between the gender gaps in both domains and a given country's economic level and degree of gender equality (two factors that, again, were found to be in congruence).

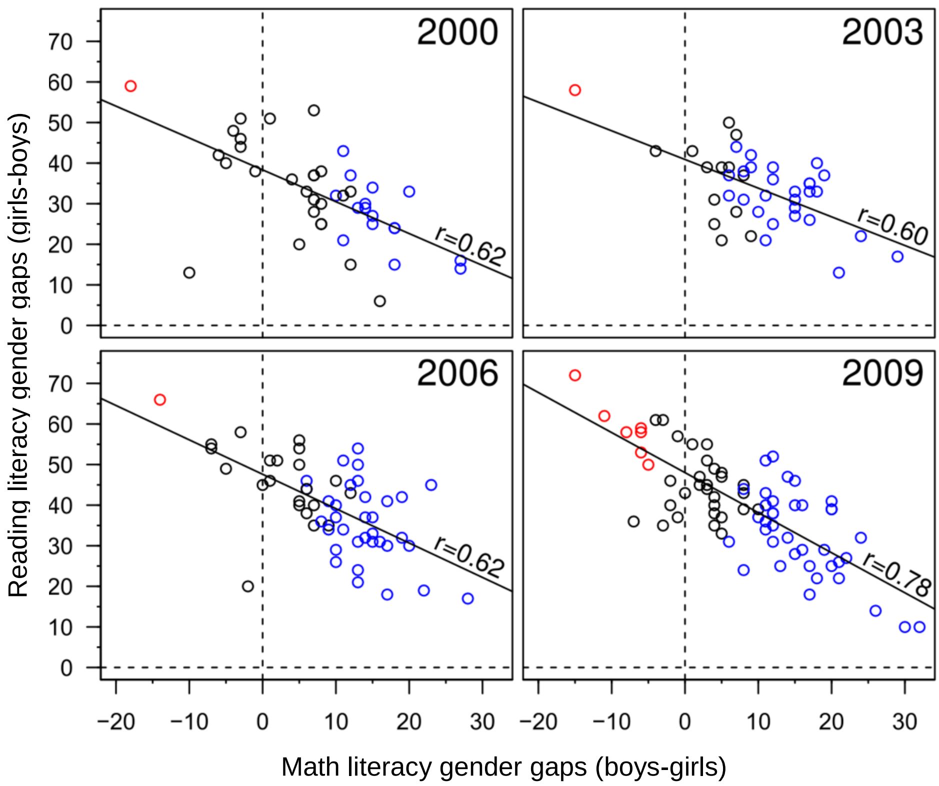
Taking an overall view of the studies that address math and language gender gaps simultaneously, two important issues emerge: A) There is a consistent correlation between the gender gaps in language (reading) and math. B) There is a fixed ranking between these two subject areas: the largest gap in favor of girls is always in the domain of language, while the largest disparity in favor of boys (or, alternatively, the smallest gap in favor of girls) is in math.

---

[9]In Israel, the percentage of "top performer" boys whose math competencies were "outstanding" (levels 5 and 6) on the 2012 PISA was 13.3, while 5.6 percent of girls reached those levels; by contrast, at the lower competency levels (under Level 2) the percentages of boys and girls were 33.6 and 33.4 respectively. For comparison purposes, the mean OECD figure for top performer boys is 14.7% and for top performer girls, 10.6%, while the percentages for boys and girls at the lowest proficiency levels were 22.1 and 23.9, respectively.

Figure 1: Correlation (scatter diagram) of the gender gap in mathematical literacy (boys' mean minus girls' mean) and reading literacy (girls' mean minus boys' mean) in PISA 2000, 2003, 2006 and 2009 (from Stoet & Geary, 2013).

Each point represents the gender gap values for a country



There are two other bodies of scholarly literature where joint consideration of math/language achievements and of the nature of the reciprocal relations between them can be found: applied studies in education and psychometric research. Applied education studies usually aim to shed light on teaching-learning products in the various domain areas in different countries, in order to advance students who have trouble in those domain areas. Teachers with experience in the field report that there is a recognized relationship between native-language studies and math (and other subjects). The prevailing assumption is that difficulties with language skills (listening comprehension, reading comprehension and written expression) are liable to affect a student's ability to learn other subjects and to be tested in them, with obvious implications for scholastic achievement in those subjects (see Estyn, 2008).

The research literature in psychometrics offers a slightly different perspective: it is concerned with the correlation between math and language achievements and with the possibility that this correlation results from math tests that are insufficiently valid. Most researchers have addressed this issue in a context of fairness to immigrant populations lacking mastery of the test language, or to people with learning disabilities who have trouble with language skills. According to the psychometric literature, the correlation between language achievements and math achievements is due to the fact that math tests are inevitably written in *a language*. In effect, they are not able to measure mathematical ability or knowledge in and of themselves. In psychometric terms, the tests “suffer” from low discriminant validity and from construct irrelevant variance[10]. The dependence appears to result from the fact that students who have difficulties with language are liable to be unsuccessful in demonstrating their mathematical knowledge, because they do not properly understand the questions posed, or because they are unable to put their answers into words (see more on this in the Discussion section).

The present work

The present work examines, both separately and jointly, gender gaps in math and language students in the Israeli education system during the past decade, as reflected in large scale standardized (national and international) tests administered in Israel. The study’s initial aim is to delineate the main trends with regard to gender gaps in math and language in Israel, and to determine the degree to which these trends accord with what the research literature has to say. Very few publications have dealt systemically or comprehensively with gender gaps in these two important subject areas in Israeli schools. This is, in effect, the first study to provide an overview of the Israeli education system’s gender gap situation – over time, at several different grade levels, across different types of exam, and based on the achievements of tens of thousands of students.[11]

Israel is a multicultural country whose students come from a range of sectors and cultural-religious-ethnic backgrounds. The main division is into two cultural-linguistic sectors: Hebrew speakers and Arabic speakers. Students who attend Hebrew-speaking schools are the majority group, while Arabic-speaking students are the minority group and represent 20-25% of the education system’s student population. Although both sectors are enrolled in the same school system, most, if not all, attend separate schools different in language of instruction, teachers, and sometimes different textbooks. In addition, significant disparities exist between the two communities in terms of social, cultural and economic status and in terms of scholastic performance in favor of the Hebrew-speaking schools.[12] Comparing

---

[10]See Haladyana & Downing, 2004 on the need for improved test validity.
[11]The number of students who take the national large scale exams (GEMS) in a given year, in a given subject and at a given grade level is around 20,000. The examinees sample size in Israel for international tests is approximately 6,000 in each program.
[12]Achievement gaps between students in the two cultural-linguistic communities, Hebrew speakers and Arabic speakers, range from half a standard deviation to one whole SD, in favor of Hebrew speakers. The gaps decline somewhat after accounting for

these two groups thus offers an opportunity to explore the gender-gap issue in a cross-cultural context but on a small scale. It is in a manner similar to international studies, but in the current case, the two groups study in the same education system and are taught the same mathematics curriculum. Hence, the advantage of the present work in contrast to international studies, is that both the education system and the curricular variable in some cases hold constant across the groups.[13]

Systematic follow-up on the Israeli data also included comparison between the findings for Israel and other countries, including comparison of data for Israeli Arabic-speaking schools versus Arab countries documented in the literature and in international tests. The two groups share a common language of instruction and assessment. Moreover, Arab-country schools also displayed a low achievement level compared with Western-country schools. A separate examination of gender gaps in both Israeli language communities and a comparison with the data for other countries provide an opportunity for better understanding the gender gaps and, perhaps, their causes – with potential insights for narrowing them. Moreover, review of the Israeli findings alongside the international picture opens up the possibility of generalizing from the present study's conclusions to the universal level.

As noted, the present study will investigate gender gaps not just separately, for each study discipline, but also together. The aim will be to deepen our understanding of the reciprocal relations between achievements and gender gaps in the two subjects, and to examine them in relation to the trends indicated by the aforementioned studies that addressed gender gaps in the two subject domains in tandem. This may shed new light on the issue of math and language gender gaps at the universal level.

Method

Instruments and data

The present work looked at gender gaps that emerged from student scores on Israeli GEMS assessment in mathematics (Grades 5 and 8) and Hebrew/Arabic native-language domains (Grades 2, 5 and 8[14]) during the six-year period 2008 to

---

socioeconomic-cultural background, but are still significant. For a complete picture, see the GEMS reports and the Israel-specific international test reports at the RAMA website: http://cms.education.gov.il/educationcms/units/rama .

[13]By contrast, "native-language" studies (Hebrew and Arabic, respectively) differ for the two linguistic communities and have separate curricula, teaching methods, supervision systems and different content in the national exams.

[14]The GEMS known as "MEIZAV" (Hebrew acronym for "School Growth and Efficiency Measures") is a national assessment system that includes student achievement exams as

2013. Gender gaps on international tests in which Israel participated were analyzed as well: PISA, TIMSS and PIRLS. For PISA, data are presented from the study years 2006, 2009 and 2012; for TIMSS the years covered are 2007 and 2011; for PIRLS, data are presented from 2001, 2006 and 2011. Generally speaking, since 2006, the studies were all managed and administered in Israel by the National Authority for Measurement and Evaluation in Education (RAMA).[15]

In the second part of the work, in which gender-gap data for both subject domains will be examined side by side, data will also sometimes be presented on gaps in a third domain – science, which is also included in the assessment systems covered in this work.[16]Adding the science domain is meant to clarify and enrich the math-language comparison. What distinguishes science in the school setting is its combination of quantitative-mathematical features (necessitating, for instance, mastery of quantitative reasoning skills, knowledge of units of measurement and their conversion from one scale to another, dealing with graphical problem etc.) and verbal features (understanding theoretical models and texts describing biological or chemical processes). This combination of two modes of reasoning characterizes both the learning stage and the assessment stage of science. For example, GEMS exams in Science and Technology include relatively long reading passages dealing with scientific topics that necessitate good reading and comprehension abilities.

On GEMS exams, students from both language communities are administered the same math and science tests. In both subjects the tests are developed in Hebrew and then translated from Hebrew into Arabic. In native-language domains students from the two groups take different tests that are developed separately and in accordance with differing curricula for Hebrew and Arabic. Regarding the international tests in which Israel participates, students from both communities

---

well as questionnaires designed to glean information about the school climate and pedagogical environment (questionnaires are administered to principals, teachers and students). The purpose of the GEMS at the school level is to provide schools with data that can be used for planning and managing and monitoring. At the system level, the GEMS is intended to provide a snapshot of the mastery level of Israel's students in the curricula of four core subjects, and to serve professional entities and policy makers in the Education Ministry. The GEMS exams are administered in four domains: Native Language (Hebrew/Arabic) Mathematics, English, and Science & Technology. Tests are administered at grade 5 and 8 in four subjects while in grade 2, only exams in native language (Hebrew/Arabic) are administered. The GEMS are developed and administered by RAMA.

[15]For more information about RAMA, its functions and the areas for which it is responsible, see: http://cms.education.gov.il/educationcms/units/rama

[16] Each year between 2006-2013 a GEMS exam on "Science and Technology" used to be administered (to Grades 5 and 8); science is tested in the TIMSS assessment along with math; and "science literacy" is one of the three literacy areas assessed in PISA, along with math and reading.

take the same tests, which are translated and adapted separately and independently from the source language (English or sometimes French) into Hebrew and Arabic as target languages.

Study population

Students in Hebrew-speaking schools belong to three different educational streams: state, state-religious or haredi (ultra-Orthodox). Students in Arabic-speaking schools belong to the Arab, Bedouin, Druze or Circassian sectors.

In all testing frameworks relevant to Israel the student samples numbered in the thousands; in each of the testing systems the test takers were representative of their cohorts (in terms of age or grade level). However, there are several differences between the various studies with regard to student sampling. In the GEMS framework and the TIMSS and PIRLS studies, test-takers represent students in the state education systems (for both language sectors) and the state-religious system[17] only. The PISA sample includes, in addition to state and state-religious school systems, students in haredi girls' institutions and a small number of students in haredi boys' institutions (a group that does not represent the entire haredi male student cohort).[18] Starting with PISA test year 2009, students in apprenticeship institutions and schools under the supervision of the Ministry of Economy were included. Students in the Ministry of Economy schools account for 3% of the students in the Hebrew-speaking community and 5% of the students in the Arabic-speaking community. In both language sectors, most students who attend these institutions are male, meaning that their inclusion in the study had an impact on the size of gender gaps.[19]

[17]Haredi students do not usually take the Israeli large scale tests, and when they do, the sample is not representative of the haredi student population as a whole. Thus, in this study their data are not included in the GEMS, TIMSS or PIRLS results. Also, special education students do not participate in any of these assessments, nor do students who have recently immigrated to Israel (less than one year in the country). For more information on the numbers of students who take each test, their distribution by language sector and gender, or the percentage of special education students, see the GEMS and international test reports published periodically by RAMA.

[18]PISA disparity data in this report include haredi students who participated in the study (mainly girls).

[19]Generally, in international studies, students who participate in the assessment are tested in the various subject areas. In PISA – math, science and reading literacy; TIMSS – science and math. However, on the GEMS exams the same student sample is tested only in math and native-language (the tests in both subjects are administered on different dates, three weeks apart), while in science a different sample is tested –all samples are a representative sample for a given grade level in a given year.

Measures

For GEMS exams, the gender gap is calculated in terms of effect size separately for each sector, that is, the gap between the girls' average and the boys' average (the girls' average minus the boys' average), divided by the standard deviation for all students – boys and girls together – in a given language sector.[20]

For international tests, the gender gap is calculated in terms of the simple difference between the girls' mean and the boys' mean (not divided by standard deviation). For these studies the score scale was determined by the international organizations responsible for them, meaning that the mean score for the participating countries was500 and the standard deviation was100.[21] A standard deviation of 100 serves as an accepted international norm for describing score variability, making it possible to assign a meaning to the size of the group and gender gaps.

---

[20]For tests included in the present study, except for the native-language Grade 2 GEMS exams, scores are calibrated, which facilitates valid comparison of achievements over the years in a given subject and for a given grade level. Thus, we can also compare gender gaps over the years for the same subject and grade level. For all GEMS exams, except the Grade 2 native-language exam, scores in each subject and grade level are calculated according to a uniform (multi-year) scale. This scale was defined separately for each subject and grade levelin the first year it was employed, such that the mean score of all test-takers was 500 and the standard deviation was 100. The standard deviation for each subject and each test is quite stable and remains close to 100 each year .

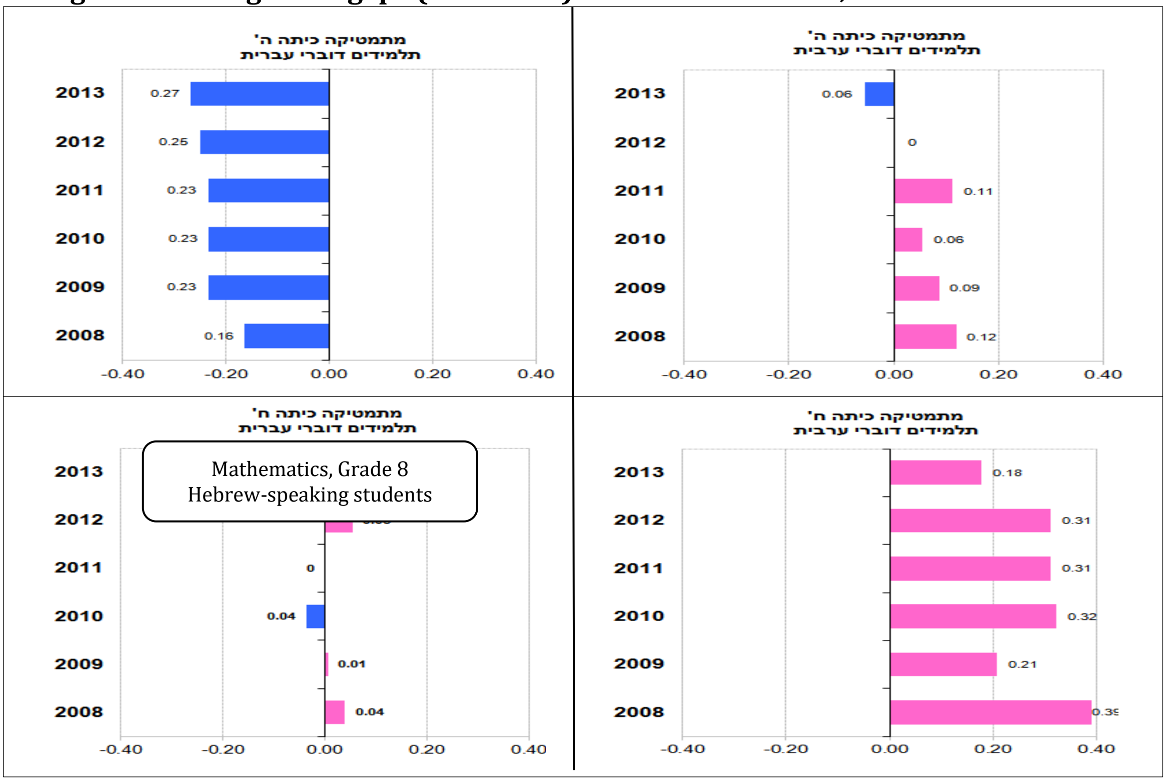
[21]For the PISA tests these mean and standard-deviation values are the mean (of the mean and SD) of the OECD countries, while for the TIMSS the mean and standard deviation are of all participating countries.

Results

Gender gaps in math on GEMS exams

Figure 2 presents the math gender gaps on the GEMS exams for Grades 5 and 8 in both language sectors over the period 2008-2013.

Figure 2: Math gender gaps (effect size) on the GEMS exams, 2008-2013



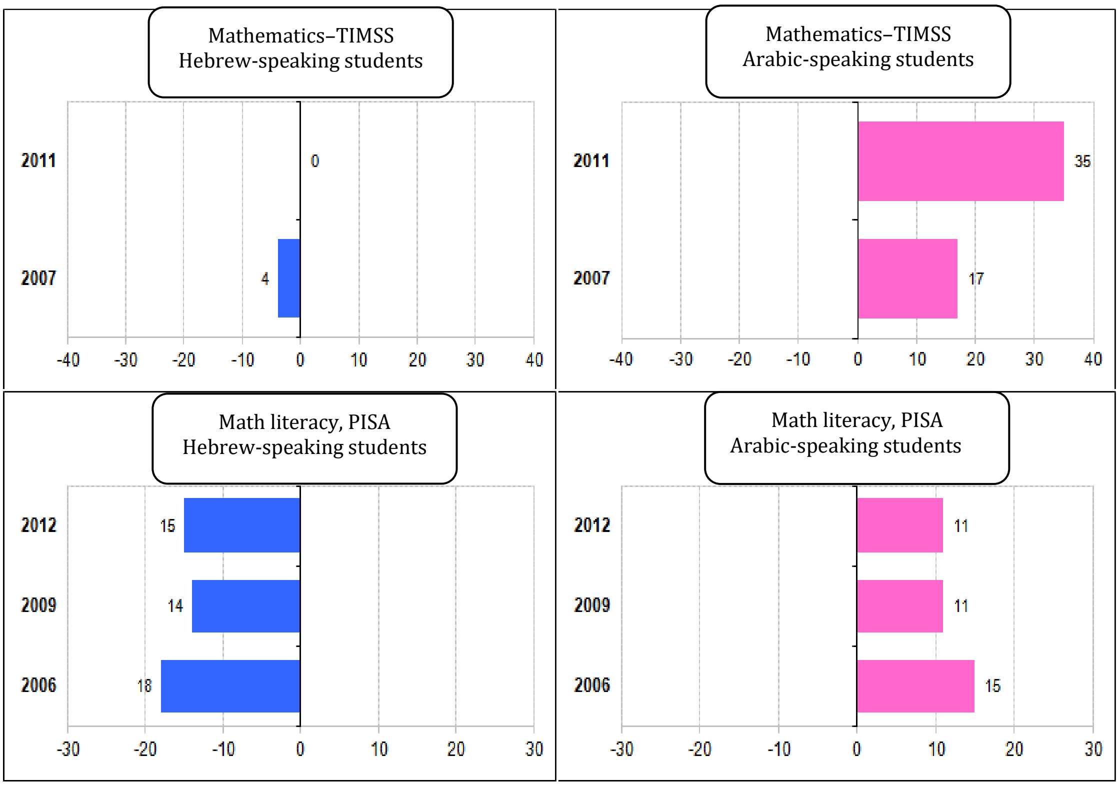
Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.

The diagram shows that the math gender-gap for Grades 5 and 8 as reflected in the GEMS test scores differs greatly between the two age groups and the two language sectors. However, the picture appears to be quite stable over the years for each given grade level and language sector. Fifth-graders in Hebrew-speaking schools in all of the test years exhibit disparities in favor of boys (a quarter of a standard deviation on average), and these gaps seem to have widened slightly over the years. By contrast, students in Arabic-speaking schools for all years (except 2013) showed small gaps of a tenth of a standard deviation in favor of girls (over the years, the mean disparity is 0.05 SD in favor of girls). The performance of eighth-grade boys and girls in Hebrew-speaking schools over the entire period is similar (the mean gap over all of the years is negligible), while eighth-graders in Arabic-speaking schools exhibited large disparities in favor of girls, ranging from one-fifth to one-third of a standard deviation).

Figure 3 presents the Israeli math-achievement gender gaps on the international tests in which Israel participated during the past decade. The diagram shows that, for each language sector, the math gender-gap situation on the TIMSS tests administered to eighth-graders reproduces, to some degree, the situation reflected in the eighth-grade GEMS exams: boys and girls attending Hebrew-speaking schools have similar achievements, while their peers in the Arabic-speaking schools show gaps in favor of girls ranging from one-sixth to one-third of a standard deviation (17 points and 35 points in favor of girls in 2007 and 2011, respectively). On the PISA mathematical literacy tests, Hebrew-speaking students were found to exhibit gaps favoring boys of 16 points on average, while their peers in the Arabic-speaking schools showed gaps in the opposite direction (favoring girls), of 12 points on average. It is interesting to note the stability of the disparity levels for both language sectors, between the two last PISA study years, 2009 and 2012. It also turns out that gender gaps in the Hebrew-speaking schools are similar to the average gaps in the PISA participating countries overall, while the boy-girl gaps (in favor of girls) found for the students in Israel's Arabic-speaking schools are exceptionally large compared with those: Figures 7 and 8 (later in this study), which present the gender gaps of the countries that participated in the 2012 PISA and the 2011 TIMSS tests, show that Israeli Arabic-speaking students are characterized by some of the largest gender gaps (favoring girls) of all participating states.

Figure 3: Israeli math gender gaps in international studies

Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.



Gender gaps in language (Hebrew/Arabic) on the GEMS exams

Figure 4 presents the native-language (Hebrew/Arabic) gender gaps on the GEMS exams for Grades 2, 5 and 8 during the period 2008-2013.[22]The figure points to a consistent picture over the years, in which both language sectors exhibit native-language gaps in favor of girls at all grade levels. The gaps clearly widen as grade level rises. In Hebrew-speaking schools second-graders exhibit a one-fifth of a standard deviation gap; by Grade 5 the gap widens to a quarter of a standard deviation, and then to a third of standard deviation in Grade 8. In the Arabic-speaking schools the gaps start at one-quarter to one-third of a standard deviation in Grade 2, widen to one-third to one-half of a standard deviation in Grade 5, and rise to one half of a standard deviation in Grade 8. Although the Hebrew and Arabic tests are not the same, and scores not comparable, in standard-deviation terms girls' advantage over boys in language, at all grade levels, is greater among the Arabic-speaking students than among the Hebrew-speaking students. It is also worth noting that fifth-graders i[illegible] schools have exhi[illegible] certain tendency for the boy-girl disparities to dwindle over the years, fro[illegible]

Math literacy, PISA
Hebrew-speaking students

Math literacy, PISA
Arabic-speaking students

[22]As noted above, it is impossible to compare the scores on the Hebrew and Arabic native-language tests, because the tests were developed separately. Please note: for Grade 2 the gender gap was calculated as the difference of the mean scores on the raw scale divided by the joint standard deviation for boys and girls, multiplied by 100, because the tests for Grade 2 are not equated from year to year in the same way as the are the tests for Grades 5 and 8.

half to one-third of a standard deviation. In 2013 there was also a narrowing of the Grade 2 gap, a finding that should be monitored and checked during the coming years.

Unlike GEMS, in the international studies in which Israel participates and that assess reading skills, tests are translated separately to each target language – Hebrew and Arabic – from the English/French source, meaning that the tests administered to both language sectors are practically the same. This makes it possible to compare the achievements of the two sectors in language as well. Figure 5 presents the gender gaps on the international tests in which Israel participated over the past decade, and which assessed scholastic achievements in the language domain: PIRLS (reading comprehension, Grade 4) and PISA (reading literacy, age 15). The diagram shows a similar picture to that presented by the GEMS exams: both assessments show gaps in favor of girls in both language communities. These gaps appear to be quite stable over the years (except for PIRLS 2011 in the Hebrew-speaking schools, where no disparity was found between boys and girls[23]). Moreover, in both studies, reading gaps favoring girls among students in the Arabic-speaking community are larger than the gaps among their Hebrew-speaking peers. In the PISA study gender gaps in the Arabic-speaking schools are especially large – two-thirds of a standard deviation, versus one-third of a standard deviation in the Hebrew-speaking schools. As with mathematical literacy, the reading-literacy gaps exhibited by students in Hebrew-speaking schools who took the PISA tests are similar to the OECD mean, while students in Israeli Arabic-speaking schools exhibit some of the largest gaps of all participating countries in both the PISA and the PIRLS studies (see Figures 7 and 8 below). Israeli data for international tests, as for GEMS exams, reveal that the gender gap favoring girls is much smaller in the lower grades (PIRLS Grade 4) than in the higher grades (PISA, 15 year olds, i.e. grades 9-10).

[23]This finding is unusual both compared with other Israeli data and compared with the data of the other participating countries; it needs to be investigated further so that it can be better understood.

Figure 4: Gender gaps (effect size) in native-language (Hebrew/Arabic) on the GEMS exams, 2008-2013

Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.

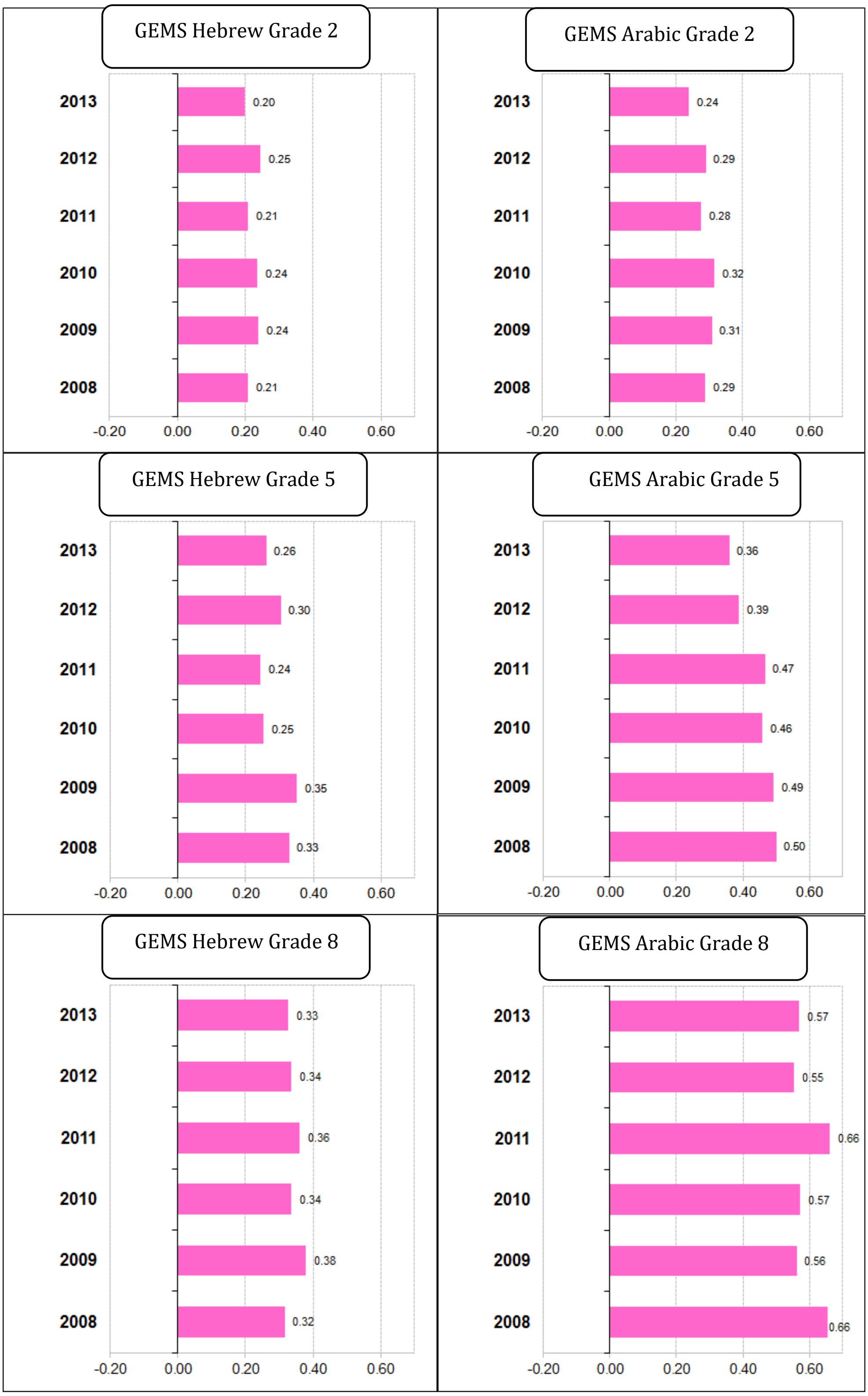
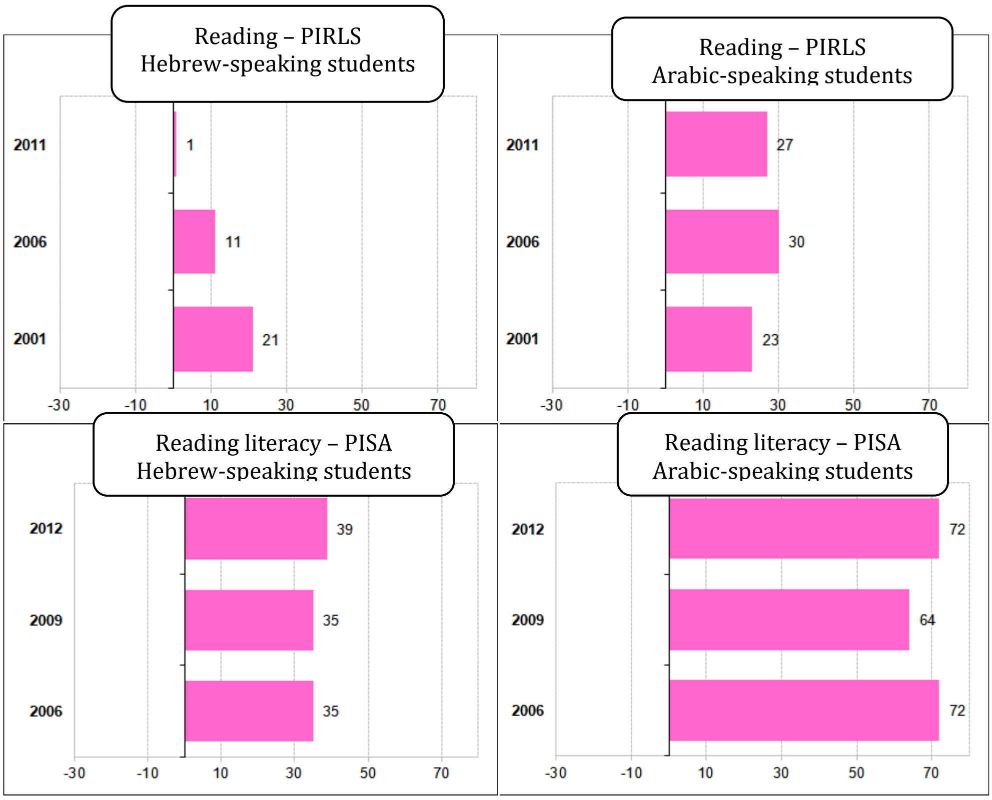


Figure 5: Gender gaps in reading comprehension (native language) in Israel, in international studies
Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.



Interim summary

Main findings from the Israeli data presented up to now:

1. The pattern of gaps among students in Arabic-speaking schools is very different from that of students in Hebrew-speaking schools. The size heterogeneity of the math gender gaps at the various grade levels and in the different language sectors is consistent with the heterogeneity reported in the relevant research literature (see Introduction).
2. <u>Students in Hebrew-speaking schools</u> – Both in math and in native-language domains the size and direction of the gender gaps are similar, on the whole, to those documented for other Western countries (comparable to the OECD average or the average of the TIMSS countries). In **mathematics** the disparity values are also similar to those reported in the research literature (primarily American data). For example, in Grade 5 there is a gap of one-fifth to one-quarter of a standard deviation in favor of boys. Moreover, the period 2008-2013 witnessed a trend toward widening gaps (in favor of boys), while Grade 8 exhibited a consistently negligible gap between the sexes over the years. Among 15 year-olds students a

disparity favoring boys in math literacy is once again detected, at an extent of one-sixth of a standard deviation; in the language domain girls outperform boys in all years, at all grade levels and in all of the studies (except for one observed aberration of parity between boys and girls in the 2011 PIRLS).

3. Students in Arabic-speaking schools – in both math and native-language domains gaps were found in favor of girls. In all cases and in all of the studies, the girls' advantage is much larger than the parallel disparities among students in Hebrew-speaking schools. These gaps are very large compared with those of other countries as documented by international studies, and are similar in size to those found for Arab countries. Recent years have witnessed a trend toward closing the gap in math among fifth-graders, and in 2013 a gap favoring boys was documented for the first time. In language the gaps widen as grade level rises (Grades 2, 5, and 8). And similar to math, a trend has been observed in recent years regarding native-language skills towards narrower gaps in Grade 5.

Examining reciprocal relations between gender gaps in the two school domains in Israeli GEMS exams and international tests

This section deals with the reciprocal relations between student achievements and gender gaps in math and language, and presents a comparison of the trends that have been found over the years in these two subject areas. For convenience purposes some of the gaps presented earlier in math and languages will be shown again, but this time side by side. Also, the data will be accompanied by information on gender gaps in science (as explained in the Introduction[24]).

Figure 6 presents gender gaps in math, science and language on the GEMS exams at two grade levels (5 and 8) and for the two language sectors. The diagram illuminates a most important aspect of the gender gaps: there is almost always a fixed ranking between the subjects in terms of gender gap size. Accordingly, native-language is the discipline where girls' advantage over boys is always the largest out of the three subject domains; by contrast, math is the school domain characterized by the largest gap in favor of boys (or, alternatively, the smallest gap

[24]As noted above, the discipline of science is not in the focus of this study, but including it sheds light on important aspects of the analysis of the reciprocal relations between the two subject areas on which the study focuses. This is because science is regarded as a subject that has both quantitative and verbal features. In the course of this section data are presented on science in the various test formats (GEMS – science and technology; TIMSS – science; PISA – science literacy).

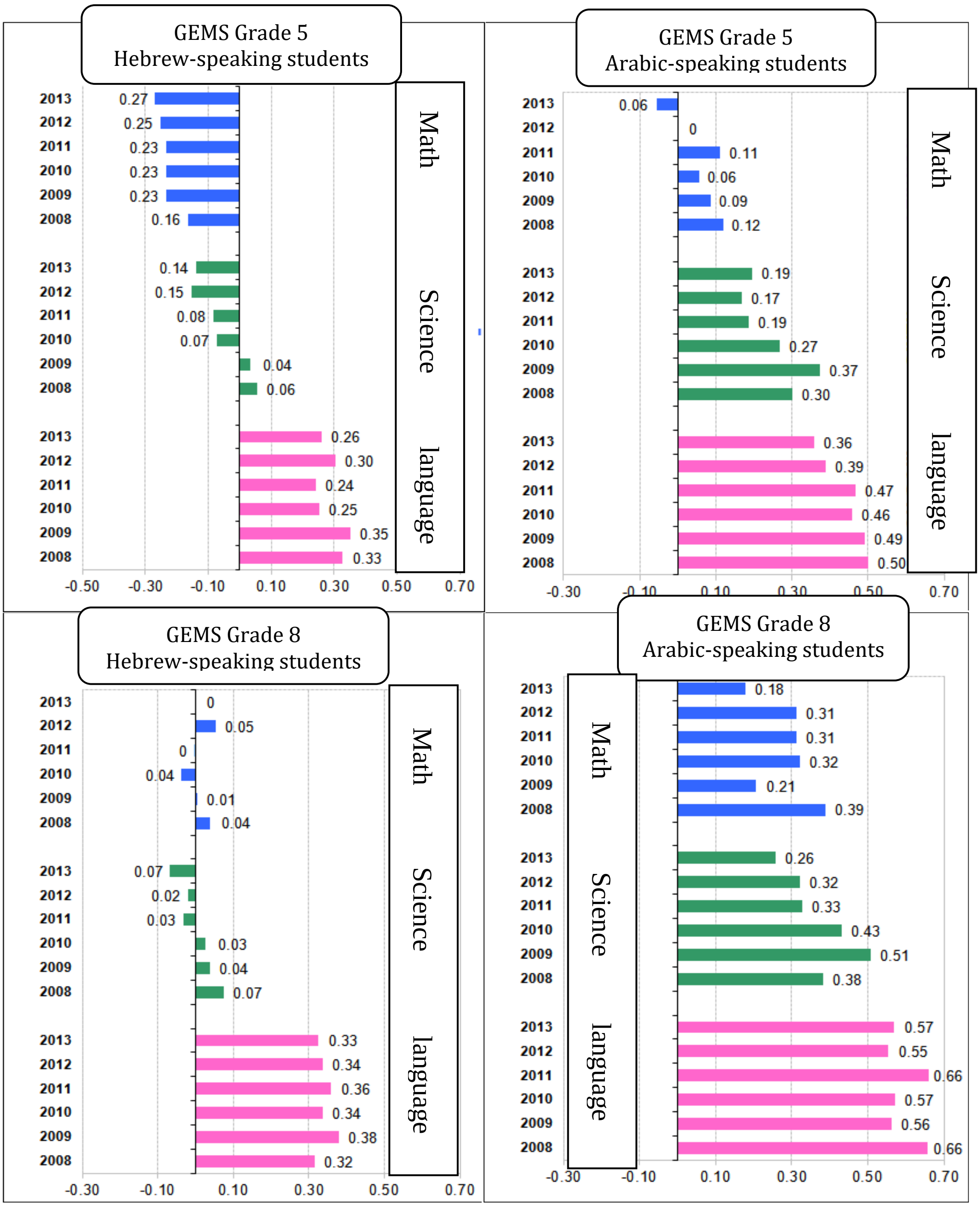
in favor of girls); science is situated at a midway point between language and math in terms of gap size[25].

As noted in the Introduction, this ranking between gender gaps in the three subject areas – language-science-math – is universal, and has been documented by the research literature via school scores (Voyer & Voyer, 2014) as well as in international studies across an array of countries (e.g., the PISA assessments). To illustrate, Figures 7, 8 and 9 present the various countries' gender gaps for the most recent test years of the international studies – PISA 2012, TIMSS 2011 (math and science, grades 4 and 8), and PIRLS 2011 (reading, grade 4). Israeli gender gaps were added to the graphs separately by language sector.[26]

Figure 7 clearly shows the ranking between the three areas of knowledge in the PISA 2012 study: in all of the countries, without exception, girls outperformed boys in reading proficiency (the mean OECD disparity was 38 points). By contrast, most of the countries displayed math gaps favoring boys (mean OECD gap of 9 points in favor of boys). In a small number of countries, such as Sweden, Norway, Kazakhstan, Slovenia and France, no gap was found between boys and girls, or the gaps were small and insignificant. In a few other countries however – e.g., Jordan, Qatar, Iceland and Finland – girls outperformed boys in math. As with the GEMS data, the gender gaps in science literacy were found to lie somewhere between the gender gaps in reading and math literacy, thus presenting a balanced picture. The mean gender gap in science is one point score; in most countries there is no significant gap between boys and girls. A few countries exhibit a gap in favor of boys, while in a few others, girls' performance is better. Another prominent finding seen in Figure 7 is that in all literacy areas, math included, the Arab countries (Jordan, Qatar, the UAE) "lead" in terms of the size of gaps favoring girls. As noted above, this finding also characterizes students in Israel's Arabic-speaking schools – disparities favoring girls in language, science and math. What is worth noting at this point is that despite the fact that in these countries girls outperform boys in all three schools domains, the aforementioned ranking between the subjects still persist.

[25]The sole exception to this recurring pattern is the outcome of the GEMS Grade 8 test in Hebrew-speaking schools: gender gaps in science and math seems to be similar in size and, in fact, are very small and without a clear direction.
[26]The TIMSS graph for Grade 4 does not present the Israeli gaps, as Israel did not participate in that assessment.

Figure 6: Gender gaps on GEMS exams in math, science and language
Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.

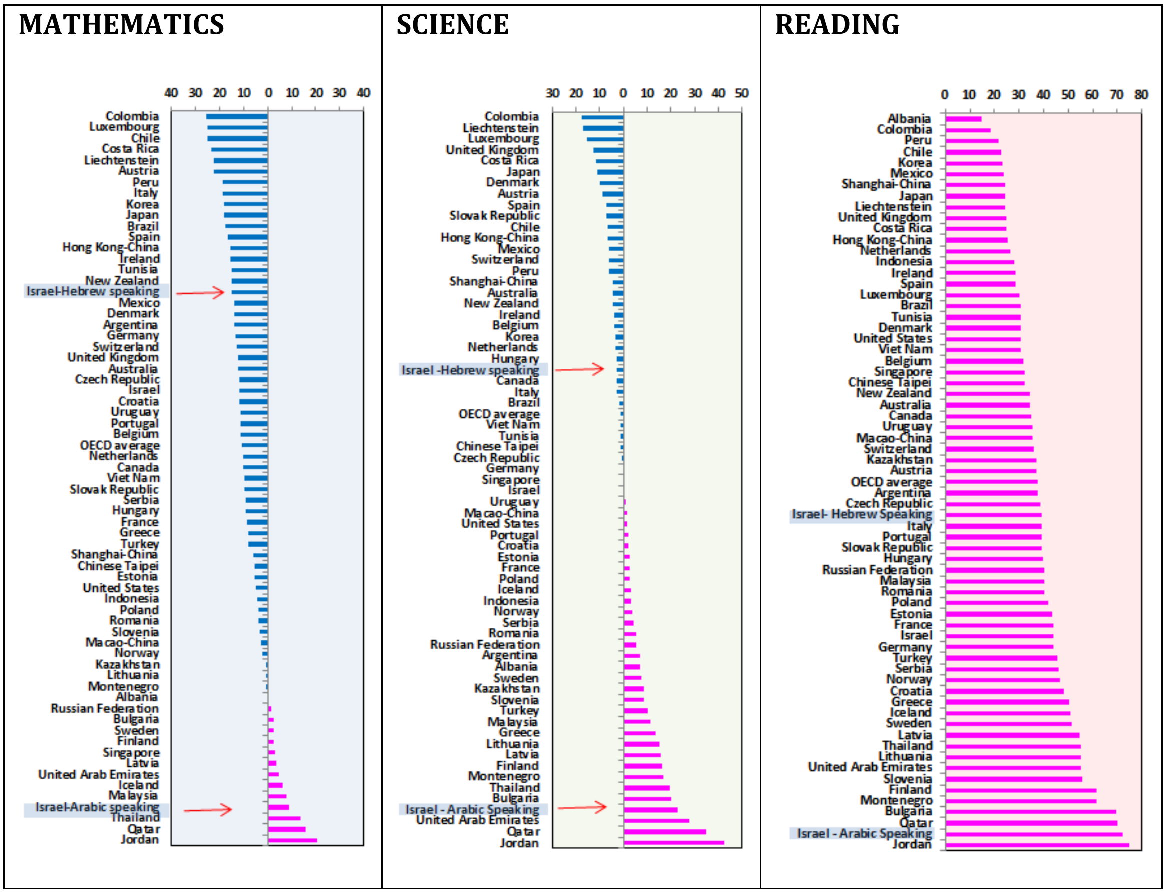


The gender-gap ranking in the three domains partly repeats itself in the Grade 4 TIMSS and PIRLS studies and in the TIMSS Grade 8 study[27] (Figures 8 and 9). The gender-gap data for Grade 4, presented in Figure 8, clearly indicate that girls consistently outperform boys in reading, while in math and science the gender gaps tend to be smaller. However, the math and science disparities lack a clear direction: in certain countries girls do better, while in others boys do better. Thus, there appears to be no clear gender-gap size ranking between the two disciplines of math and science. A balanced (and on average – negligible) international gender-gap picture in science and math was also found in the TIMSS Grade 8 study (Figure 9). It should be recalled that a similar picture was also found for Israeli eighth-graders in Hebrew-speaking schools, on both the GEMS exams and the TIMSS math and science tests. Finally, when the focus is on the Arab countries that participated in the TIMSS assessments (Morocco, Jordan, Bahrain, the UAE, Qatar, Oman, Saudi Arabia, the Palestinian Authority and others) and on students in Israeli Arabic-speaking schools, the reading-science-math ranking order(or only the science-math ranking) turns out to exist both in the TIMSS Grade 4 study and in the TIMSS Grade 8 study: in both cases the math gap for these countries (in favor of girls) was smaller than the girl-favoring gap in science. And the TIMSS study (like the PISA) again found the Arab countries to have strikingly large gender gaps favoring girls in all subject areas assessed.

[27]In 2011 several countries participated in both the TIMSS Grade 4 tests (math and science) and in the PIRLS study (Grade 4 reading), meaning that there is a basis for comparing those countries' findings in the relevant disciplines.

Figure 7: Gender gaps in the PISA 2012 study in the three literacy areas[28]

Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.



[28]The diagrams are based on data from the PISA 2012 report (OECD, 2013).

Figure 8: Gender gaps in the PIRLS and TIMSS Grade 4 studies, 2011[29]

Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.

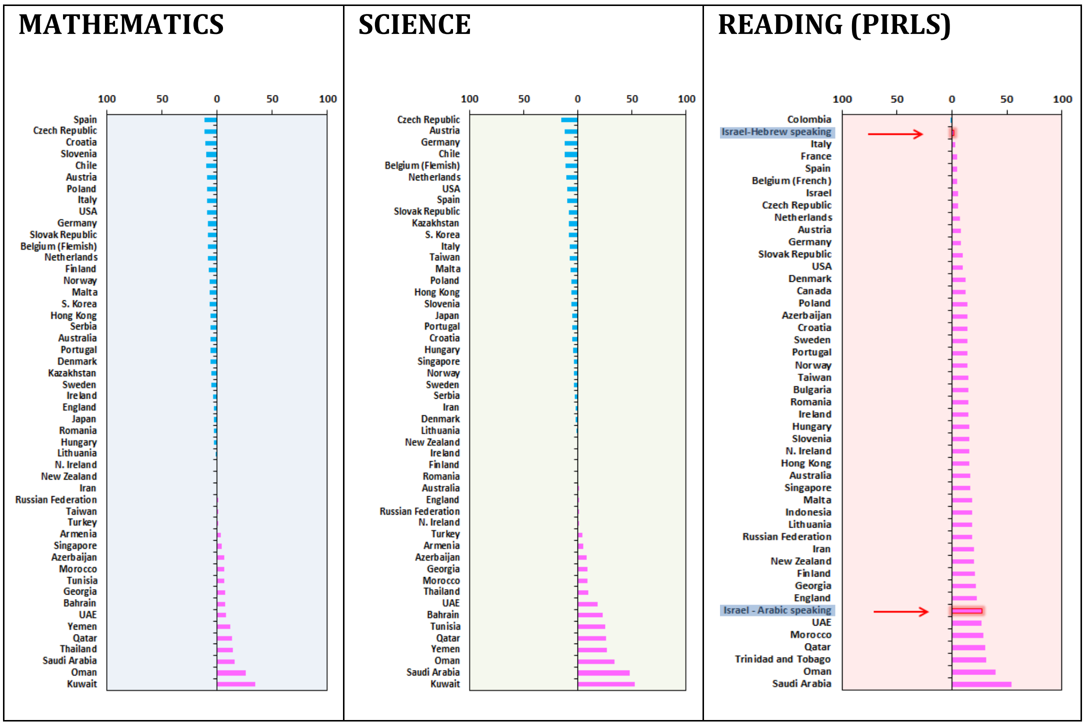
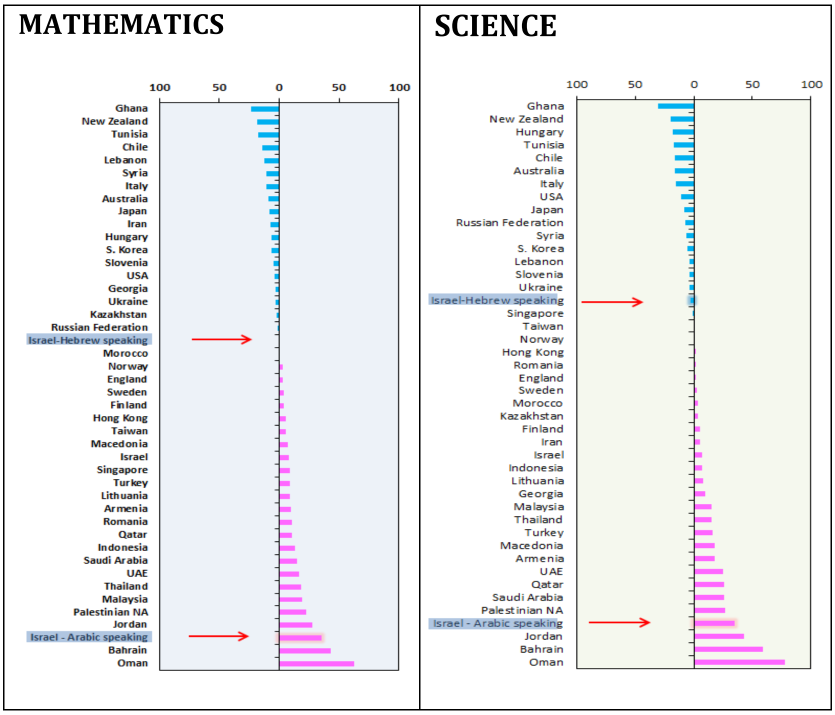


Figure 9: Gender gaps (boy-girl achievements) in the TIMSS 2011 Grade 8 study

Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.



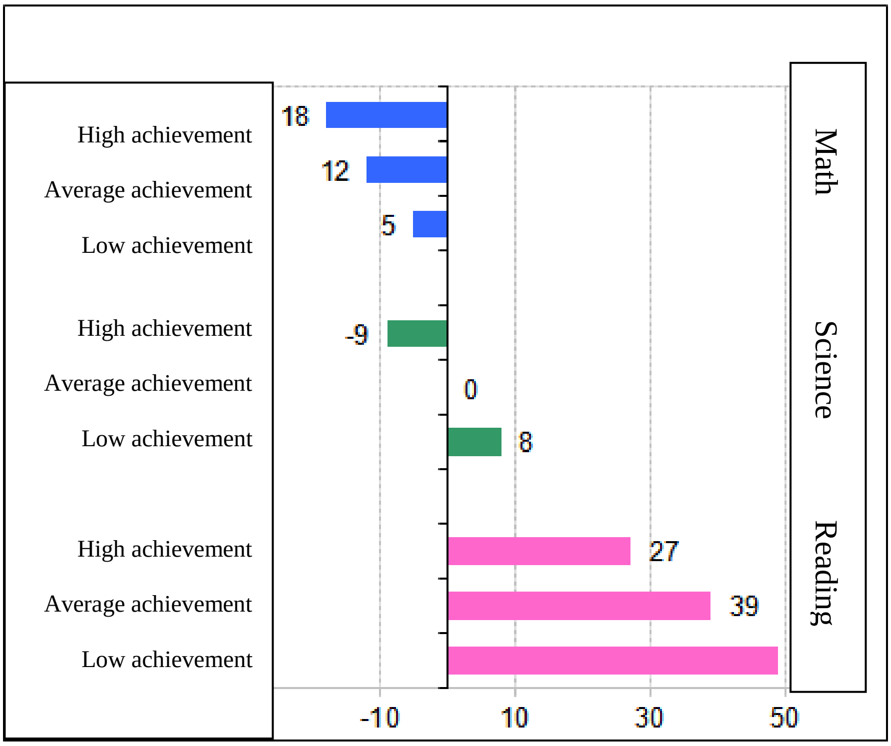
[29] Graphs are based on data from the TIMSS 2011 and PIRLS 2011 reports (Mullis, Martin, Foy & Arora, 2012; Martin, Mullis, Foy & Stanco, 2012 and Mullis, Martin Foy & Drucker, 2012). Please note that Israel did not participate in the TIMSS Grade 4 study.

The gender-gap ordered ranking between reading-science-math is so consistent that we find it not only across grade levels, sectors, countries and tests (as well as school marks as presented by Voyer & Voyer, 2014, reviewed above), but also when test-takers are divided into ability groups. Del Pero & Bytchkova (2013) divided students in the PISA 2009 study into three groups by performance level, and calculated the gender gap in the three literacy areas for each of the groups (data presented in Figure 10). In the three literacy areas – reading, science and math – the gaps favoring girls were found to widen the lower the group's achievement level. In the present context, however, what is interesting is that within each separate student group (high-achievement, average-achievement and low-achievement), the gender gap in favor of girls is largest for reading, followed by science, and then math – in which boys outperformed girls. For example, in the low-achievement group there was a gender gap favoring girls of 49 points in reading and 8 points in science, but a 5-point gap favoring boys in math[30]

[30]Stoet & Geary (2013) also calculated the math and reading gender gaps for four PISA test years. Their findings were similar to those presented here. Their calculation of the gaps among students in the fifth percentile, the fiftieth (median) percentile and the 95th percentile indicate that the higher the percentile, the greater the disparity favoring boys in math and the smaller the reading advantage favoring girls. Stoet & Geary also called attention to the fact that the percentage of high-performing boys in math is more than twice that of girls, while the percentage of low-performing boys in reading is nearly 5 times that of girls. The PISA 2012 data (OECD, 2013) also showed that, on average, boys outperform girls in math competencies, but that the boy-girl gap is larger among high-performing students and almost nonexistent among low performers. These findings are also consistent with another finding that is not discussed in detail in this study: among students of low socioeconomic status (who also usually have lower scholastic achievement levels), girls tend to perform better scholastically than boys. This is true of all subject areas, but less so of math and more so of native-language skills (see Rapp et al., 2013).

Figure 10: Gender gap in the PISA 2009 study in the three literacy areas, for three levels of student ability: high, average and low (data from del Pero & Bytchkova, 2013)

Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.

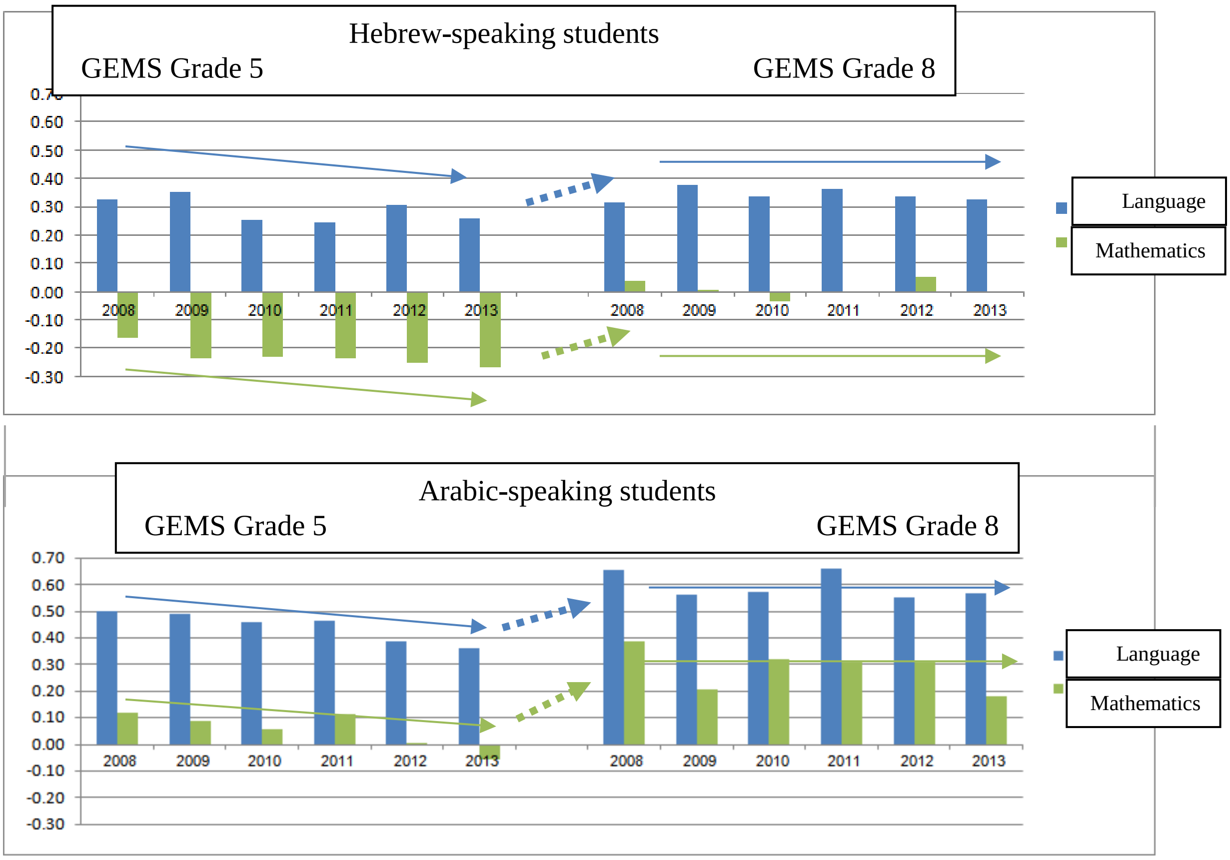


Correspondence between gender-gap sizes in different areas of knowledge

Another finding that emerges from a comparison of gender gaps in the different subjects assessed by the GEMS exams, and which touches on the fact that the ranking between the domains holds up in different situations, is that the gender-gap trends in the subject areas correspond. That is, when the gender gap in one area is found to be changing in a particular direction, growing or narrowing, a gender-gap trend in a similar direction is seen in the other school domains. Figure 11 presents gender gaps on the GEMS math and language exams for Grades 5 and 8 over the years, and change trends are delineated across years and grade levels. The graph clearly shows that the gender gap favoring girls in language is larger in Grade 8 than in Grade 5, and that, accordingly, the math gaps vary between Grades 5 and 8 in the same direction: among Hebrew-speaking students the gap (in favor of boys) in Grade 5 declines by Grade 8, while among their peers in the Arabic-speaking schools the gap (favoring girls) widens. Multi-year trends clearly indicate that on the GEMS Grade 5 exams for both language sectors there is a trend toward smaller boy-girl gaps in language tests, accompanied by a trend in the same direction in math[31]

Figure 11: Trends in gender gaps in language and math on the GEMS exams 2008-2013, and between Grade 5 and Grade 8

Right-facing columns represent advantage in favor of girls; left-facing columns – in favor of boys.



This congruence between the size of gender gaps in the different subject areas is consistent with the correlation that was found between gender gaps across (and within) different countries, as reported some time ago by Stoet & Geary (2013), who looked at PISA data for test years during the period 2000-2009 (earlier calculations had been performed by Guiso et al., 2008 for PISA 2003). In order to reinforce the validity of this consistent finding, we re-checked, in this study, the correlation between the gender gaps for the most recent PISA study, 2012, in which math literacy was the main discipline assessed. Correlations across the 65 countries that took part in the PISA assessment were as follows: 0.75 between the math-literacy gender gap and the reading-literacy gender gap; 0.84 between the

[31]In Hebrew-speaking schools the math gaps (favoring boys) widened over the years, while in Arabic-speaking schools during the same years the math gaps (favoring girls) diminished to the point that, in 2013, boys slightly outperformed girls for the first time.

reading gender gap and the science gender gap: 0.86 between the math gender gap and the science gender gap. Similar correlations were also found for earlier PISA test years[32]. The meaning of these positive correlations is, again, that in countries for which no boy-girl math gap was found (regarded as countries where the gap has been eliminated) a relatively large gap favoring girls is documented in reading literacy; similarly, in countries where boys have a large advantage in math a relatively small gap favoring girls in reading is found. To illustrate, Figure 12 presents a scatter plot of the gender gaps in reading and mathematical literacy in the PISA 2012 study, across the participating countries; the diagram indicates a positive correlation between the gaps. In any case, it is important to note that existence of a statistical correlation between the gaps across the various countries does not, in and of itself, testify to a causal relationship or effect in a particular direction between those gaps.

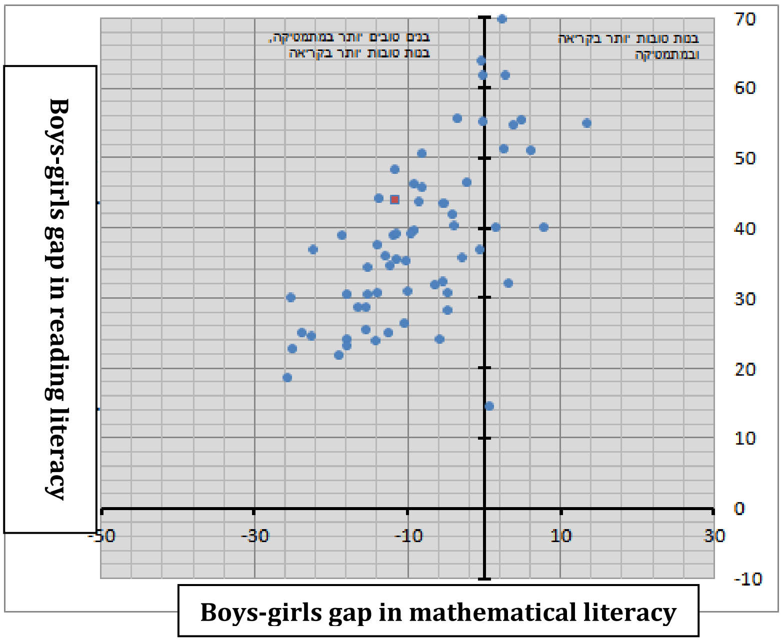
Figure 12: Math and language literacy gaps in countries that participated in PISA 2012 (each point represents a participating country)

Positive values represent a gap in favor of girls; negative values represent a gap in favor of boys. The red square-shaped point represents the values for Israel. The points on the upper-right portion of the diagram represent outlier countries where girls outperform boys in both reading and math competencies, e.g. Arab countries.

Boys better in math, girls better in reading

Girls better in reading and math

[32]Correlations of similar magnitude are reported by Stoet & Geary between the reading and math gender gaps. According to them, the correlations in the four test years covered range from 0.60 to 0.78. In a calculation that we performed using data from PISA 2009: between the math-literacy gender gap and the reading-literacy gender gap – 0.75; between the reading and science gaps – 0.78; and between the math and science gaps – 0.81.



Correlations between achievements in the various subject areas

It is conjectured that the correlation between gender gaps in the various school domains is rooted in a pre-existing statistical correlation between individual student scores in those subjects (see Introduction). That is, a student who tends to have high achievements in one subject (e.g. math) will be expected to have high achievements in the other subject (e.g. language) too. So if Student A is better than Student B in math, likely he will also outperform Student B in language. This can be generalized to groups of students: if Group A tends to have higher mean achievements than Group B in one school domain, the same can be expected in the other domain. Hence, if there is an achievement gap between two groups of students in one school domain, there will most likely also be a gap in the other discipline, which achievements in it are correlated with achievements in the first discipline at the individual-level. Due to this correlation, a widening (or narrowing) of the gap between two groups in one school subject will usually be accompanied by a widening (or narrowing) of the gap in the other, meaning that we will ultimately expect to find a correspondence (and even a statistical correlation) between gaps in the two subjects.

We indeed found a strong positive correlation between individual scores in the various literacy areas among all PISA 2012 test-takers (from all participating countries), and this correlation was actually stronger than the correlation found between gender-gap sizes in the various literacy areas that was reported in the previous section: the correlation between math and reading scores was 0.85; between science and reading it was 0.86, and between science and math it was

0.89. Similar correlations were found in the PISA 2009 study[33]. Positive correlations between the large scale Israeli national tests (GEMS) in math and language scores, though somewhat weaker, were found as well[34]. These correlations reproduce the achievement correlations reported in the literature (see Introduction). Here as well it is important to remember that a positive correlation between scores in different subjects does not signify a causal relationship in a given direction between the two subjects, and that there may be a third variable influencing achievements in all study disciplines. We will expand on this idea in the Discussion section.

Analysis of gender gaps in math among students with identical proficiency levels in reading

Because achievements in the various subject areas, especially in math and language, are strongly positively correlated, it becomes necessary, as explained in the Introduction, to look at gender gaps in math while adjusting for the effect of achievement levels in language; that is, to determine the math gaps between boys and girls whose language skills (mainly, reading proficiency) are similar. In the PISA assessments, the reading-proficiency score scale is divided into 7: 1a, 1b, 2, 3, 4, 5, 6, 7, with an additional category for "undefined" students (below level 1). Reading proficiencies are accompanied by explanations of what students in each of the categories are generally able to do[35]. Figure 13 presents boys' and girls' mathematical literacy average score for PISA 2009, among students (boys and girls) of the same reading proficiency level (as measured by the study). The diagram presents the data for Israeli students (separately for the two language sectors) and for students in several countries known to have no math gender gaps (in a few of them, girls actually outperform boys in math). For example, in Finland and Sweden, famous for their gender equality, the boy-to-girl gap in math as tested by PISA was negligible[36]. Arab countries, like students in Israel's Arabic-speaking

---

[33]Correlations between the reading and math scores – 0.83, between the science and reading scores – 0.87, and between the math and science scores – 0.91 (PISA 2009 Technical Report, http://dx.doi.org/10.1787/9789264167872-en.

[34]For example, correlations between GEMS math and language exam scores in 2013 were 0.57 and 0.57 for Grade 5 and 0.66 and 0.65 for Grade 8 in Hebrew-speaking and Arabic-speaking schools, respectively. It may be assumed that these correlations are lower than those documented in the PISA study because the GEMS math tests and the GEMS language tests are administered at different times, about one month apart. By contrast, the PISA measurements for both subject areas are obtained from the very same test (different sections).

[35]For more information about what students at different proficiency levels are able to do, see the PISA international reports (OECD, 2010, 2013).

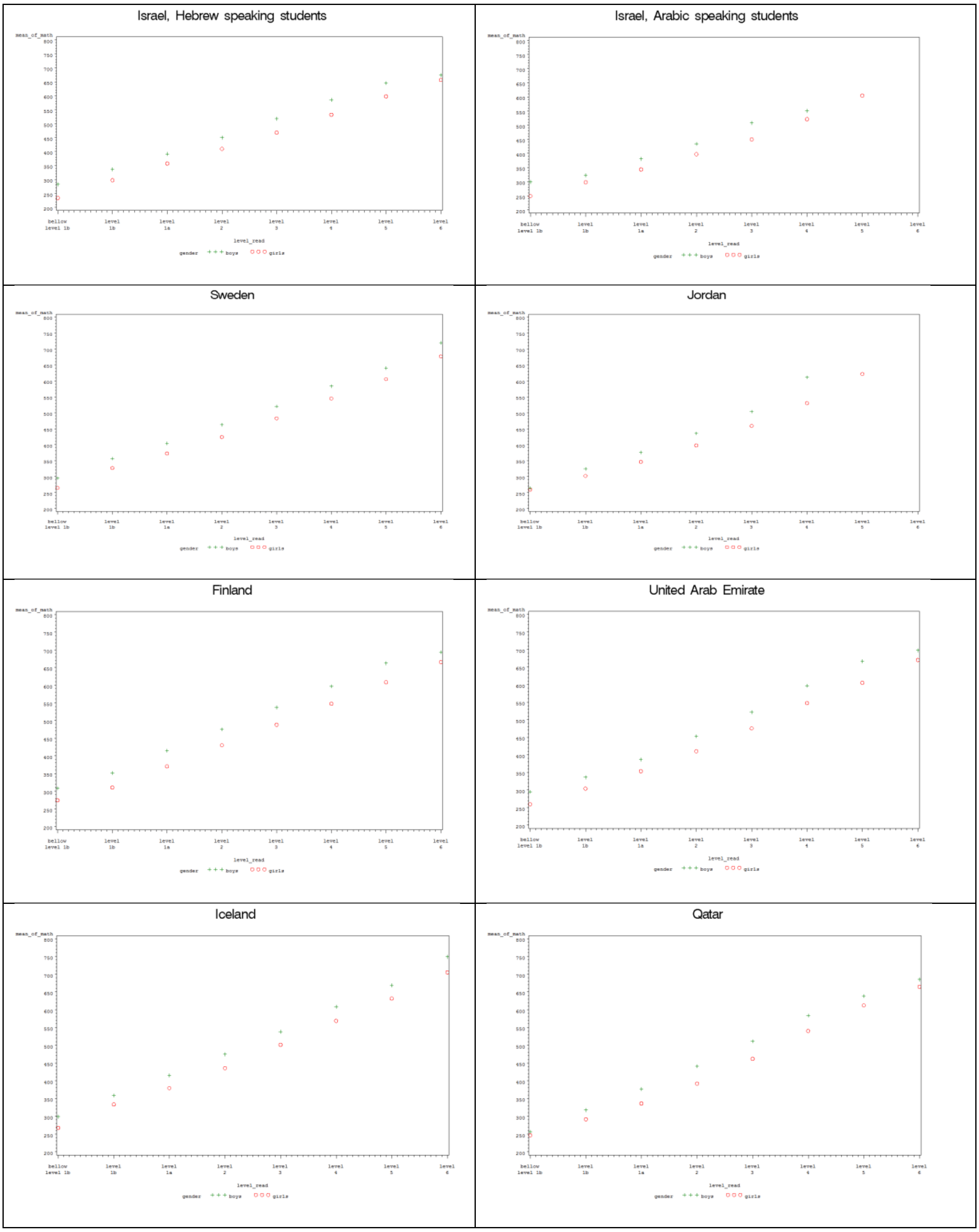
[36]In these countries the mean gaps were: Finland: 2.5 points in favor of boys in 2009 and 3 points in favor of girls in 2012; Sweden: 2 points in favor of girls in 2009 and 3 points

schools, were selected since they exhibited the largest math gender gap favoring girls of all the PISA-participating countries[36]

in favor of girls in 2012; Iceland: 3 points in favor of boys in 2009 and 6 points in favor of girls in 2012; Jordan: 0.5 points in favor of girls in 2009 and 21 points in favor of girls in 2012; Qatar: 5 points in favor of girls in 2009 and 16 points in favor of girls in 2012; UAE: 2.5 points in favor of boys in 2009 (Dubai) and 5 points in favor of girls in 2012).

Figure 13: Mean math literacy achievements of boys and girls as a function of reading proficiency in PISA 2012, in Israel (both language sectors) and selected countries

Black asterisks show boys' mean, red circles show girls' mean.



The graphs in Figure 13 clearly show that, at nearly all reading literacy levels and in all selected countries for which data were presented, boys' performance in math outperforms girls' at any given reading proficiency level. It is important to stress that boys' advantage in math within each reading-proficiency level persists despite the fact that in all countries and cultures represented in Figure 13 (except for Hebrew-speaking students in Israel), girls' mean achievements are similar, or even superior, to those of boys. Apparently, this picture stands in contrast to the differences between boys' and girls' average performance[37]. A similar picture is found for the GEMS Grade 5 and Grade 8 exams[38] among students in Hebrew-speaking schools. As noted, in this group the girls' mean math score is generally higher than that of the boys.

Discussion

The present work has explored the Israeli gender gaps in math and in native-language, as manifested in national and international large-scale standardized tests at schools. The study's original goal was to investigate and describe the main gender-gap trends as reflected in the data over the course of several years and with reference to a number of opportunities and studies, and to determine the degree to which these gaps are consistent with the international research literature and with data from international studies. However, because Israel's education system encompasses two main cultural-linguistic groups and since these two groups differ largely in regards to social, cultural and economic aspects as well as in scholastic performance in general, comparing trends in the two sectors sheds light on important aspects of the gender-gap issue as a whole. New universal insights emerged from analyzing the data of gender gaps in the different school domains mutually while looking at the relationships between them and the way they co-varied. . It is important to note, again, that the vast majority of publications on scholastic gender gaps usually deal with gender gaps in a single school subject (e.g., only in math or only in language), and that only recently have researchers begun to address more than one discipline at a time and to compare gender gaps in two different domains. This kind of approach offers new ways of

[37]There is an apparent contradiction here, but the two findings can, in fact, coexist. The difference between the means in favor of girls is due to the fact that girls' presence at the higher reading-proficiency levels is greater than that of boys in those categories and, since there is a correlation between language and math, the higher the reading-proficiency level, the higher the achievements in math. Thus, girls' mean achievements in math can be higher than those of boys, even if, at each specific proficiency level, boys' achievements are superior to those of girls.

[38]Due to the findings' similarity, the graphs are not presented here. It was possible to obtain this information for the GEMS exams because the same students were tested in math and native-language skills in any given year. GEMS does not divide its score scale by proficiency level; for the purposes of this study we divided the students by native-language score decile (each language sector separately). The finding repeated itself as in the PISA assessment, except for one instance in the GEMS Grade 8 exams (in the lowest reading/writing decile).

analyzing and understanding data with regard to gender gaps in scholastic achievement (especially in math), and opens up new directions for further study. The present study's findings also have pedagogical-practical implications, and suggest new channels for helping students fulfill their intellectual potential and close gaps between them and their grade-level peers (male or female) and for improving the education system as a whole.

The present study's main findings are as follows: (1) Israel's gender gaps in different areas of knowledge display a correlation across different sectors, test years, grade levels and testing systems. This is consistent with the existence of a correlation between these gender-gaps across and within different countries. (2) Students in Israeli Arabic-speaking schools exhibit a consistent gap favoring girls in math, alongside an even larger gap in language. This situation resembles the state of affairs in the Arab countries and differs from the situation in Israeli Hebrew-speaking schools and in developed countries as a group. (3) In both of Israel's language sectors and in most countries there is a consistent ranking between the various school domains in terms of gender-gap size (favoring girls), in the following order: language-science-math. Given the correlation that exists between achievements in these different domains, when one compares the math performance of boys and girls while controlling for their proficiency level in reading, boys always outperform girls. These three findings will be discussed at length below.

1. What might explain the relationship between the gender gaps in math and in language?

Using data on Israeli students, over a multi-year period and encompassing both Israeli language communities, we have demonstrated a congruity of trends regarding gender gaps in the various study subjects. We devoted particular attention to the disciplines of mathematics and native-language domains. We argued that this correspondence accords with the significant and consistent correlation that exists between gender gaps in math and reading across and within different countries (a correlation that was previously reported with regard to the PISA studies by Stoet & Geary (2013) and, before them, by other researchers – and which we reproduced on the basis of the data of PISA 2012). The hypothesis was raised that this relationship stems from, or reflects, the fact that students' math and language achievements are positively correlated in any case. The correlation between achievements in the two school domains may be attributed to a number of causes, singly or combined: (A) the existence of a causal relationship between language skills and math learning, i.e., the relationship that arises at the learning stage; (B) the existence of another (more general) factor that influences both language and math learning (and, in fact the learning of all other school subjects); (C) a psychometric artifact stemming from poor discriminant validity and high verbal loading of the math tests. More specifically:

A. **Language affects math learning.** Language skills ("native-language" as school subject) are the basis for all disciplinary learning at school, meaning that linguistic difficulty or proficiency can affect performance (for better or worse) in all other subjects, including math. There is a body of scholarship that explains how language ability affects the study of "academic" school subjects (see review by Chen, 2010), math included. Researchers have primarily been concerned by the impact on students with poor language skills (immigrants or those with reading disabilities), but students highly proficient in language may also be affected. One way or another, education professionals generally agree that if students are to fulfill their potential in math, they must be helped to acquire language skills and achieve mastery of them. For example, a report issued by Estyn, Her Majesty's Inspectorate for Education and Training(Wales, UK) calls attention to significant gender disparities favoring girls on the tests administered in both languages studied in Wales (English and Welsh), as well as on math and science exams. According to the report authors, "The most crucial factor in explaining the greater difficulty that some boys have in coping with the demands of learning and teaching in school is that fewer boys than girls acquire the level of literacy necessary to succeed […] Literacy is critical for educational success at school. Because more boys have trouble with literacy than girls they also have problems in accessing the wider curriculum. This difficulty affects progress not only in subjects that are highly language-based […], but across the whole curriculum […]" (Estyn, 2008).

   Education researchers indeed see a causal relationship between language skills and success in other subject areas (including math), in a clear direction: difficulty with language causes problems in all other subjects, and strong language skills aid in the mastery of other disciplines. Yet it should be noted that some scholars believe there to be an impact in the other direction as well. For instance Sfard (2012) argues that developing mathematical competence can contribute to linguistic competence. According to her, "mathematics can contribute to language ability in several ways: mathematical literacy enhances the ability to create complex, content-rich expressions; it raises awareness of the problem of ambiguity, and it teaches routines that increase communicative efficiency."

   Another possible means of linking language achievements with math achievements is to take a broad view of the term "mathematics." Recent years have seen growing numbers of experts in math teaching adopt an approach whereby the use and mastery of language are an inseparable part of math. This approach is reflected mainly in the sphere of mathematical literacy (see, for instance the way in which mathematical literacy is defined in the PISA study framework – OECD, 2013). From this perspective, math encompasses the need to contend with the verbal character of math literacy; that is, the need to read and understand the questions and the tasks, to explain the solution and set forth the reasoning behind it in writing. When math is defined in this broad way, student's mastery of language skills is clearly critical to their success in mathematics as well.

B. **Existence of a general cause that affects scholastic performance in all subjects.** Students' scholastic achievements in all academic disciplines, including native-language and math, are affected by personal-emotional and motivational factors (Nasser & Birenbaum, 2005; Birenbaum & Nasser, 2006). There is thus a general attribute, to be referred to hereinafter as "schooling," that encompasses factors such as student self-capacity, attitudes toward school and studies, diligence, motivation and willingness to make an effort regarding studies and tests in general, especially tests that do not have important consequences for the student, such as standardized tests[39]. Israeli and international school curricula comprise a broad array of school domains. A student's scholastic performance in all disciplines depends on his/her (positive or negative) attitude toward schoolwork, his/her investment in schoolwork and tests, and more. We cannot rule out the possibility that this schooling factor differs between boys and girls and between age levels (or, of course, between societies and countries). It is likely, for instance, that adolescent boys (Grade 8) have a lower schooling level than do girls, and take school and schoolwork less seriously, are less willing to make an effort on tests – and especially those tests that are not important (e.g. have not high stakes) from the students' perspective. This may explain, for instance, why the gaps favoring girls on the Grade 8 GEMS exams are larger in all subjects than on Grade 5 exams (or to put it another way, why the gap in favor of boys gets smaller or disappears, as it does in math in Hebrew-speaking schools). As noted above, most research on sex differences is based on the results of these kinds of large-scale tests, meaning that scholastic differences in general, and testing differences in particular, among boys and girls at different ages may be a critical factor with regard to the gender-gap issue.

C. **Math tests have poor discriminant validity.** The 1999 edition of the *Standards for Educational and Psychological Testing* states explicitly, in its

[39]In Israel, on GEMS exams and international tests the student does not receive an individual score or any personal feedback; the degree to which he/she makes an effort on these tests has to do with his/her level of interest, goodwill and sense of obligation to his/her teachers and school.

first sentence, that “for all test takers, any test that employs language is, in part, a measure of their language skills[40]” Psychometric researchers sometimes voice the concern that math tests, whether school-based or at the large-scale, system-wide level, have poor discriminant validity, arguing that they have too much loaded on verbal abilities. In other words, math tests do not measure math achievements exclusively but rather are intertwined with the test taker’s level of language proficiency (for a more in depth discussion, see Chen, 2010). Abedi & Lord (2001) argue that "the interaction between language and mathematic achievement is real" (p.232) and Chen (2010) and that and "this interaction is real for all grades, for both genders and for various ethnic groups."Like scholars in the education field, they also feel that students with poor command of language are vulnerable and might fall behind in math because their language difficulties result in ineffective learning. However, as psychometricians, they emphasize that the gap also stems from the language difficulty that they face when being tested. For example, Kieffer, Lesaux, Rivera & Francis (2009) and Wright & Li (2008), who addressed this issue from a perspective of fairness to students with language problems (e.g., immigrants or students with reading disabilities), found that students’ performance on math tests is affected by the tests’ verbal level. They showed that students who do not have complete mastery of the test language suffer from undervaluation on math tests. Similar findings have emerged from international studies. One such study (Mulis, Martin & Foy, 2013) looked at students from countries that, in 2011, participated in both the TIMSS study (Grade 4 science and math) and the PIRLS study (Grade 4 reading). The authors found that poor achieving students in reading had more trouble with verbally-loaded math and science test items (i.e.items with “high reading demands”) than they had with items characterized by low reading demands; while students with strong reading skills showed no difference in performance on items of high or low verbal load. Sato, Rabinowitz, Gallagher & Huang (2010) demonstrated that lowering the language level and using simpler language for math test items improve the grades of students with language difficulties, without affecting students whose language skills are strong.

PISA tests are known for their high verbal level, as is the portion that measures mathematical literacy. This high level is inevitable given the nature of the study. The test questions are supposed to be authentic and rooted in real life (that is, they cannot be phrased solely in formal mathematical language). In order to understand and answer them, the test taker thus needs not only math skills but, before anything else, the ability

[40]Ibid, p. 91.

to comprehend the text of the questions. A student whose reading literacy is poor will likely have trouble understanding the questions. PISA defines students with low reading proficiency (Level 1b) as students who are able only "to locate a single piece of explicitly stated information in a prominent position in a short, syntactically simple text with a familiar context and text type, such as a narrative or a simple list. The text typically provides support to the reader, such as repetition of information, pictures or familiar symbols." (OECD, 2010[41]). Because countries that perform poorly in the PISA study (including Arab countries), as well as students in Israeli Arabic-speaking schools, have a high proportion of boys with poor reading-literacy proficiencies, this may affect their mathematical literacy levels and explain the gender gap favoring girls that these countries exhibit not only in language but also in math.

It is important to note the difference between these three explanations. The two first explanations are related to the learning stage, while the third has to do with a feature of the testing stage – the tests' verbal-load level. This latter explanation is essentially psychometric, and improvable by those who compose the tests (on condition that they do not define mathematics in the broad sense of the term, that which also encompasses language abilities). Alternatively, it is possible (and even necessary) to adjust for the effect of language on math test scores when researching achievements in this discipline. Another difference is that one may expect the tests' verbal load mainly to affect students who have difficulties with language, but the explanations regarding mastery of language as the basis for learning of all other subjects, and schooling as a general trait that affects achievements in both language and math, refer to students at all levels. One way or another, whether the relationship between language and math achievements is a real one or merely a statistical artifact of the testing stage, the pedagogic implications of this relationship are important. This is because we can expect that, had a way been found to improve boys' language skills (which in all countries are weaker than those of girls), their learning would have improved in all subject areas. Their math test scores would, consequently, have improved and the gap (if any) between them and the girls would have narrowed (or widened, if the gap were already in their favor). It should be recalled that, according to the data presented in this report (Figure 11), recent years have witnessed a trend of this kind among fifth-graders in Arabic-speaking schools. Gender gaps on GEMS native-language exams gradually grew smaller; at the same time the GEMS math gaps also narrowed, and in 2013, for the first time,

[41]See also the report on Israeli PISA 2009 data at the RAMA website.

boys in Arabic-speaking schools actually outperformed their female classmates in this subject.

2. How the scholastic-achievement gender gap favoring girls in Arab societies can explained?

One finding of this study that came up repeatedly with regard to Israeli Arabic-speaking students – a finding that is also typical of some Arab countries – is that in all tested school subjects (including mathematics) girls outperform boys. Girls' superior scholastic achievements in Israeli Arabic-speaking community have been documented for two decades (see Birenbaum, Nasser & Tatsuoka, 2007; Birenbaum & Nasser, 2006; Nasser & Birenbaum, 2005 – studies based on national feedback data gathered by the education system from the mid-1990s[42]on). Disparities favoring girls have also been documented over the years in a number of Arab countries, including Jordan, the United Arab Emirates and Lebanon (El-Hassan, 2001; Green & Alkhateeb, 2001; Ridge, 2010); see also a report of the Hashemite Kingdom of Jordan Ministry of Education[43]). However, these gaps took on special prominence in the present study, when we demonstrated that disparities between students in Israeli Arabic-speaking schools are among the world's largest, and that Arab countries "lead" the world in the international TIMSS, PIRLS and PISA studies with regard to the size of scholastic gaps favoring girls in all subjects for which data were obtained.

The finding that it is precisely in Arab countries – whose culture is regarded generally as patriarchal and discriminatory toward women – that girls outperform boys of the same age and grade level in mathematics, runs counter to the relationship that has been documented in various countries between the size of the math gender gap and the degree of gender equality. The expectation is that the lower the degree of a country's gender equality, the larger will be its math gender gap (in favor of boys). But as shown by Kane & Mertz (2012), Guiso et al. (2008) and Fryer & Levitt (2010), this positive correlation is found only if Arab states are omitted from the calculations. These studies also demonstrated that disparities favoring girls in Arab countries (and in some

[42]See Abiram, Kfir and Ben-Simon, 1998.

[43]*Report on Learning Disparities between Boys and Girls in Jordan on National and International Tests of the Past Decade: The Hashemite Kingdom of Jordan* (2014). The report was a joint project of World Education and Jordan's National Center for Human Resources Development, with funding from USAID and the Jordanian Ministry of Education.

Muslim countries) are due mainly to boys' poor scholastic performance, not necessarily to girls' outstanding achievements.[44]

Senor & Singer (2009) and the report of the Hashemite Kingdom of Jordan argue that the advantage to girls in Arab countries stems from the custom of separating boys and girls in public schools. They note that in these schools men teach boys and women teach girls. Because men find teaching a less attractive occupation than do women, boys' schools are forced to employ less-qualified teachers, which negatively affects boys' scholastic performance.[45] An attempt to link girls' superior scholastic performance in Arab countries to single-sex education was also made by Fryer & Levitt (2010), but refuted later on by Kane & Mertz (2012). We should note that the existence of disparities favoring girls in Israeli Arabic-speaking schools where single-sex education is *not* the norm indicates that the reason for the gaps cannot be a difference in the teaching cadre or in the resources allocated; nor, of course, can it be differing curricula for the two gender groups.

Past studies that documented girls' advantage over boys in Israeli Arabic-speaking communities with regard to math achievements, attitudes toward math and math-related emotional issues (Ayalon, 2002; Mitelberg & Lev-Ari, 1997; 1999; Birenbaum et al., 2007; Birenbaum & Nasser, 2006; Nasser & Birenbaum, 2005) attributed girls' superior performance to socio-cultural factors in Israeli Arab society that affect boys and girls differently at the motivational and emotional levels. The main argument that these studies advanced, in varying forms, was that Israeli-Arab society is traditional-patriarchal and treats boys and girls differently, with a clear preference for boys and an expectation that girls will take on traditional domestic roles when they grow up. Thus, for Arab girls, scholastic success is nearly the only means

[44]As noted in the Introduction, one explanation that was subjected to close scrutiny by Kane & Mertz (2012) links the given country's economic level to its scholastic achievement level. Kane and Mertz summarized the results of international studies showing that the weaker a country is economically, the poorer its scores on these tests, with boys showing the greatest reduction in achievements. Stoet & Geary (2013) argue that the lower the student level (a factor related to country economic level), the greater will be the disparity favoring girls in reading, and the smaller the disparity favoring boys in math. Per their report based on PISA data, the OECD countries display larger math disparities in favor of boys than do the other PISA participant states. Figures for Israel (not published here) are similar: on both GEMS exams and international tests in which Israel participates not only were scholastic achievements found to decline along with socioeconomic level, but within the socioeconomically-disadvantaged group it was found that girls tend to reach higher scholastic achievements than boys in all study disciplines. The nature of this correlation is unclear, nor is it known why socioeconomic distress negatively affect boys' scholastic performance more than it does girls'.

[45]For a discussion of achievements by gender in the United Arab Emirates, see Ridge (2010).

of escaping the doubly-inferior social status to which females are relegated, in that they belong to a patriarchal society as well as Israel's Arab minority. It was also suggested that boys in Arab culture have greater freedom to "move about," rather than toiling away at their studies, while girls are expected to stay at home and will naturally prefer to work on their school assignments than to engage in customary housekeeping tasks. However, these explanations have stayed within the realm of researcher hypothesis or common knowledge among education professionals, without having been seriously investigated in depth; it is, therefore, hard to determine the degree to which they are currently valid. Similar arguments were raised in the Jordanian gender-gap report[43].

In the meantime, another factor that has not been posited by any study, despite the fact that it is a common feature of all Arab countries and of Israeli-Arab society, is Arabic's status as native language and instructional language in schools. Arabic has certain unique qualities that could potentially affect boys and girls differently. Given the importance of language as a basis for learning the various school domains, and the relationship that we have noted between language-proficiency level and math achievements, it appears justifiable to consider the idea that the exceptionally large gap (compared with the average gaps in the international studies) in Arabic as a native-language between girls and boys is related to (or perhaps even the cause of) the disparity favoring girls in math. But if this is the case, then why is there such a large disparity in Arabic language, being a native language between boys and girls at schools?

Arabic diglossia and its possible connection to boys' poor scholastic performance

Ferguson (1959) explains that Arabic is a diglossic language characterized by two different modalities: literary (classical) Arabic – the language used in formal settings including the education system; and spoken Arabic – the language of everyday communication (generally oral). Arabic's two modalities are, in fact, quite distinct linguistically; they differ in terms of vocabulary, phonology, syntax and grammar. Children use spoken Arabic at home; they generally encounter literary Arabic only when they start school, and they experience it almost as a second language (though they may be exposed to it earlier if their parents read aloud to them). Arabic diglossia is recognized today as a major cause of poor language achievements in most Arabic-speaking countries, at all rungs of the socioeconomic ladder (Maamouri, 1998), and as a factor that disrupts the acquisition of reading skills (Saiegh-Haddad, 2003). Reading achievements in Arab countries, as measured by PIRLS studies, are among the lowest of all participating countries (see the PISA and TIMSS reports). It has been conjectured in Israel as well that diglossia is the main reason for Arabic-speaking students' lower achievement in reading skills – and not necessarily student socioeconomic background (Zuzovsky, 2010). But in the present context, the question to be asked is whether diglossia poses a greater problem for boys than for girls in terms of language skills acquisition and performance.

As we demonstrated in the Introduction, boys develop language skills more slowly in all cultures, and the gap in their disfavor with regard to reading and writing emerges in the early elementary-school years, in all languages (see the GEMS data on native-language gender gaps in Grade 2, and fourth-grade girls' superior reading achievements in the PIRLS study in most countries). This being the case, it is not inconceivable that because of the diglossia that characterizes Arabic Arab boys would be more vulnerable to the difficulties it create, especially at the early stages of acquiring basic reading/writing skills, which take place at a young age. Because reading and writing are the foundation for all school-based learning, and because acquiring these skills is the main objective at the lower elementary grades, failure or difficulty at this early stage in reading/writing acquisition is potentially detrimental not only to language-acquisition itself, but also to schoolchildren's emotional and motivational status (e.g. boys' self-perception and sense of self-efficacy). That is, difficulty in this area at such a critical stage of learning may cause boys to develop, early on, a lower level of "schooling" than do girls, and poorer language skills. These problems are liable to afflict boys throughout their entire school career and may ultimately compromise their achievements compared with girls – even in the school domain that they find easiest – mathematics (we will have more to say about the domain ranking below).

In order to assess the validity of this explanation, additional research must be carried out on students in Arabic-speaking schools, focusing on the issue of whether boys in this language sector have more trouble than girls in acquiring reading and writing skills at the very beginning of school, and whether, in consequence, they tend to develop poorer schooling (e.g., poor self-concept), which may continue to afflict them later on. Positive answers to these questions may have far-reaching pedagogic implications regarding how basic language skills should be taught and what kind of interventions should be used with boys in the early stages of reading acquisition. One way or another, the research literature on gender differences has preferred to focus on the issue of the boy-girl gap (favoring boys) in math, and has not adequately addressed the boy-girl gap (to boys' disadvantage) in language, or the way in which this gap may affect them. The present study brings into sharper focus the conclusion explicitly affirmed by Stoet & Geary (2013), namely, the need to invest in boys' reading skills, especially in the early years. This is because "Sex differences in reading are not only persistent and growing, they are particularly large for the most vulnerable boys at the bottom of the reading performance continuum."

3. The consistent ranking between subject areas and the possible connection between it and the stereotype of boys outperforming girls in math

The present study's most important finding appears to be that in Israel (in both language sectors), as elsewhere (countries for which documentation exists), there is a consistent ordered ranking between various school domains in terms of gender-gap magnitude. Thus, a comparison of gender gaps in different subjects revealed that, regarding language in its broader form (verbal abilities, reading literacy and reading comprehension), girls almost always greatly outperform boys. By contrast, in math, boys' achievements are usually higher than those of girls (or the gender gap favoring girls in math is smaller than in other subjects). In the middle, between language and math, lies science – a study discipline that draws on skills from both language and math. In science the gender gaps are negligible or balanced – sometimes favoring girls and sometimes favoring boys. This constant subject-area ordered ranking– language-science-math – characterizes the gaps found in the various test years (GEMS exams), across different grade levels and in different countries. This finding is documented not only for large-scale standardized tests (national and international, e.g. PISA), but also for school grades (Voyer & Voyer, 2014).

The meaning of this fixed hierarchy is that, relative to girls, math is the subject in which boys reach the highest achievements (even if their performance in math is inferior to that of their female peers). In other words: from the boys' point of view, math is their "best" subject, while for girls language is the area where they are strongest. These perceptions hold even if we believe that there are other variables (such as a general schooling trait) that similarly affect boys' (or girls') achievements in all subjects.

It could be argued that this language-science-math ranking endures in different situations due to the existence of a correlation between gap size (and between achievements) in the various disciplines, as explained above. After all, if girls, for instance, improve in one subject, they will most likely improve in other school subjects as well (math include), and vice versa. Alternatively, it may be argued that the ranking's durability across different situations is what creates the correlation. Another expression – or, perhaps, an outcome – both of this ranking between the subject matters and of the correlation between the gender gaps in them, is that when the math achievements of boys and girls at the same level of reading proficiency are compared, boys always outperform girls. It is interesting that this disparity in their favor emerges even when, on average, there is no gender gap in math (a situation that, for instance, characterizes the gender-equal Scandinavian countries), and even when on average there is a math gap in favor of girls (e.g. in Arab countries and in Israeli Arabic-speaking schools).

The invariability of this latter finding and, in any case, of the finding with regard to the ranking among the school domains, conflicts with the variability of the gender gaps in math as documented in the literature. In the Introduction to the present study it was explained that this variability as such is sometimes used to rule out any biological explanation for boys' advantage over girls in math. The present study's findings undermine the validity of using the math gender gap's heterogeneity to rule out a biological model and to support an environmental model. If in all situations, even when girls close gaps (on average) vis-à-vis boys in math or even surpass them, math still remains boys' strongest subject, it is important that we understand why. If boys at a given level of reading proficiency have higher achievements in math than girls at the same reading level, we need to find out what is special about this subject area for boys as opposed to girls. Ultimately, perhaps, the variability of math gaps, as emphasized by Kane & Mertz (2012), can be explained by the disparities that exist in various countries between boys and girls who differ in their language ability and/or overall motivational-scholastic characteristics. Thus, in countries where the math gender gap was closed, girls also displayed a general improvement in their scholastic achievements, including those in language (an argument that had been raised long before, Guiso et al., 2008[46]). Given the aforementioned findings, it appears that researchers studying gender gaps in math will need, from now on, to look at the issue in a broader school context. As Chen (2010) has noted, we can no longer ignore the relationship between language achievements and math achievements (whatever the explanation for their existence may be), and when we investigate disparities between different population groups, we must also take into account language achievements (data that, like math performance data, are generally obtained via large-scale tests).

Stoet & Geary (2013) presented new findings that are of value to a discussion of the possible impact of environmental-social conditions on math disparities and on the nature of that impact. According to them, girls have trouble closing gaps in math precisely at the upper end of the ability scale, where boys outnumber girls and where the gender gap favoring boys is largest – a finding that is consistent with the findings of PISA 2012 and with the larger proportion of high-performing boys in math competitions and in institutions of higher education. Another finding presented by Stoet & Geary is that the trend toward a narrowing of gender gaps in the upper range of ability – a trend that had been noted in the relevant professional literature – has halted in recent years.

[46]For more on this and on the counterarguments mounted by Stoet & Geary, 2013, see the Introduction.

Finally, the present study's findings offer a new answer to the disturbing question of why the stereotype of boys outperforming girls in math still prevails, despite the many studies showing that the stereotype lacks empirical support (Else-Quest et al., 2010).

Student self-concept (or belief in their own abilities) and social attitudes exist within a multidisciplinary school environment. Because human perceptions are relative, so are students' self-concepts about whether they are good at a particular subject – and so, accordingly, will their self-image as learners develop. When students in the process of establishing a self-image as learners ascribe to themselves competency (high or low) in a specific subject, they do it within a broader school context. They most likely look at their achievements in all of the other subjects. Even when they fall behind the girls in math, most boys experience math as a subject that, relative to themselves, is easier for them[47] and in which they are more successful. And, conversely: girls succeed at math – relative to themselves and even, at times, in absolute terms – to a lesser degree than in language-intensive subjects (especially native language). This may cause them to perceive math as a harder subject for them than it really is. Clearly, this might make them like the subject less, impair their self-concept as "good" math students, and instill anxiety or negative feelings toward math (see OECD 2013). Parents and instructors who teach and grade them are also subject to this perceptual bias and form similar impressions. They can be expected to encourage boys in math because that is the subject in which boys do relatively well (even in the lower elementary grades); by contrast, they will expect girls to make less effort in math because there are other subjects in which they can excel. Moreover, in light of the data repeatedly documenting boys' greater presence among high achievers in math and at the upper end of the ability scale (Stoet & Geary, 2013; OECD, 2013), there is a basis for assuming that boys will be more numerous among outstanding math students, not only at the countrywide level but at the school level as well. This in turn will be reflected in boys' and girls' self-image and reinforce the stereotypical view held by male and female students and teachers that "math is a boys' subject." In this circular manner the prevailing stereotype that boys are better than girls at math is perpetuated.

These far-reaching conjectures should, of course, be investigated further. In the meantime, we need to be aware that the socialization process described in the literature starts at the very earliest stages of primary education (and it makes no difference whether boys and girls learn together or separately, since

[47]In this study we did not deal with the affective aspects of math learning among girls and boys. An extensive literature and empirical evidence show that boys suffer less math anxiety, believe in their self-capacity more than girls, and so forth. See, for example, the international 2012 PISA report.

they always learn both math and native-language skills). This process appears to cause girls to devote less scholastic effort to (and to be less successful in) mathematical-quantitative disciplines, with the consequence that later in life they will be less likely to pursue advanced study or careers in fields that require quantitative thinking. If we want a more egalitarian society in which these kinds of stereotypes no longer prevail, we need to clearly understand the source of such problematic attitudes and try to combat them where possible. If we are to persuade girls that they can be just as good at math as boys (and it is irrelevant whether boys' advantage in math is biologically-based or not), it is important that we help them improve their math achievements; at the same time, we need to help boys do better with their language capabilities. As we have seen, cultivating boys' language abilities could help them realize their potential in math while also loosening the hold of stereotypical attitudes – in society at large and among the boys themselves. This kind of approach would foster more egalitarian and less stereotypical perceptions of boys' and girls' scholastic capabilities.

References

(Hebrew)

Abiram T., Kfir, R. and Ben-Simon, A (1996). ***National Feedback on the Education System: Mathematics – Grade 8***. Ministry of Education, Culture and Sport, Office of the Chief Scientist and National Institute for Testing and Evaluation, Jerusalem.

Sfard, A. (2012). "Linguistic and Mathematical Literacy – What is the Connection?" in: ***Between Language and Disciplinary Literacy: Workshop Report***," Pollak, I. (ed.). Initiative for Applied Education Research, Israeli Academy of Sciences and Humanities.

Rapp, J., Notea-Koren, E., Ron-Kaplan, I., Gelbart, H., Awadyeh, I., Rogel, A. (2013). ***Mathematics Achievement Gaps between Boys and Girls on Israeli Standardized Tests***. Ninth Conference of the Israel Psychometric Association (IPA), Jerusalem.

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*, 219-234.

AERA, APA, NCME. (1999). *Standards for Educational and Psychological Testing*.Washington, DC: AERA.

Aiken, L. R. (1971). Verbal factors and mathematics learning: A review of research. *Journal for Research in Mathematics Education, 2*(4), 304–313.

Aiken, L. R. (1972). Language factors in learning mathematics. *Review of Education Research, 42*, 359–385.

Ayalon, H. (2002). Mathematics and sciences course taking among Arab students in Israel: A case of unexpected gender equality. *Educational Evaluation and Policy Analysis*, 24: 63-80.

Baron-Cohen, S. (2003). *The Essential Difference: Men, Women, and the Extreme Male Brain.* Allen Lane, London.

Berenbaum, S. A., Martin, C. L, Hanish, L. D., Briggs, P. T. & Fabes, R. A. (2008). Sex differences in children's play. In: Becker, J. B., Berkley, K. J., Geary, N., Hampson, E., Herman, J. P. & Young, E. A, (editors). *Sex differences in the brain: from genes to behavior*. New York: Oxford University Press;

Brody, L., & Mills, C. (2005). Talent search research: What have we learned*? High Ability Studies,16*, 97-111.

Burrelli, J. (2008). *Thirty-Three Years of Women in S&E Faculty Positions, InfoBrief, Science Resources Statistics, NSF 08-308, National Science Foundation Directorate* *for Social, Behavioral, and Economic Sciences*. Retrieved from: http://www.nsf.gov/statistics/infbrief/nsf08308/nsf08308.pdf.

Birenbaum, M., &Nasser, F. (2006). Ethnic and gender differences in mathematics achievement and in dispositions towards the study of mathematics. *Learning and Instruction, 16*, 26-40.

Birenbaum, M.,&Nasser, F. (1994). On the relationship between test anxiety and test performance. *Measurement and Evaluation in Counseling and Development, 27*(1), 293-301.

Birenbaum , M., Nasser, F. &Tatsuoka, C. (2007). Effects of gender and ethnicity on fourth graders' knowledge states in Mathematics.*Journal of Mathematical Education in Science and Technology, 38*(3), 301-319.

Bowers, J. M., Perez-Pouchoulen, M., Edwards, N.S., &McCarthy, M.M, (2013). Foxp2 mediates sex differences in ultrasonic vocalization by rat pups and directs order of maternal retrieval.*The Journal of Neuroscience, 33*(8), 3276-3283.

Ceci S. J., &Williams W. M. (2010). *The Mathematics of Sex: How Biology and Society Conspire to Limit Talented Women and Girls*. Oxford University Press.

Chen, F. (2010). *Differential Language Influence on Math Achievement.* Ph.D. Dissertation, The University of North Carolina at Greensboro. Retrieved from: http://libres.uncg.edu/ir/uncg/f/Chen_uncg_0154D_10511.pdf.

Cole, N. (1997). *The ETS gender study: How females and males perform in educational settings.* Princeton, NJ: Educational Testing Service.

del Pero, A. S., &Bytchkova A. (2013), *A Bird's Eye View of Gender Differences in Education in OECD Countries,*OECD Social, Employment and Migration.Working Papers, No. 149, OECD Publishing. http://dx.doi.org/10.1787/5k40k706tmtb-en

Eagly H.A., &Wood W. (2013). The Nature–Nurture Debates: 25 Years of Challenges in Understanding the Psychology of Gender.*Perspectives on Psychological Science 8*(3), 340–357.

El-Hassan, K. (2001). Gender issues in achievement in Lebanon. *Social Behavior and Personality, 29*(2), 113-124.

Ellis, H. (1894). *Men and Women: A Study of Secondary and Tertiary Sexual Character.* Heinemann, London.

Ellison, G., & Swanson, A. (2010). The gender gap in secondary school mathematics at high achievement levels: Evidence from the American Mathematics Competitions, *The Journal of Economic Perspectives, 24*, 109-128.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis. *Psychological Bulletin*, *136*(1), 103–127.

Estyn (2008).*Closing the gap between boys' and girls' achievement in schools. Her Majesty's Inspectorate for Education and Training in Wales*. Retrievedfrom: http://www.estyn.gov.uk/publications/Gender_Gap_Report_March_2008.pdf.

Ferguson, C. A. (1959). Diglossia. *Word, 14*, 47–56.

Fryer R.G., &Levitt S.D.,(2010). An Empirical analysis of the Gender Gap in Mathematics. *American Economic Journal: Applied Economics, 2*(2010): 210-240.

Gallagher, A.M., De Lisi, R., Holst, PC, McGillicuddy-De Lisi, A. V., Morely, M., & Calahan, C. (2000). Gender differences in advanced mathematical problem solving. *Journal of Experimental Child Psychology, 75*, 165-190.

Geary, D. C., (1996). Sexual selection and sex differences in mathematical abilities. *Behavioral and Brain Sciences, 19*, 229-247.

Geary, D.C. (2010). *Male, Female: The Evolution of Human Sex Differences, 2nd edition.* Washington DC: American Psychological Association.

Goetz, T., Bieg, M., Ludtke, O., Pekrun, R. H., & Hall, N. C. (2013). Do girls really experience more mathematics anxiety? Conflicting evidence from trait vs. state perspectives. *Psychological Science, 24*(10), 2079-2087.

Green, B., & Alkhateeb, H. M. (2001). Gender differences in mathematics achievement among high school students in the United Arab Emirates (1991-2000). *School Science and Mathematics, 101*(1), 5-9.

Guiso, L., Monte, F., Sapienza, P., &Zingales L. (2008). *Science, 320*, 1164-1165.

Haladyna, T.M., & Downing, S.M. (2004). Construct-irrelevant variance in high-stake testing. *Educational Measurement: Issues and Practice, 23*(1), 17-27.

Harasty, J., Double, K. L., Halliday, G.M., Kril, J.J., &McRitchie, D.A. (1997). Language-associated cortical regions are proportionally larger in the female brain. *Archives of Neurology. 54*(2),171-176.

Hausmann, R., Tyson, L. D., &Zahidi, S. (2006). *The Global Gender Gap Report, World Economic Forum*. Geneva, Switzerland.

Hyde, J. S., Fennema, E., & Lamon, S. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *(107)*2, 139-155.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A., & Williams, C. (2008). Gender similarities characterize math performance*, Science, 321*, 494–495.

Hyde, J.S., & Linn, M.C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104*, 53-69.

Kane, J. M., & Mertz, J. E. (2012). Debunking Myths about Gender and Mathematics.*Notices of the AMS, 59*, 10–21.

Kieffer, M.J., Lesaux, N.K., Rivera, M., &Francis, D.J. (2009). Accommodations for English language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research, 79*(3), 1168-1201.

Kimura, D. (1999). *Sex and Cognition*. MIT Press, Cambridge MA.

Lietz, P. (2006). A meta-analysis of gender differences in reading achievement at the secondary school level. *Studies in Educational Evaluation,32*, 317–344.

Lindberg, S.M., Hyde, J. S., Peterson, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis, *Psychological bulletin,136*, 1123-1135.

Maamouri, M. (1998). *Language education and human development: Arabic diglossia and its impact on the quality of education in the Arab region*. Discussion paper prepared for the World Bank Mediterranean Development ForumMarrakesh, Philadelphia: University of Pennsylvania International Literacy Institute.

Maccoby, E. E., & Jacklin, C.N. (1974). *The Psychology of Sex Differences.* Stanford University Press: Stanford, California.

Machin, S., & Pekkarinen, T. (2008). Global sex differences in test score variability, *Science 322,* 1331-1332.

Martin, M.O., Mullis, I.V.S., Foy, P. and Stanco, G.M. (2012). TIMSS 2011 International

Results in Science. Chestnut Hill, MA: Boston College, TIMSS and PIRLS

International Study Center

Mittelberg, D., &Lev Ari, L. (1999). Confidence in mathematics and its consequences: Gender differences among Israeli Jewish and Arab youth. *Gender and Education 11*, 75-92.

Mittelberg, D., &Lev Ari, L. (1997). Gender differences in mathematics among Jewish and Arab youth in Israel. *Mathematics Education Research Journal, 9*(3), 347-351.

Mosconi, N.(2001). Comment les pratiques enseignantes fabriquent de l'inégalité entre lessexes. *Les Dossiers des Sciences de l'Education*, *5* (in French).

Mullis, I.V.S., Martin, M.O., Foy, P., & Drucker, K.T., (2012). *PIRLS 2011 International Results in Reading.* Chestnut Hill, MA: Boston College, TIMSS and PIRLS International Study Center.

Mullis, I.V.S.,Martin, M.O., Foy, P. & Arora, A. (2012). TIMSS 2011 International Results in Mathematics. Chestnut Hill, MA: Boston College, TIMSS and PIRLS International Study Center.

Mullis, I.V.S., Martin, M.O.,& Foy,P. (2013). The impact of reading ability on TIMSS mathematics and science achievement at the fourth grade: an analysis by item treading demands. In *TIMSS and PIRLS 2011: Relationships among Reading, Mathematics and Science Achievement at the Fourth Grade - Implications for Early Learning.* Michael O. Martin and Ina V.S. Mullis, Editors. TIMSS and PIRLS International Study Center. Lynch school of education, Boston College and International Association for the Evaluation of educational Achievment (IEA).

Nasser F., &Birenbaum, M. (2005). Modeling mathematics achievement of Jewish and Arab eighth graders in Israel: The effects of learner-related variables. *Educational Research and Evaluation, 11*(3), 277-302.

Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives, 24*(2), 129-144.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B., Wiers, R. W., Somogyi, M., Akrami, N., Ekehammar, B., Vianello, M., Banaji, M. R., & Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences, 106*, 10593-10597.

OECD (2010).*PISA 2009 Results: what students know and can do – Students performance in Reading, Mathematics, and Science*. PISA, OECD Publishing.

OECD (2013).*PISA 2012 results : what students know and can do – Students performance in Mathematics, Reading, and Science.* PISA, OECD Publishing.

OECD (2004). *Learning for Tomorrow's World: First Results from PISA 2003.* Retrieved from http://www.oecd.org/pages/0,3417,en_32252351_32236173_1_1_1_1_1,00.html.

Preckel, F., Goetz, T., Pekrun, R.,& Kleine, M. (2008). Gender differences in gifted and average-ability students: Comparing girl's and boy's achievement, self-concept, interest, and motivation in mathematics. *Gifted Child Quarterly, 52*(2), 146-159.

Ridge N. (2010). *Teacher Quality, Gender and Nationality in the United Arab Emirates: A Crisis for Boys.* Working paper 10-06 . Dubai School of Government.

Saiegh-Haddad E. (2003). Linguistic distance and initial reading Acquisition: The case of Arabic diglossia. *Applied Psycholinguistics, 24*, 431-451.

Sato E., Rabinowitz S., Gallagher C.,&Huang C.W. (2010). *Accommodations for EnglishLanguage Learner Students: The Effect of Linguistic Modification ofMath Test Item Sets*. NCEE 2009-4079. U.S. Department of Education.

Secada,W.G. (1992). Race, ethnicity, social class, language, and achievement inmathematics. In D. A. Grouws (Ed.), *Handbook of Research on MathematicsTeaching and Learning* (pp. 623–660). New York: Macmillan.

Sells W. L., (1973).*Sex Differences in Graduate School Survival*. ERIC Document Reproduction Service.

Senor, D., &Singer, S. (2009*). Start-up Nation: The Story of Israel's Economic Miracle .* Twelve, Hachete book group. New York.

Shields, S. A. (1982). The variability hypothesis: The history of a biological model of sex differences in intelligence. *Journal of Women in Culture and Society, 7*, 769–797.

Spelke, E.S. (2005). Sex differences in intrinsic aptitude for mathematics and science: A critical review*. American Psychologist, 60*(9), 950-958.

Stoet, G., &Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: within- and across-nation assessment of 10 years of PISA data. *PLoS ONE* 8(3): e57988. Retrieved from: doi:10.1371/journal.pone.0057988

The Hashemite Kingdom of Jordan; Education Reform for Knowledge Economy II (2014). *Gender Gap in Student Achievement in Jordan, Study Report*. (unpublished).

Voyer, D., & Voyer, S. D. (2014, April 28). Gender Differences in Scholastic Achievement: A Meta-Analysis. *Psychological Bulletin.* Advance online publication. Retrieved from: http://dx.doi.org/10.1037/a0036620

Wright, W.E., & Li, X. (2008). High-stake math tests: How No Child Left Behind leaves newcomer English language learners behind. *Language Policy, 7*, 237-266.

Zuzovsky, R. (2010). The impact of socioeconomic versus linguistic factors on achievement gaps between Hebrew-speaking and Arabic-speaking students in Israel in reading literacy and in mathematics and science achievements. *Studies in Educational Evaluation, 36*, 153-161.