THE HASHEMITE KINGDOM OF JORDAN

EDUCATION REFORM FOR KNOWLEDGE ECONOMY II (ERfKE II)

Gender Gap in Student Achievement in Jordan Study Report

Monitoring & Evaluation Partnership (MEP) Project*

Prepared by:

Dr. Ahmad Tweissi
Dr. Khattab Abu Lebdih
Dr. Imad Ababneh
Dr. Sheren Hamed

August 2014

NCHRD Publication Series No. (179)

* This report is the product of a collaboration between National Center for Human Resources Development (NCHRD) and World Education, Inc. (WEI) researchers under the Monitoring & Evaluation Partnership (MEP) project. MEP is a five-year (2010-2015) USAID-funded project implemented by World Education. Its goal has been to strengthen the technical capacity of NCHRD and to provide financial support for a series of program quality evaluations for the Government of Jordan's Education Reform for Knowledge Economy (ERfKE II) program.

ACKNOWLEDGEMENTS

This study was conducted by the research team working under the Monitoring and Evaluation Partnership (MEP) Program, which is supported by USAID in Jordan, and implemented by World Education, Inc. in the U.S. and the National Center for Human Resource Development in Jordan. The authors wish to acknowledge the important contributions made by Jordan's Ministry of Education (MoE). In particular, we are very thankful to the staff members from the Gender Department within the Directorate of Planning and Educational Research, the selected field directorates, schools, principals, teachers and students who provided invaluable information about their schools and the education system in Jordan as a whole. We would especially like to express our sincere gratitude to Nour Abu Al-Ragheb, USAID Education Specialist, and Eng. Firyal Aqel, Director of the Donor Coordination Unit, for their unwavering support and coordination. Without their support this study would not have been completed.

ACRONYMS

ANOVA	Analysis of Variance
DET	Department of Examinations and Tests
ERfKE	Education Reform for Knowledge Economy
FGD	Focus group discussion
HLM	Hierarchical Linear Model
IEA	International Association for the Evaluation of Educational Achievement
M&E	Monitoring and Evaluation
MEP	Monitoring and Evaluation Partnership project
MoE	Ministry of Education
NAfKE	National Assessment for the Knowledge Economy
NCHRD	National Center for Human Resource Development
NT	National Test (Jordan)
OECD	Organization for Economic Cooperation and Development (OECD)
PISA	Program for International Student Assessment
SES	Socio-economic status
TIMSS	Trends in Mathematics and Science Study
UNRWA	United Nations Relief and Works Agency

CONTENTS

EXECUTIVE SUMMARY

This study, which is entitled "The Gender Gap in Student Achievement in Jordan," is intended to contribute to the overall mission goal of the Education Reform for Knowledge Economy (ERfKEII) by identifying student and school level factors, such as school, teacher, and student characteristics that might affect male and female student achievement differently. Moreover, it provides information that will directly inform national and school-based reform plans designed to ensure quality education for all students in Jordan, irrespective of gender.

Study Objectives

The study focuses on three main objectives:
1) Determining the magnitude and trend in the achievement gap between males and females in TIMSS, PISA, NAfKE, and National Test (NT) over time[1];
2) Investigating the factors associated with the gender gap in different assessments over time, and understanding how those factors operate; and
3) Based on the findings, proposing policy options that will help narrow the gender gap in student achievement.

Methodology

Quantitative and qualitative research methods were utilized to ascertain the gender gap in student assessment and its related factors in other assessments. The quantitative analyses strategies consisted of: 1) comparative analysis of male and female scores in TIMSS (2003, 2007, 2011), NafKE (2006, 2008 and 2011), PISA (2006, 2009, 2012), and NT (2007, 2011, 2013) datasets[2]. Overall scores in Math, Science, Arabic, and English were compared by gender and grade over time. 2) analyses of school, and teacher background and school variables from NafKE (2011), TIMSS (2011), and PISA (2012) data sets. The research team employed the Hierarchical Linear Model (HLM) to ascertain the factors associated with the gender gap in achievement.

The qualitative component comprised focus group discussions with teachers, parents, principals, supervisors, and students in order to understand what they perceived to be the reasons for the gender gap and how the identified factors contribute to differences in achievement between males and

[1] Two previous reports have addressed the gender achievement gap in early grades (EGRA/EGMA) and among older students (Tawjihii exam). Therefore, analysis of those tests results were not be carried out and discussed in this report. For more details, see 1) *Student Performance in Reading and Mathematics, Pedagogic Practice, and School Management in Jordan* (2012) by RTI International; and 2) *Gender Gap in Tawjihii* (unpublished) by Gender Department, Ministry of Education, Jordan.

[2] Regarding the NT, it was not possible to build a model, as the dataset does not contain detailed student, school, and/or family background data.

females. In addition, focus group participants answered a questionnaire about their opinions regarding school, parental attitudes, and cultural perceptions/expectations. The results from the questionnaire tabulation were used to support the FGD notes.

Findings

Up until grade 3, most females and males study in co-ed classes housed in female government schools. However, after grade 3, most Jordanian public schools are single-sex. Female schools have only female teachers and principals; the opposite is true for all male schools. The results presented here confirm that female students have been outperforming males in TIMSS, NAfKE, PISA, and NT in all subjects and grades for several assessment cycles. We observe that the mean score differences and the effect sizes for Math are smaller than for Science and Arabic. In addition, we learned that the gap is more pronounced in government schools than non-government schools, and as students move up through the education system and reach higher grades, the gender gap tends to increase. However, despite the pervasiveness of the gender gap in student achievement phenomenon in Jordan, we cannot say that the magnitude of the gender gap is increasing consistently over time. We observe, for example, that the gap in TIMSS math scores decreased from 2003 to 2007 (27 and 21 points difference, respectively). The gap then increased again in 2011 (28 points).

Regarding the factors associated with students' performance and the gender gap, we can conclude there are both student and school-level factors associated with students' achievement and the differences between the sexes. First, we notice that at the descriptive level, students, teachers, and principals in female schools reported behaviors and attitudes more conducive to learning than their counterparts in male schools. For example, female students were more likely to do homework, and have higher academic ambitions, repeat grades less frequently, and skip classes less often than male students. However, it is important to highlight that those positive behaviors and academic ambition on the part of girls does not translate into or lead to higher ambition to participate in the labor market. Conversely, males were more likely to receive out-of-school lessons (probably as a remedial intervention) and have less parental involvement in their education than females.

At the school level, differences between male and female schools were also apparent. Female teachers reported higher job satisfaction than male teachers, and female schools were reported to be safer than male schools. In addition, female schools seemed to have more qualified teachers to teach subject-specific classes, better student-teacher relations, less teaching limitations due to disruptive students, and more supervision of homework. Finally, male principals reported fewer resources in their schools than female principals. Female principals reported closer supervision of teachers, less teacher turnover, and higher parental participation.

The HLM models for TIMSS, PISA, and NAfKE further revealed that some factors are key predictors of males' and females' achievement (controlling for the other factors). More precisely, the models revealed that at the school level, school gender is an important predictor of achievement in all assessments. In other words, school gender is significantly and positively related to mean achievement.

Female schools score, on average, higher than male schools. School SES is also a significant predictor of achievement for both males and females. At the student level, the following factors are positively associated with higher achievement: 1) high academic ambitions, i.e., the goal of attending university and/or beyond; 2) parents who make sure the student sets time aside for homework; 3) frequent use of a computer at home, or general use of a computer for drawing, painting, using graphics program, and for networking; 4) the feeling of being safe in school; and 5) a strong sense of belonging (i.e., students are not left out of school activities and/or do not feel lonely or "out-of-place"). On the other hand, having a history of repetition was adversely associated with students' performance.

Despite the findings highlighting the gender gap and its underlying causes, it is important to emphasize that Jordanian students (both males and females) are performing below the average (center-point) in national and international assessments. That is particularly true in Math and Science. Therefore, we suggest that, when it develops educational policies for the Kingdom to address the gender gap, the MoE promote the improvement of education for all children and not focus solely on males. More specifically, action plans should not reallocate resources from female schools to male schools. In fact, while the MoE maintains and enhances female schools in a continuous basis, it must heavily invest in male schools to accelerate males' performance improvement. If they are followed, the options presented below will have a trickle-down effect that has the potential to improve the education system for all.

Recommended Options for Improving School Performance and Decreasing the Gender Gap in Academic Achievement:

1. ***Invest in teachers***. It will be necessary to (re)shape teaching profession requirements and career path options in order to attract talented individuals of both sexes and to stimulate current teachers to perform at high levels. Unless the teaching profession becomes more rewarding, not just financially, but also by offering a better work environment and career path, there will be no incentives for talented individuals to enter the profession or for current teachers to perform at the desired levels aimed by ERfKE II's stated levels or any educational reform project. Currently, male teachers are forced to hold multiple jobs to fulfill their social and financial obligations and that affects the time they can spend on preparing lessons and meeting the academic needs of male students. To improve the status of the teaching profession to attract and retain committed individuals, particularly males, we recommend several key measures, including.[3] a) Engaging in serious dialogue that will promote strategies to provide financial incentives to all teachers so both males and females can embrace the teaching profession as their primary (or sole) source of income. Financial incentives must be combined with accountability systems to ensure results. Teachers should be ranked according to their performance in the classroom; b) Enhancing the status of the teaching profession by utilizing the media to change teacher´s image in society; c) Offering

[3] These suggestions were presented in detail in the McKinsey *Report How the World Best-Performing School Systems Come Out on Top* (2207) by Mona Mourshed, as essential measures undertaken by high-performing school systems around the world.

scaffolding for low-skilled teachers through scripted lesson plans and coaching on curriculum; d) Strengthening pre- and in-service training and finding appropriate times to conduct those trainings; e) Promoting peer-led learning for teachers; and f) activating the role of supervisors as coaches. Although that has been a key strategy under the School and Directorate Development Program (SDDP), much remains to be done with regards to increasing the number of supervisors and training them effectively. Further, supervisors have little incentive to supervise male schools.

2. ***Invest in principals***. A strategic plan to select principals should also be put in place. Ideally, principals should be selected from a pool of excellent teachers or supervisors in the system and be trained accordingly. Best principals should be assigned to most challenging male schools and receive adequate compensation for working in difficult environments. An accountability system with incentives mechanisms should be in place to encourage leaders to perform at their best. Further, principals should share administrative tasks with a vice-principal (or equivalent), so s/he can supervise teachers to improve instruction.

3. ***Improve overall safety in schools***. Safety is a major concern in male schools. Educational environments where students, teachers, and administrators feel unsafe decrease their motivation to learn and work. Therefore, it will be essential to address this issue in a multifaceted manner in order to:

   (a) Increase awareness about current school-based initiatives to deter school violence in Jordan. Currently, some male schools have been experimenting with strategies to involve parents and communities in school life in an attempt to make schools safer and to improve the overall quality of education. Those experiences have been yielding promising results[4] by emphasizing family involvement and educating parents about values of a male student. Some schools in Jordan have held workshops that adopted an Islamic approach to school violence and vandalism where they educate both the parents and students on the values of school preservation and respect to the teacher. The MoE might benefit from such local initiatives by promoting a national forum for shared experiences about improving safety in schools;

   (b) Activate councils at the central and directorate levels to search for innovative solutions and exchange information regularly on how to tackle the safety problem. Under ERfKE II Component 1, the MoE established educational councils at the school cluster and directorate levels. In addition, parents´ and students´ councils have also been established. However, not all councils are in fact active. One key strategy adopted by the councils should be to listen to and include children's views, perspectives, and experiences before developing ways to deter violence. A reward system could be implemented to recognize successful initiative to decrease violence. The committees would suggest

[4] Alaref, J., Linnemann, H., Quota, M., & Woolcock, M. (2014). Jordan Education Case Study, for MENA Regional Report. The World Bank.

and approve a general framework for the strategies and mechanisms that are expected to contribute to the improvement of the overall safety of schools. The MoE would have to visit schools often to ensure councils are active.

(c) Ensure there is appropriate legislation that will safeguard the rights of victims (teachers, administrators, and students);

(d) Create an accountability system in cooperation with schools, communities, and law enforcement;

(e) Involve relevant institutions (such as the judicial system, social services, and/or Islamic religious authorities) in forming a partnership with schools and supporting families and communities when violence extends beyond the school premises;

(f) Coordinate efforts with existing programs (funded locally or by international agencies) to avoid duplication of efforts and to enhance the effects of the initiatives;

(g) Equip teachers, principals, and supervisors with relevant competencies and decision-making mechanisms to deal with violence in schools. Both capacity building and development of mechanisms to deal with violence should be included in the school development and school action plans.

(h) Activate the role of school counselors to offer psychological support to students and parents and to generate reflection about the roots of violence and how students and parents may react to it.

(i) Increase awareness about violence by utilizing the media, publications inside the school, and university radio.

4. ***Make education more relevant to the needs of students***. Based on the findings presented in this report it is clear that parents, teachers, and students perceive the returns of education for males and females differently, due to existing socio-cultural and economic factors. However, we strongly believe that education is a critical variable for enabling males and females to gain the skills necessary to meet the demands of the labor market, no matter what profession they choose. Therefore, it becomes necessary to adapt teaching methodologies and streamline the curriculum in a way that emphasizes the direct application of education to "real life" necessities and skills that do not fall under gender stereotypical lines. It is clear that the traditional notions of teaching, even when they are "student-centered," may not be sufficient to raise students' interest in learning and thus their performance. In a technology-driven world it means better utilization of ICT in the classroom to enhance students' familiarity with that tool, and application of more innovative strategies to capture the attention and interest of students, particularly males. Close relationships with businesses can enhance schools' understanding of the private sector expectations and also improve students' awareness about the importance of education and motivation, and their likely direct relationships with future earnings.

5. ***Increase students´ motivation to learn.*** This study has shown that there are significant differences between female and male learning environments, which might negatively affect males´ motivation to attend school and learn. Some measures to change that situation might include: a) improve the physical learning environment of male schools by ensuring schools are clean, organized, and have the necessary learning materials and resources; b) provide extra-curricular activities for male students (such as sports, army sponsored programs, "values" and "morals" classes); c) create remedial programs to support failing students; d) reward high-performing students.

6. ***Initiate interventions in directorates with large gender performance gaps***. Although the gender achievement gap is a phenomenon observed across Jordan, some locations identified in this report are in need of immediate actions to ameliorate the problem. Education programs currently implemented in those areas might expand their scope of work to address the gender inequity problems in student achievement. For example, school principals, supervisors, and MoE staff from the central Gender Department could assist male schools in developing specific action plans that contain initiatives and strategies designed to decrease the gender achievement gap.

7. ***Create initiatives to involve parents and communities in children's education.*** Many parents do not visit their children's schools and are not aware of their children's school performance in national assessments. It is recommended that the MoE invest in communication campaigns and strategies to invite parents and community members to visit the school. In addition to sponsoring *open school days* for community activities, it will be necessary to develop mechanisms to share school-level assessment results with both parents and the communities at large in order to increase school accountability.

1. INTRODUCTION

1.1 Background

The second phase of the Government of Jordan's Education Reform for the Knowledge Economy Project, also known as ERfKE II, was launched in 2010. The overall objective of ERfKE II is to provide students enrolled in pre-tertiary education schools (basic and secondary levels) with increased skills that will allow them to participate in the knowledge economy. To achieve this ambitious goal, the program focuses on five integrated and comprehensive components: 1) Establishment of a National School-based Development System; 2) Monitoring and Evaluation and Organization Development; 3) Development of Teaching and Learning; 4) Development of Special Programs; 5) Improvement of Physical Learning Environment.

This, "Gender Gap in Student Achievement in Jordan Report," will contribute to the overall mission goal of improving student learning for the knowledge economy by identifying school level factors, such as school, teacher, and student characteristics that might affect male and female student achievement differently. Moreover, it provides information that will directly inform national and school-based reform plans designed to ensure quality education for all students in Jordan, irrespective of gender.

This study is aligned with the Ministry of Education´s (MoE) efforts to institutionalize all issues related to gender, such as the creation of a Gender Department within the Directorate of Planning and Educational Research, in order to develop a national strategy for gender mainstreaming in the MoE. This initiative has been developed to achieve the following objectives:

- "Ensure that the roles of males and females complement each other in a way that achieves the desired positive change;
- Enable the MoE to ensure a gender perspective is taken into account in the development and implementation of the policies, strategies, work plans, programs, and activities of the Ministry;
- Unleash the human potential and equal rights to both males and females;
- Enable individual staff at MoE, of both genders, to actively participate in making decisions that take into consideration the different status and roles of men/boys and women/girls; and
- Follow-up on Education for All and the Millennium Development Goals".[5]

The Gender Department was established by the MoE in late 2010 to implement the National Strategy for Gender Inclusion, which aims to achieve gender equity at the central government, directorates and schools by working on related polices, plans, projects and programs. During the past three years, the

[5] Ministry of Education (November 2010). National Gender Mainstreaming Strategy for the Ministry of Education. Page 9.

Gender Department has launched and/or implemented several initiatives to ensure gender equity. In addition to participating in and producing studies to highlight gender inequities in Jordan, the Department has attempted to bring gender issues to the forefront of the educational agenda by undertaking the following tasks:

1. Participating in workshops at both national and international levels to disseminate information about gender issues in Jordan.
2. Participating actively in the development of national and international reports for the Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW), in cooperation with the Committee for Women's Affairs, and the United Nations Committee for Human Rights.
3. Assigning a gender focal point in every Directorate to ensure the gender mainstreaming strategy is implemented at the local level.
4. Developing training materials to bring gender awareness to the school, Directorate, and MoE levels.
5. Training Directorate level staff on how to incorporate the gender perspective in their Annual School Plans.

Trainings of Directorate and school-level staff comprise the core of the Gender Department activities at the present time. The trainings focus on the following tasks: (1) Identifying and reflecting on males' and females' work-related tasks at all levels of the Ministry, including directorates and schools; (2) Collecting school- and directorate-level data disaggregated by gender, and creating gender indicators; (3) Identifying evidence-based gender gaps and questioning why and where they occur; (4) Developing action plans, activities, resource allocation plans, and implementation strategies with the intention of decreasing gender gaps and enabling both males and females to benefit from MoE's programs and resources; and (5) Monitoring and evaluating progress on activities related to gender on a regular basis.

Currently, the Gender Department is tasked with understanding and participating in the proposal and planning of remedial interventions to improve education and curb the educational achievement gap between males and females. Jordanian females have been consistently outperforming males in all subjects, including in: Trends in Mathematics and Science Study (TIMSS), Programme for International Study Assessment (PISA), Tawjihii, National Test, and National Assessment for Knowledge Economy (NAfKE). Additionally, there is evidence that at the tertiary level of education, females' GPAs are significantly higher than those of males in most areas of study, including Mathematics, Engineering, and Computer Information Systems.[6]

While women´s high performance on standardized tests and at universities are very positive attributes of educational development in Jordan (and other countries, as will be discussed), the low performance of males and low participation of women in the economic and political arenas are both reasons for

[6] Khwaileh, F.M. & Zaza, H.I. (2011) Gender differences in academic performance among undergraduates at the University of Jordan: Are they real or stereotyping? *College Student Journal*, 45(3), 633-648.

significant concern. First, the stagnation of male educational attainment might be a precursor to low employment and weak earning opportunities in the future, particularly for those with low socio-economic status (SES).[7] Second, in addition to the human capital deficiencies associated with low educational attainment, a multi-country household survey has found that male educational attainment is highly associated with attitudes toward gender roles.[8] According to the study's findings, men with low educational attainment were more likely to have more rigid gender attitudes, more likely to use domestic violence against their partners, and less likely to participate in the daily care of children. They were also less likely to support policies related to gender equality.[9] Thus, to ensure an equal learning environment where they can succeed academically, it is important to understand the reasons for the educational gap between males and females, the gendered nature of schools, and disparities that affect both males and females. A stronger understanding of these issues also informs the MoE's Strategic Plan goals and ERfKE's II education strategy objectives of developing policies that address these issues and increase equity between men and women.

1.2 Regional and Jordanian Trends in Overall Gender Achievement Gap

The 2012 World Development Report[10] reveals that most countries have achieved gender parity in primary education enrollments. Further, in more than one-third of all countries that have been studied, females significantly outnumber males in secondary education, and tertiary enrollment growth has been increasing faster for females than males worldwide. Nevertheless, many inequalities persist. The Global Gender Gap Report (2012) developed a Gender Equity Index which examines inequality between men and women in 135 countries with respect to four critical areas (sub-indices): economic participation and opportunity, educational attainment, health and survival, and political empowerment.[11] The report concludes that, although countries have closed almost 96% of the gender gap in health outcomes and almost 93% of the gender gap in educational attainment, gaps in economic participation and political empowerment outcomes remain large (40% and 80%, respectively). For example, "the share of female directors in institutions ranges from 40% percent in Norway, where the government imposed a quota, to less than 2% percent in Bahrain, Japan, Jordan, the Republic of Korea, Qatar, Saudi Arabia, and the United Arab Emirates."[12]

---

[7] Autor, D. H. & Wasserman, M. (2013). Wayward Sons: The Emerging Gender Gap in Labor Markets and Education. Cambridge: Third Way. http://economics.mit.edu/files/8754

[8] Barker, G., Verma, R., Crownover, J., Segundo, M., Fonseca, V., Contreras, J.M., Heilman, B., & Awlak, P. (2012). Boys and education in the global south: Emerging vulnerabilities and new opportunities for promoting changes in gender norms. Journal of Boyhood Studies, 6(1-2), 137-150.

[9] Ibid.

[10] World Bank. (2011). World Development Report 2012: Gender equality and development. Washington, DC: The World Bank.

[11] Hausmann, R.; Tyson, L.D.; Zahidi, S. (2012). The Global Gender Gap Report, 2012. World Economic Forum.

[12] World Bank. (2011). World Development Report 2012: Gender equality and development. Washington, DC: The World Bank. p. 204.

Around the globe, Arab and North African countries rank the lowest on this Indexe's four sub-indices (above) and listed in Table 1, although progress has been made in the last few years, particularly in the field of education. For example, the United Arab Emirates is the only country in the region that has completely closed the educational attainment gap. Kuwait is one of the top 20 highest scoring countries with regards to secondary education enrolment and one of the top 10 in tertiary education enrollment.[13] The Hashemite Kingdom of Jordan is close to achieving gender parity in education and health (Index Scores = 0.9859 and 0.9706, respectively, out of 1.0) as well as the goals of Education for All. However, the political participation (0.0552) and economic index (0.4296) scores for women need improvement. In Jordan, only 17% of 20- to 45-year-old women work, compared with 77% of men in the same age group.[14] This gap exists even among the more educated, despite the fact that 93% of recent female graduates claim that they want to work, and 91% claim they would like to work outside the house after they are married.[15]

Table 1: Overall Gender Inequality Rankings in the Middle East and North Africa

Country	Overall Gender Equity Index (0-1*)	Ranking (Out of 135 countries)
Israel	0.6989	56
United Arab Emirates	0.6392	107
Kuwait	0.6320	109
Bahrain	0.6298	111
Qatar	0.6264	115
Mauritania	0.6129	119
Algeria	0.6112	120
Jordan	**0.6103**	**121**
Lebanon	0.6030	122
Oman	0.5986	125
Egypt	0.5975	126
Morocco	0.5833	129
Saudi Arabia	0.5731	131
Syria	0.5626	132
Yemen	0.5054	135

Source: The Global Gender Gap Report, 2012. World Economic Forum, p. 18
*The highest and lowest possible scores are 1 (equality) and 0 (inequality), respectively.

1.3 Overall Assessment Trends and Reasons for Gender Gap in Achievement

Previous studies around the world, such as the ones presented below, have indicated that understanding the reasons for the gender gap in assessment scores is a complex exercise that involves

[13] http://forumblog.org/2013/05/top-5-countries-for-gender-equality-in-the-middle-east/
[14] World Bank. (2011). World development report 2012: Gender equality and development. Washington, DC: The World Bank. p. 301.
[15] World Bank. (2011). World development report 2012: Gender equality and development. Washington, DC: The World Bank.

understanding the socio-economic, cultural, educational, and psychological environments in which students live and go to school.

In another in-depth evaluation, an analysis of changes in gender differences among 8th graders from 16 countries participating in the Trends in International Mathematics and Science Study (TIMSS) between 1995 and 2003 suggested that there were no major changes in gender gap scores for mathematics during that period[16]. Further, the most recent Programme for International Student Assessment (PISA) results (2012) seem to indicate that females outperformed males in only five out of 44 PISA countries and economies[17]. The PISA 2012 report also suggests that some reasons for the gender gap in favor of males, in most countries, are the gender differences around mathematics and self-beliefs in mathematics. Females tend to report "less perseverance, less openness to problem solving, less intrinsic and instrumental motivation to learn mathematics, less self-belief in their ability to learn mathematics and more anxiety about mathematics than males, on average; they are also more likely than males to attribute failure in mathematics to themselves rather than to external factors."[18]

With regards to science, the overall analysis of student performance in TIMSS between 1999 and 2003 suggested that the gender gap seemed to be closing, especially in chemistry and physics.[19] More recent assessment results have also revealed that the gender gap in science in PISA 2006 and 2012 has remained stable during the period under study.[20] However, in countries where the magnitude of the gender gap increased, the change always favored females. Although many factors may be at play in explaining female students' overall progress in science, the case of Australia illustrates how policy can affect the gender gap. During the 1970s and 1980, there were multiple campaigns to encourage females' participation in mathematics and the sciences, and this resulted in a significant narrowing of the gap that favored males in those subjects in either TIMSS 1995 or TIMSS 1999. Yet as the Australian campaign came was phased out over time and the country focused on other issues, it became clear that by secondary school males were outscoring females in all science domains, with the exception of life science in TIMSS 2003.[21] In addition, females reported lower levels of self-confidence in science and did not value science as much as males. In that particular study, teacher and school factors, including class size, location, and availability of resources, did not seem to affect achievement.[22] The case of Australia highlights how detailed and complex the explanation for gender differences can be. More importantly, it emphasizes that policy may influence how females perceive science and perform in assessments.

---

[16] Neuschmidt, O., Barth, J., & Hastedt, D. (2008). Trends in gender differences in mathematics and science (TIMSS 1995–2003). Studies in Educational Evaluation, 34, 56–72.
[17] OECD (2013), PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I), PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264201118-en.
[18] Ibid, p. 18.
[19] Neuschmidt, O., Barth, J., & Hastedt, D. (2008). Trends in gender differences in mathematics and science (TIMSS 1995–2003). Studies in Educational Evaluation, 34, 56–72.
[20] OECD (2013), PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I), PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264201118-en
[21] Thomson, S. (2008). Examining the evidence from TIMSS: Gender differences in Year 8 science achievement in Australia. Studies in Educational Evaluation, 34 (2008) 73–81.
[22] Ibid.

Finally, an analysis of gender differences in reading scores points to more straightforward and consistent results across countries over time. Results from the Australian Studies, the National Assessment of Educational Progress Studies (NAEP) and PISA, conducted in 2006, revealed that female secondary students performed above their male peers in reading skills, regardless of age or language of instruction.[23] More recently, PISA 2009 revealed that females outperformed males in reading in every PISA country. In OECD countries, the average gender gap was 39 score points, or over half a proficiency level.[24] PISA 2012 confirms that females continue to outperform males in every participating country by 38 score points, on average. In Jordan, females outperform males by 75 score points, an entire proficiency level.[25]

Several studies have attempted to explain gender differences in reading abilities. Lynn and Mikk[26] have found that higher achievement among females might be associated with school and home conditions. In their analysis of PISA (2003) data, they have learned that females read more at home and school while males spent more time using computers and DVD players. Another study conducted in 2009[27] compared males' and females' behaviors and their relationship to reading performance. The findings revealed that males were less likely to read at home and more likely to play sports than females. Males were also less likely than females to report positive attitudes toward reading and writing, less likely to complete homework, and more likely to use electronic entertainment. Teachers and principals interviewed for that study shared strong beliefs about males' and females' maturational and educational differences as well. Males were described as being "more tactile and kinesthetic learners, to be slower to develop motor skills, to have difficulty transferring ideas to paper and to have a higher need for structure" (p.24). Females were perceived to have "more organized thought processes, [to be] better listeners, [to be] able to sit and sustain attention for a longer period of time and capitalize on earlier language development" (p. 24).

Given these varying performance results across multiple subjects, it is clear that several explanations might exist for the gender gap or equity in males' and females' performance in student assessments across countries. In fact, the causes for this gender gap or lack thereof can be complex, and gender is only one of many factors that account for achievement. Students´ socio-economic status and school characteristics are other very strong factors that can influence achievement and these might be interconnected with socio-cultural factors such as one´s choice of career path. This in turn may

[23] Lietz, P. (2006). A meta-analysis of gender differences in reading achievement at the secondary school level. Studies in Educational Evaluation, 32, pp. 317–344.
[24] OECD (2010), PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I). http://dx.doi.org/10.1787/9789264091450-en
[25] OECD (2013), PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I), PISA, OECD Publishing. http://dx.doi.org/10.1787/9789264201118-en
[26] Lynn, R. & Mikk, J. (2009). Sex differences in reading achievement. Trames, A Journal of the Humanities and Social Sciences, 13(63/58), 1, pp. 3–13.
[27] Klinger, D.A., Shulla, L.A., Wade-Woolley L. (2009). Towards an understanding of gender differences in literacy achievement. Toronto: Education Quality and Accountability Office (EQAO).

influence the aspirations and the academic performance of males and females differently.[28] Our study will attempt to investigate the quantitative and qualitative factors that might explain the gender gap in Jordan and will propose policy-related measures that can decrease the gender gap in favor of either sex.

2. STUDY OBJECTIVES AND EVALUATION QUESTIONS

2.1 Study Objectives

This study was commissioned by the MoE in Jordan in order to ascertain appropriate measures for improving male and female academic performance in national and international standardized tests. The case of the gender gap in Jordan is somewhat unique because most males and females attend single-sex government schools after Grade 3. There is anecdotal information that male and female schools differ greatly in terms of quality. There is also a common belief that social and cultural gender factors might be responsible for the difference in academic performance between the sexes. Through this research, the MoE hoped to obtain empirical evidence of the gender gap and its related causes in order to be able to address student and school-related factors that might contribute to overall education inequality in Jordan. Therefore, this study focuses on three main objectives:

1) Determine the *magnitude* and *trend* in the achievement gap between males and females in TIMSS, PISA, NAfKE, and National Test (NT) over time[29];

2) Investigate the factors associated with the gender gap in different assessments over time so as to understand how those factors operate;

3) Based on the findings, propose policy options that will help narrow the gender gap in student achievement.

2.2 Evaluation Questions

The study's main questions are:

1. What are the *magnitude* and the *trend* of the gender gap in achievement between males and females on national and international assessments (TIMSS, PISA, NAfKE, NT) over time?
2. What are the main factors associated with the gender gap in achievement in TIMSS, PISA, NAfKE, and NT assessments? Can the gender gap be explained by school, location, grade,

[28] Education, Audiovisual and Cultural Executive Agency (2010). Gender differences in educational outcome: Study on the measures taken and the current situation in Europe. DOI: 10.2797/3598.

[29] Two previous reports have addressed the gender achievement gap in early grades (EGRA/EGMA) and among older students (Tawjihii exam). Therefore, analysis of those tests results were not be carried out and discussed in this report. For more details, see 1) Student Performance in Reading and Mathematics, Pedagogic Practice, and School Management in Jordan (2012) by RTI International; and 2) Gender Gap in Tawjihii (unpublished) by Gender Department, Ministry of Education, Jordan.

subject, teacher characteristics, or other factors related to students' practices and attitudes in/outside schools?

3. How do the main factors associated with the gender gap contribute to the variance in achievement?

3. METHODOLOGY

Both quantitative and qualitative research methods were employed to ascertain the gender gap in student assessment and its related factors.

3.1 Strategy for Quantitative Analyses

The study utilized available TIMSS, PISA, NAfKE, and NT datasets to conduct quantitative analyses, as explained below.

1) To answer **Question 1**, the research team conducted a comparative analysis of male and female scores by analyzing TIMSS (2003, 2007, 2011), NafKE (2006, 2008 and 2011), PISA (2006, 2009, 2012), and NT (2007, 2011, 2013) datasets. Overall scores in Math, Science, Arabic, and English were then compared by gender and grade over time. The research team conducted appropriate tests to assess the significance level of male and female differences and the effect sizes of the gender gaps across subjects.

2) To answer **Question 2**, the research team analyzed student, school, and teacher background and school variables from NafKE (2011), TIMSS (2011), and PISA (2012) data sets (as presented in Table 2). Variables derived from those assessments about student, teacher, and school information were included in the final models after three considerations: 1) conceptual importance; 2) correlation between those variables and students' scores in the specific assessments and subjects; and 3) statistically significant differences between males and females with respect to each specific variable. The choice to analyze specific years and subjects was guided by the following rationale:

a) The model for TIMSS scores focused on 8th grade Science, because the effect size in that subject was the largest among MoE schools.

b) The model for PISA focused on Math for 15 year-olds (10th grade) for various reasons. The effect size of the gender difference in PISA was similar to the effect size in other assessments in MoE schools. Further, PISA 2012 focused on Math, and most background variables focused on that subject. For example, when students were asked about reading habits, the question was: How much do you enjoy reading about math?

c) The model for NAfKE focused on Arabic in 5th grade, as the effect size of the difference in that subject in MoE schools was relatively large. In addition, we selected the 5th grade because it is

the earliest grade that is assessed, and we would like to determine if there were specific factors that affected gender differences in an earlier grade.

d) Regarding the NT, it was not possible to build a model, as the dataset does not contain detailed student, school, and/or family background data. In the future, the MoE might consider adding specific background information to the questionnaires as to enable more advanced analysis.

The research team then utilized Hierarchical Linear Model (HLM) to ascertain the factors associated with student achievement in each assessment as well as the gender achievement gap. HLM was considered to be the most appropriate statistical technique for this hierarchical structure -- i.e., students nested within schools. It allowed evaluators to use a two-level model to assess the factors influencing gender differences in the assessment by identifying both student and school perspectives.

Table 2: Summary of Stakeholders´ Information Data Available through the Datasets

Datasets (Years)	Participants[30]	Stakeholders' Data				
		Student	Teacher	Principal and Supervisor	Parents (mothers and fathers)**	School Characteristics
PISA Dataset (2006, 2009, 2012)	Representative Sample (2012) (Student n=7,038 Principal n=233)	✓	-	✓	✓	✓
NAfKE Dataset (2006, 2008, 2011)	Representative Sample (2011, 5th grade) (Students n=1,525) Teachers n=169 Principals n=169)	✓	✓	✓	✓	✓
TIMSS Dataset (2003, 2007, 2011)	Representative Sample (2011) (Students n=7,694) (Teachers n=253) (Principals n=230)	✓	✓	✓	✓	✓
NT (2011, 2012, 2013)	Census (2013) (Students N= Arabic : 24,975 English: 23,385 Math: 24,342 Science: 25,151)	✓ Basic EMIS info.	Not Available			

**Data about parents was collected through Students´ Questionnaires.

[30] The total numbers represent all schools in Jordan. The final models include only government schools.

3.2 Strategy for Qualitative Analysis

- To answer **Question 3**, we carried out focus group discussions with teachers, parents, principals, supervisors, and students to understand what they perceived as the reasons for the gender achievement gap and how the identified factors contributed to differences in achievement between the sexes.

Instruments and Procedures

Focus Group Protocols (FGPs) and short questionnaires were developed to gather data to answer Question 3:

FGPs and Questionnaires

FGPs were comprised of semi-structured questions to be discussed by FG (focus group) participants. Although questions differed for each group of stakeholders, all protocols centered around common themes: a) attitudes of males and females regarding school; b) perceptions about education and career opportunities; c) parental attitudes/behaviors; (d) cultural perceptions/expectations; e) general activities outside school.

Participants in each FGD were also asked to fill out a short questionnaire about their opinions/attitudes regarding school, parental attitudes, and cultural perceptions/expectations. The number of items per questionnaire varied from 19 (Students) to 24 (Principals). Each item could be rated as (1) Yes; (2) Somewhat; (3) No. The results from the tabulation of those items were used to support the FGD notes. In addition to these items, questionnaires also included two open questions to ascertain: (a) participants' opinions about the three most important reasons why female students outperform male students academically; and (b) participants' opinions about the main reasons for the overall decline in student academic achievement (males and females) in the last few years.

Procedures

Three field directorates were selected for FGs: Aqaba (South), Irbid (North), and Amman (Center). These locations were selected based on those schools that had participated in TIMSS 2011, PISA 2012 and/or NAfKE 2011. To be selected for FGs, the school should have participated in at least two of the assessments mentioned. If a school had participated in all three studies, the school was given priority in the selection process. It is important to note that all schools had participated in the NT, as that assessment is census based. A summary of the schools and stakeholders selected for the qualitative component of this evaluation is presented in table 3. Additional information can be found in Appendix A.

Table 3: Characteristics of Schools Selected for Focus Group Discussions

School ID	# FGD Participants				School Sex	Authority	Region
	Teachers	Students	Parents	Principals/ Directorates Supervisors			
113230	10	12	10	14 Supervisors and 4 Principals	Male	MoE	South
112470	10	10	8		Male	MoE	South
112487	10	11	9		Female	MoE	South
171707	11	10	14		Co-ed	Private	South
111037	10	10	6	14 regional Supervisors and 4 Principals	Male	MoE	North
111049	10	10	9		Female	MoE	North
150002	10	10	9		Co-ed	Private	North
160143	10	10	9		Female	UNRWA	North
110149	10	9	11	15 regional Supervisors and 4 Principals	Male	MoE	Middle
110243	10	10	10		Female	MoE	Middle
170217	9	10	11		Co-ed	Private	Middle
160027	9	10	9		Male	UNRWA	Middle
Totals	**119**	**122**	**115**	**43**			

For this part of the study, NCHRD hired a total of 12 "supervisors" to conduct FGs (6 facilitators and 6 note-takers) with school teachers, principals, supervisors, students, and parents. One facilitator and one note-taker jointly carried out each focus group. Facilitators and note-takers were MoE Supervisors, who were trained by NCHRD for an entire day for their respective roles as either facilitators leading the focus groups or as note-takers summarizing what participants were sharing with the group. All note-takers were English language supervisors and they translated the notes into English after the focus groups ended.

In each Directorate, two pairs of supervisors (2 facilitators and 2 note-takers) carried out focus group discussions in four schools. All interactions were voice-recorded. Although the voice recordings were not transcribed, the NCHRD research team verified the notes with random sections of focus group tapes. Each focus group had between 9 and 12 participants, with similar distribution of males and females in all groups. The exception was the supervisors, which were predominantly male, and the parents, who were for the most part separated by gender.[31]

In addition, NCHRD selected a different group of supervisors to organize the data by themes and to process them for analyses. The NCHRD research team supervised the data entry process and later conducted the qualitative data analysis in collaboration with the MEP and The Evaluator's Institute (George Washington University) team accordingly.

[31] Mothers usually participate in school-related activities, including meetings in female schools. Fathers participate in male school-related activities.

3.2 Limitations

It is important to keep in mind that schooling and learning are complex phenomena. The analysis presented in this study is limited to the set of variables available in TIMSS, PISA, and NAfKE background questionnaires. Therefore, many student and school characteristics that might influence achievement may not be accounted for in the final models presented below. Although the qualitative data presented may support quantitative findings and shed light on additional reasons for the gender achievement gap, that information cannot be imputed into quantitative models, and it cannot be generalized to the overall student, teacher, and school population in Jordan.

An additional limitation is the lack of background information available to conduct in-depth analysis of NT results. Consequently, when building HLM models, the research team could not explore the relationship between within and between school variables, gender and students' achievement in NT.

4. FINDINGS

4.1 Characteristics of the Gender Gap in Student Achievement

To address the first study objective, the research team compared male and female performance on several standardized tests conducted in Jordan over time, namely TIMSS, PISA, National Assessment for Knowledge Economy (NAfKE) and National Test (NT). TIMSS and PISA are international assessments, while NAfKE and NT are national assessments. More information about each test will be provided below under the "Appropriate Results section." The objective of comparing students' scores on those assessments was to determine whether or not females outperformed males in all assessments, grades, and subjects. If differences actually existed, the research team established the magnitude of the difference between the two sexes and determined whether there was a clear trend (i.e., an increase or decrease of the gender achievement gap across time).

In addition, in this section the research team compared male and female scores at the directorate level and in different types of schools (e.g., MoE, UNRWA, private, other government schools, urban and rural). For the policymaker, obtaining information on specific characteristics of the gender achievement gap may assist in making important decisions, such as focusing on specific administrative and academic practices carried out in different kinds of schools and/or determine which directorates should be targeted first for any specific intervention that might decrease the gender achievement gap.

Some general findings, presented in Figures 1-8 below, suggest that males and females are both performing below the expected average in every subject or grade across all participating nations. However, females outperformed males in all assessments, grades, and subjects.

TIMSS 2003, 2007, and 2011 Results

Figure 1 demonstrates the results for TIMSS 2003, 2007, and 2011. TIMSS is an international sample-based assessment organized by the International Association for the Evaluation of Educational Achievement (IEA); it assesses students' achievements in science and mathematics. The assessment is aligned with the curriculum of each participating country. Generally, TIMSS assessments at the eighth

grade level are organized by content and cognitive dimensions (knowing, applying, and reasoning). TIMSS scores can range from 0-1000. As Figure 1 suggests, both males and females scored at or below the TIMSS center point of 500. The performance average in science for Jordanian students was 449, significantly lower than the international average of 477. Among the Arab countries, Jordan ranked third (out of 11), and the Jordanian performance average in science was higher than the Arab average of 428. Similarly, the average performance score for Math in Jordan was 406, which is below the international average of 467. In Math, Jordan ranked 35th (out of 45) at the international level and sixth out of the eleven Arab nations that were measured. [32]

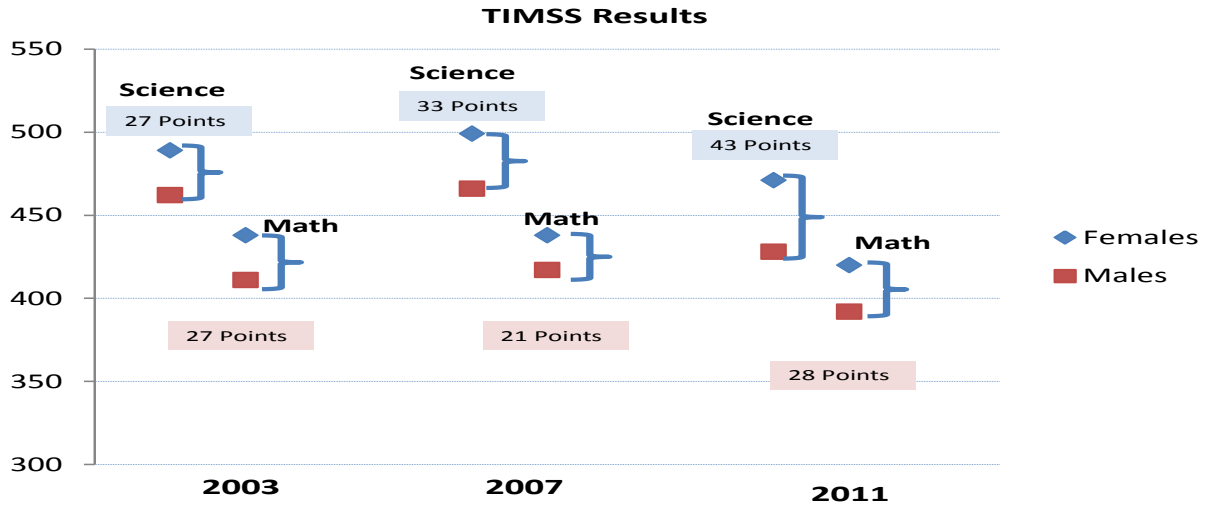
The differences between male and female TIMSS scores were significant in all cycles (2003, 2007, and 2011) and subjects (Math and Science). In 2003, the gaps between males and females in Math and Science were the same (27 points)—equivalent to 2.7 percentage points in a 100-point scale[33]. In 2007 and 2011, however, the gender achievement gap in Science had grown to be larger than that in Math. In 2011, for example, the difference between female and male average scores in Science (471 and 428, respectively) was 43 points (the equivalent of 4.3 percentage points). The difference in Math was smaller (28 points)—equivalent to 2.8 percentage points, with females scoring 420 points and males scoring 392.

Although the evidence above points to changes in the gap between male and female scores in both subjects over time, those changes (increases or decreases) from 2003—2011 were not statistically significant. Thus, these results suggest that the gaps for Math and Science in TIMSS have remained somewhat stable, showing no consistent signs of significant increases or decreases over the three assessment cycles.

[32] NCHRD (2013). Jordanian National Report on the Trends in Math and Science Study (TIMSS) 2011. Amman: NCHRD.

[33] In order to compare TIMSS and PISA scores with NAfKE and NT, we may convert the TIMSS scale from 1000 to 100 points. For example, the differences between male and female scores in Science (43 points) and Math (28 points) are equivalent to 4.3 and 2.8 percentage points on a 100-point scale.

Figure 1: TIMSS Results and the Achievement Gap in Math and Science for Males and Females in 2003, 2007, and 2011



PISA 2006, 2009, and 2012 Results

Figure 2 demonstrates the results for PISA 2006, 2009, and 2012. PISA is an international standardized assessment sponsored by the Organization for Economic Cooperation and Development (OECD). It assesses the abilities of 15-year-olds in reading literacy, mathematics literacy, and science literacy. More specifically, it assesses students' acquired knowledge and the skills that are essential for full participation in society. Each cycle government includes assessments of all three subjects, but it assesses one of the subjects in depth. In 2012, for example, the assessment focused on math literacy. PISA scores can range from 0-1000. Similar to TIMSS, the center point is 500, and for NAfKE and NT comparability purposes, we may convert the PISA scale from 1000 to 100 points.

Figure 2 shows that in Jordan, both male and female scores are below PISA's center point (500) and below the 2012 OECD averages in Math (494), Arabic (496), and Science (501). Males scored lowest in Arabic (in 2006 and 2012) and Science (2009). Females scored lowest in Math during the three assessment years.

The largest gender achievement gaps occurred in Arabic. In fact, in 2012, the 75-point difference in Arabic (or 7.5 percentage points if we convert the scale to 100) was the largest in the world, with females scoring 436 points and males scoring 361. That difference corresponds to an entire proficiency level[34]. Jordan also led the gender achievement gap world rank in Science. In 2012, the difference between female and male average scores in Science (430 and 388, respectively) was 42 points (or 4.2

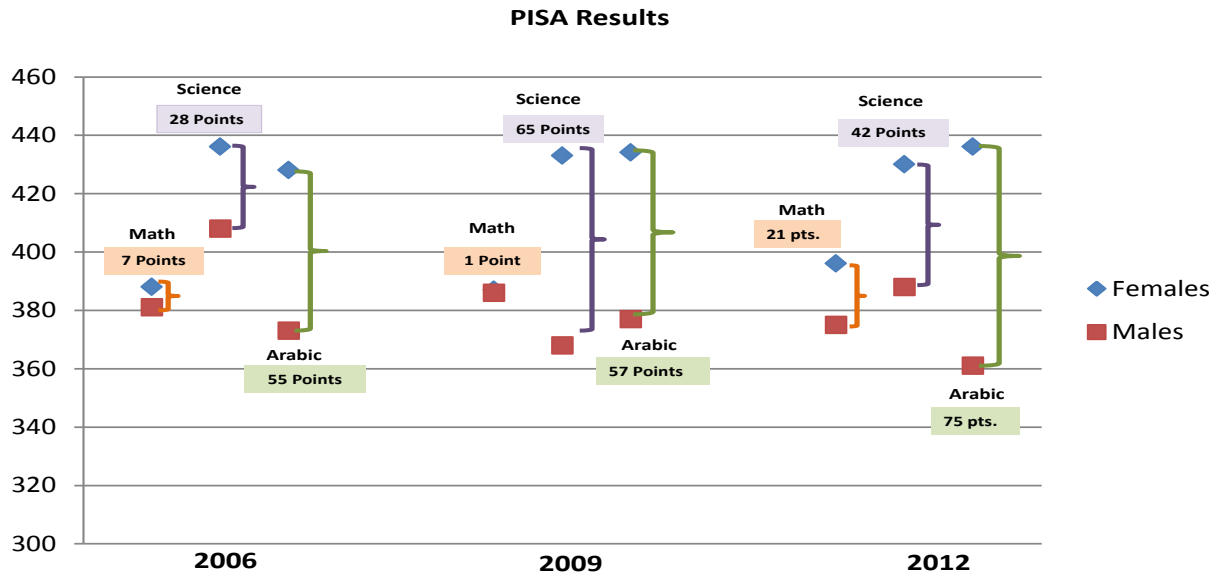
[34] According to the PISA 2012 report (p. 190), the seven proficiency levels used in the PISA 2012 reading assessment are described according the three processes that students use to answer the questions. These three processes are classified as access and retrieve (skills associated with finding, selecting and collecting information), integrate and interpret (processing what is read to make sense of a text), and reflect and evaluate (drawing on knowledge, ideas or values external to the text).

percentage points). The differences in scores between males and females in those subjects were statistically significant in all years.

For Math, the differences between male and female scores were not statistically significant in 2006 or 2009. Only in 2012 did the differences become significant, with females scoring higher. In fact, in that year, the difference between female and male average scores in Math (396 and 375, respectively), was 21 points (or 2.1 percentage points). That is a significantly higher difference than in 2006 (7 points/0.7 percentage points) and 2009 (1 point/0.1 percentage point).

An additional characteristic of PISA is the fluctuation of the gender achievement gap magnitude over time. As can be observed, the gender achievement gap in math scores decreased from 2006-2009 (from 7 to 1 point) and increased greatly from 2009-2012 (from 1 to 21 points). Only the latter change was statistically significant. For science, there was an increase in the gender achievement gap between 2006-2009 and then a decrease from 2009-2012. Neither of those changes was statistically significant. Only the gap in Arabic seemed to increase consistently over the three assessment cycles, but just the change between2009-2012 was statistically significant. The percentage change between those two cycles was 31.6% (from 57 to 75 points). These results suggest that there is no clear pattern in gender achievement gap changes over time.

Figure 2: PISA Results and the Achievement Gap in Math, Science, and Arabic for Males and Females in 2006, 2009 and 2012



NAfKE 2006, 2008, and 2011 Results

Figures 3 through 5 display NAfKE results for Grades 5, 9, and 11 in 2006, 2008, and 2011. NAfKE was a survey designed by NCHRD in collaboration with MoE at the beginning of the Education Reform for Knowledge Economy project (ERfKE I) in 2006. NAfKE was designed to identify students' skills and cognitive abilities in three subjects (Mathematics, Science, and Arabic language) across three grades.

NAfKE items present real life situations to assess students' knowledge economy skills. Scores can range from 0-100.

Overall, males scored below 50 (the center point) in all subjects each year, except in 11th grade Arabic (in 2008 and 2011). Females performed better than males overall, but the highest average score across years and subjects was 66 in Literacy (in 2008). As presented below, math scores were collectively the lowest compared to Arabic or Science across all grades, years, and students (males and females). In fact, the mean scores did not reach 50 (mid-point) in any grade or year in that subject, indicating that students on average answered less than 50% of the questions correctly.

As Figure 3 indicates, the smallest gender achievement gaps in Grade 5 across the years have been in Math, followed by Science and then Arabic. In 2011, the difference between female and male average scores in Science (45 and 42, respectively) was 3 points. The difference in Arabic was 8 points, with females scoring 42 and males scoring 34. For Math, the difference was 2 points (females scored 28 and males scored 26). Differences in scores between males and females in 5th grade were statistically significant for all years and subjects.

In addition, we observe that the magnitude of the gap for 5th graders fluctuated across years. The gap for all subjects seems to have increased from 2006-2008 and then decreased from 2008-2011. However, those increases in gap during those periods were not statistically significant for Math and Literacy. Only in Science did we observe that the gap decreased significantly between 2008 and 2011 (from 11 points in 2008 to 3 points in 2011).

Figure 3: NAfKE Results and the Achievement Gap in Math, Science, and Arabic for 5th graders in 2006, 2008 and 2011

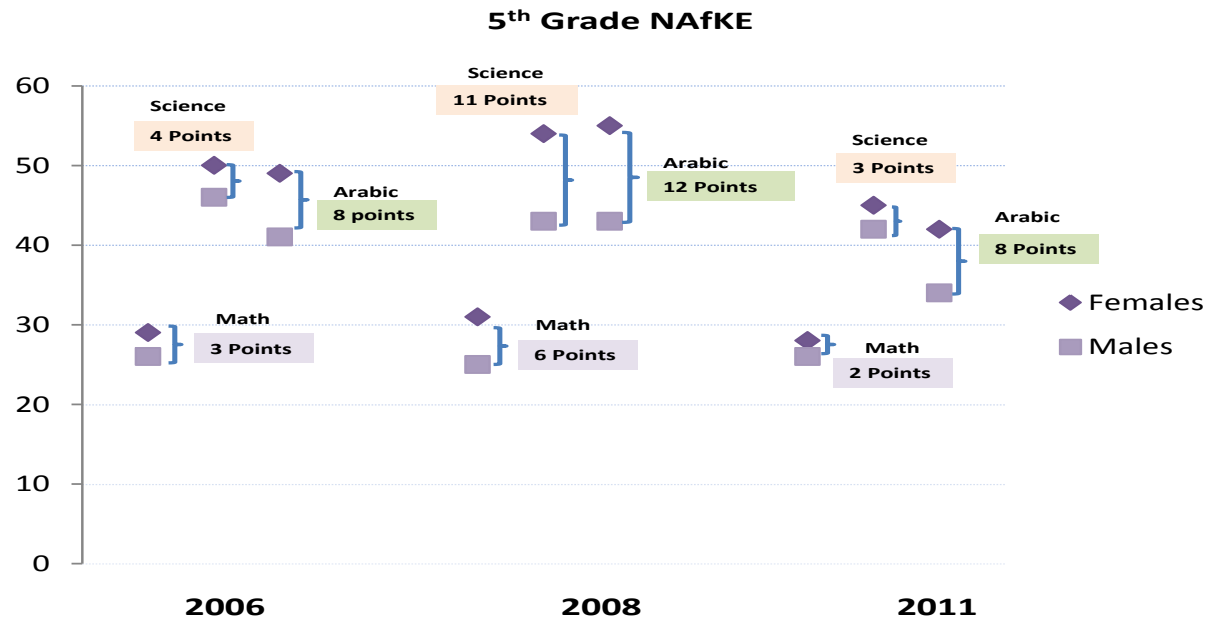


Figure 4 demonstrates that Math scores in Grade 9 continue to be the lowest compared with scores in Arabic or Science. However, Science scores are also very low (at or below the mid-point) for males and females in any given year. The largest gap in scores between males and females in Grade 9 across years can be observed in Arabic, followed by Science and then Math. In 2011, the gap in Arabic scores was 13 points, with females scoring 46 and males scoring 36 points. For Science, the gap between male and female scores was lower at 8 points, as females scored 40 and males scored 32. Finally, the smallest gap was 4 points in Math, with both females and males scoring very low (35 and 31 points, respectively). The differences in scores between males and females in all years and subjects were statistically significant. We can also observe from Figure 4 that for Grade 9, the magnitude of the gender achievement gap seems to have been increasing consistently over time for both Arabic and Science. In 2011, the gap in Arabic scores in 2011 was 18% larger than that in 2008. The gap in Science scores was 60% larger in 2011 than in 2008. However, only the change in Science was statistically significant. For Math, none of the changes in magnitude were statistically significant, indicating that the gender achievement gap is not increasing for that particular subject in Grade 9.

Figure 4: NAfKE Results and the Achievement Gap in Math, Science, and Arabic for 9th graders in 2006, 2008, and 2011

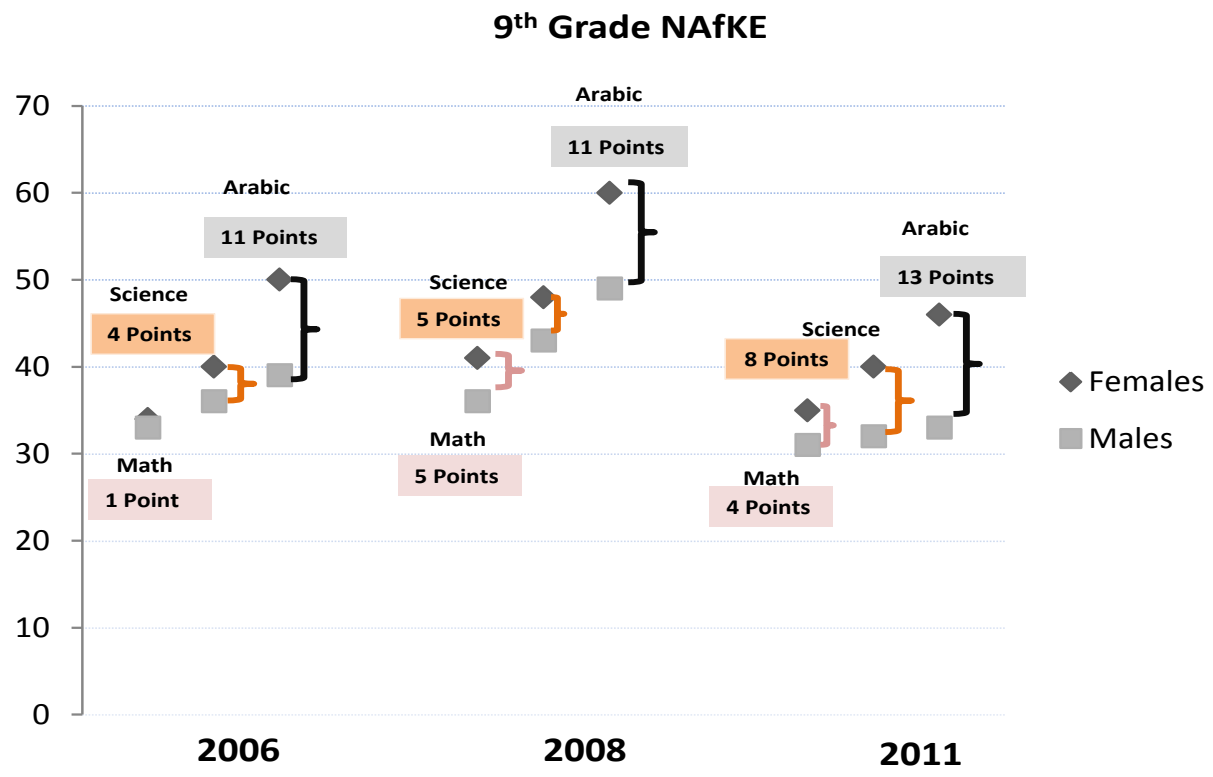
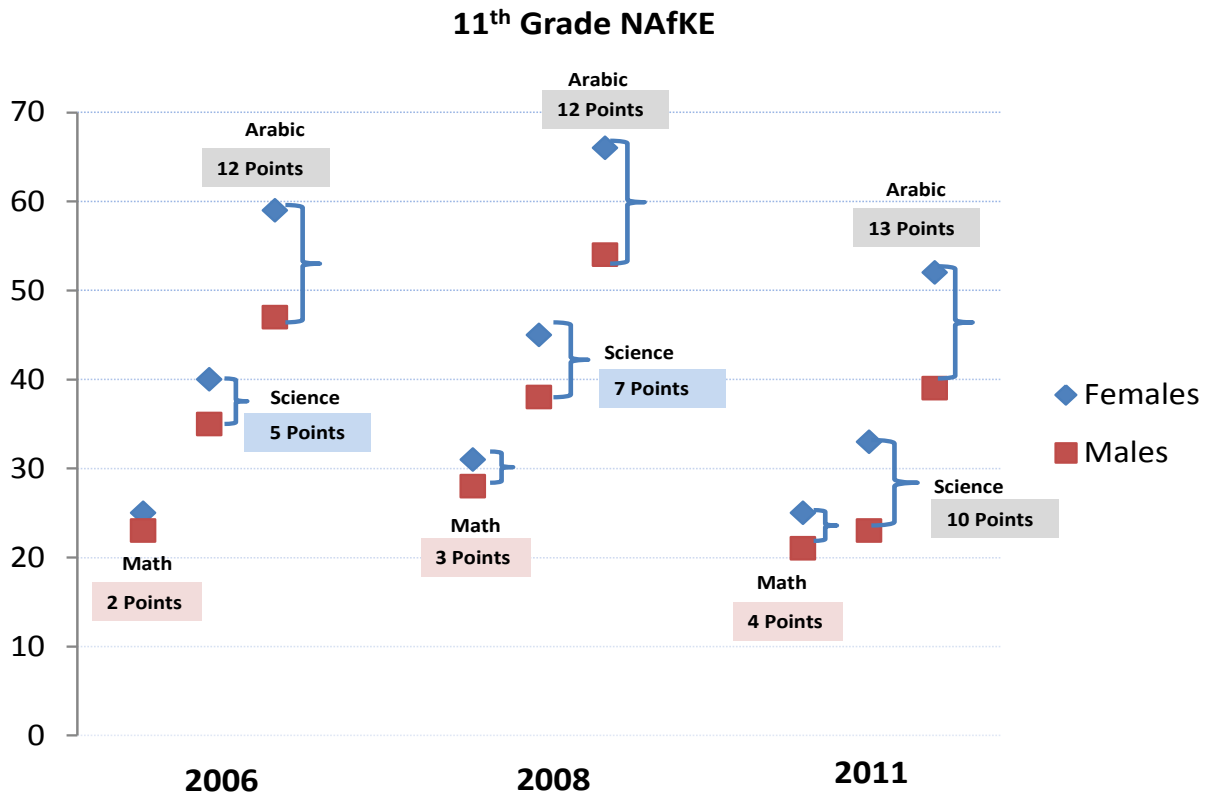


Figure 5 displays NAfKE scores for males and females in grade 11. Cumulatively, we observe that both male and female literacy scores were higher in Grade 11 than in Grades 5 and 9. However, overall performance in Science and Math in Grade 11 was much lower than in the earlier grades. In addition,

both males and females scored below the center point in those two subjects in all three cycles (2006, 2008, and 2011).

Similar to the earlier grades, the smallest gaps between males and females were in Math, followed by Science and Arabic. This is true for the three assessment cycles. The largest gaps between male and female scores can be observed in 2011 in Arabic (13 points), Science (10 points) and Math (4 points). The gender differences in all subjects and years are statistically significant. We can also observe that the magnitude of the gender achievement gap in all subjects has increased over time. However, only the changes in magnitude for Math were statistically significant.

Figure 5: NAfKE Results and the Achievement Gap in Math, Science, and Arabic for 11th Graders in 2006, 2008, and 2011



The NAfKE results presented in Figures 3 through 5 above (by grade) also allow us to observe another interesting pattern—the gender achievement gap in student assessment by grade. In 2011, the latest assessment cycle, the differences between female and male scores in Math for 5th, 9th, and 11th graders were 2, 4, and 4 points, respectively. For Science, the differences in the same grades were 3, 8, and 10 points, respectively. Finally, for Arabic, the differences were 8, 13, and 13 points. These data suggest that the differences in scores between males and females seem to increase as students move up in the education system; i.e., in earlier grades, the gender achievement gap is smaller than in later grades.

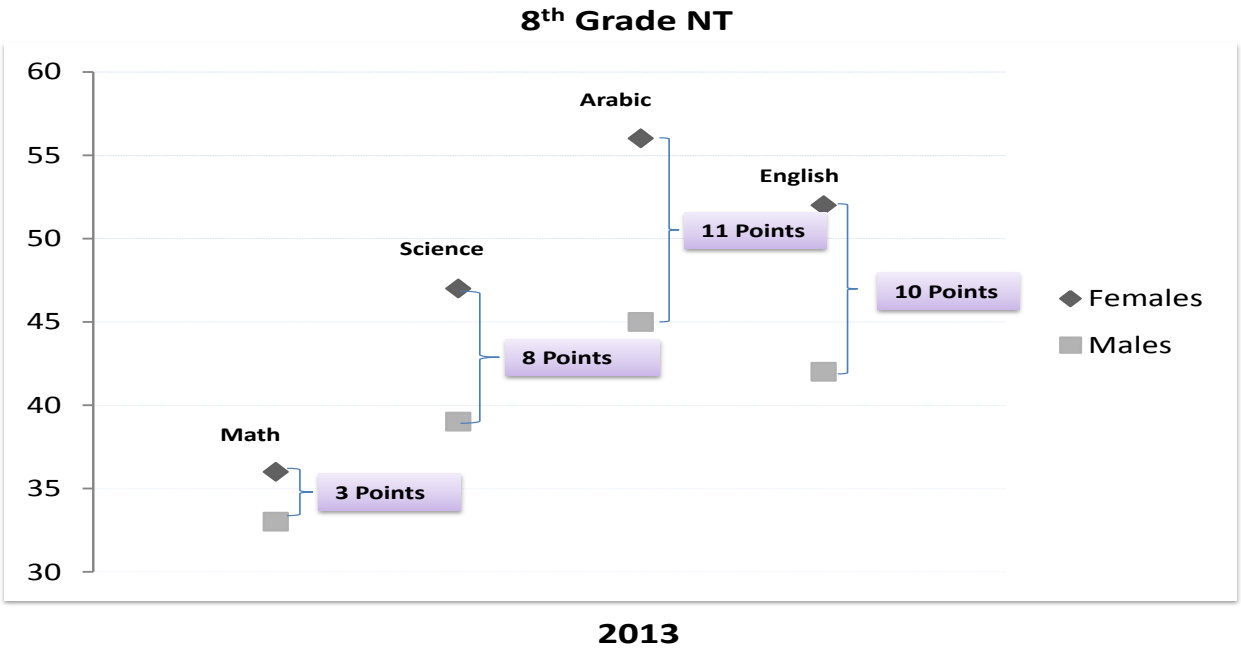
NT 2008, 2011, 2012, and 2013 Results

The National Test (NT) is a census-based test organized and administered by the MoE's Department of Examinations and Testing (DET), and this assessment is aligned with the national curriculum. National Test (NT) assesses all students in grades 4, 8, and 10 in all government schools on their performance in Arabic and English languages, mathematics, and science. However, NT only assesses one single grade in any given year and it takes three years to "repeat test" one grade. For example, NT assessed 10th graders in 2011, 4th graders in 2012, and 8th graders in 2013. Thus, 10th graders will not be assessed again until 2014.

Figure 6 shows 8th grade performance in NT in 2013. Similar to other studies, the lowest scores for males and females were in Math (below 40) and Science (below 50), scores in Arabic and English were higher. For females, average scores in Arabic and English were 56 and 52, respectively. For males, scores in those subjects were significantly lower, at 45 and 42, respectively.

Similar to other assessments, the smallest gap between males and females is in Math (3 points), followed by increasing gaps in Science (8 points), English (10 points) and Arabic (11 points), respectively.

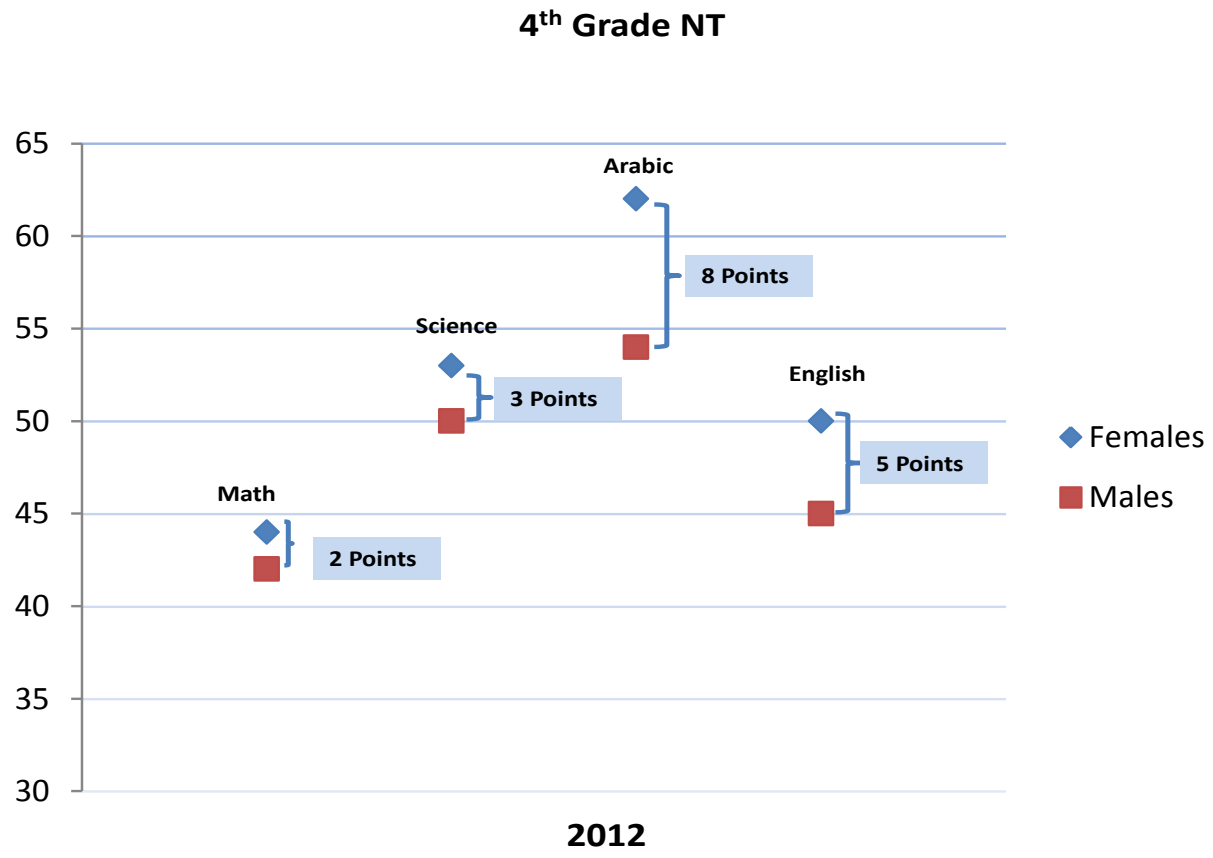
Figure 6: NT Results and the Achievement gap in Math, Science, Arabic, and English for 8th Graders in 2013



Regarding NT 4th grade scores (2012), displayed in Figure 7, we observe that females and males scored at or above the center point (50) in Arabic and Science. For Arabic, female and male scores were 62 and 54, respectively. For English, the scores were somewhat lower: 50 and 45 points, respectively.

Similar to other tests, the gender achievement gaps in Arabic and English were the largest (8 and 5 points, respectively), while the smallest gaps were in Math (2 points), followed by Science (3 points). All differences in the scores between males and females were statistically significant.

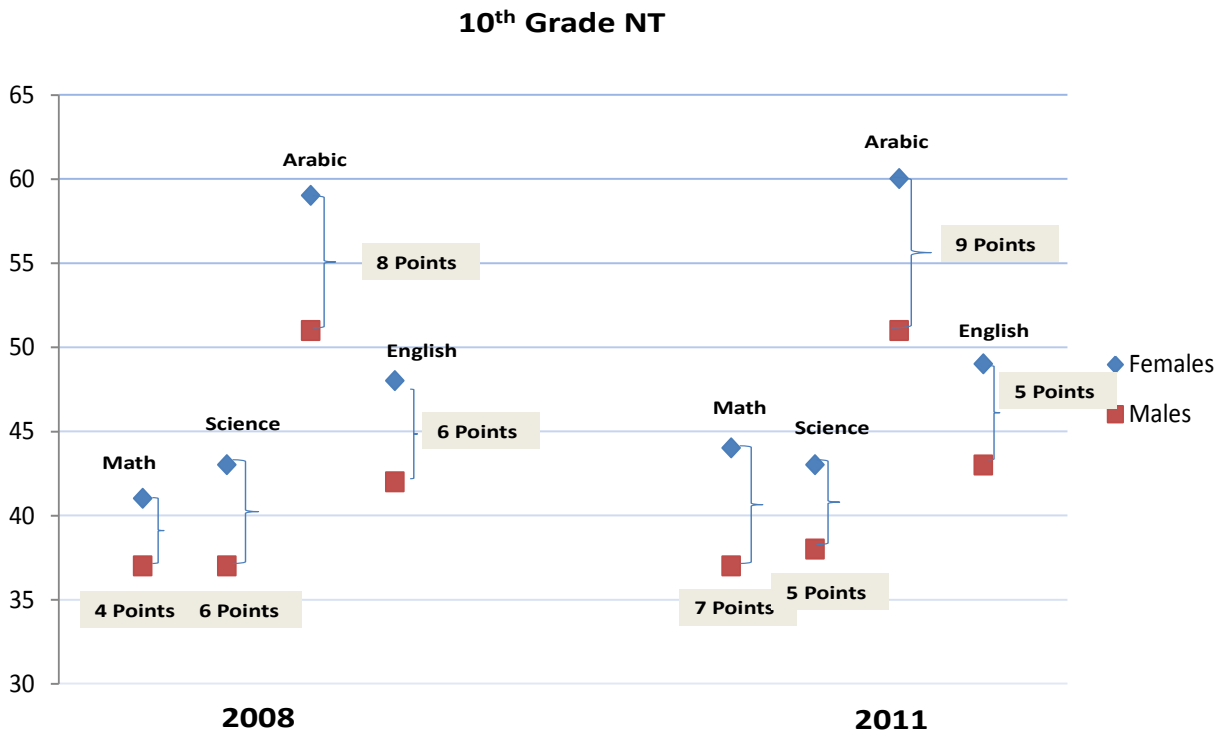
Figure 7: NT Results and the Achievement gap in Math, Science, Arabic, and English for 4th Graders in 2012



NT results and the gender achievement gap for 10th graders can be compared in two data collection cycles: 2008 and 2011. As Figure 8 indicates, overall student performance in Arabic ranged from 51-60 in both assessment cycles. Arabic was the highest scoring subject for both males and females. Performance in Math, Science, and English was low (averages below 50) for both sexes in 2008 and 2011.

In 2008, the smallest gender achievement gap was in Math, and in 2011, the smallest gender achievement gap was in Science. However, it is important to highlight that in both years the gap was statistically significant (4 and 5 points, respectively). The largest gaps in 2008 and 2011 were for Arabic scores (8 and 9 points, respectively).

Figure 8: NT Results and the Achievement Gap in Math, Science, Arabic, and English for 10th Graders in 2012



Finally, NT results presented in Figures 7 through 9 above (by grade) also allow us to observe once more the gender achievement gap in student assessment by grade. The differences between female and male scores in Math for 5th, 8th, and 10th graders were 2, 3, and 7 points, respectively. For Science, the differences in the same grades were 3, 8, and 5 points, respectively. For Arabic, the differences were 8, 8, and 9 points. Finally, for English, the differences were 5, 10, and 5 points. As for NAfKE, these data suggest that the gender achievement gap seems to increase as students move up in the education system, particularly if we compare the gaps between 4th and 8th grades. Between grades 8 and 10, the gap continued to increase for math and Arabic, but it decreased in Science and English for those same grades.

SUMMARY OF FINDINGS

- Females outperformed males in all subjects and grades across assessments and cycles. Further, the differences in scores between males and females are all statistically significant, except for PISA math scores in 2006 and 2009.

- The magnitude of the gender gap is not increasing consistently across assessments and subjects over time. For example, the magnitude of the gender gap in PISA Science results increased from 28 to 65 points from 2006-2009. However, the gap decreased to 42 points in 2012. Similar patterns can be observed in NAfKE Arabic and NT Science and English for 10th grade.

- NAfKE and NT results suggest that the gender gap seems to increase as students move up through the education system; i.e., at earlier grades, the gender gap is smaller than in later grades.

4.2 Gender Differences by School Type and Location

In addition to comparing overall male and female performance in the aforementioned assessments, we analyzed specific characteristics of subgroups of schools (by type and location). For this comparison, we examined the latest results for PISA (2012), TIMSS (2011), NAfKE (2011), and NT (2012). Moreover, we reported the mean scores for each subgroup and the effect size (magnitude of the difference between male and female scores) in order to ascertain whether differences between those two groups were considered "small, d = .2," "medium, d = .5," or "large, d = .8."[35] The effect size is simply a way of quantifying the size of the difference between two groups, taking into account the standard deviation.

As can be observed in Table 4, even though each assessment measures different kinds of specific subject knowledge at different grade levels, we can see subject- and gender-specific trends. First, we observe that even though females are outperforming males in all subjects and grades, the mean score differences and the effect sizes for Math are smaller than for other subjects. This finding is consistent the results presented in the previous section. Moreover, there are clear differences between government (MoE and other government) and non-government schools (private and UNRWA), particularly when we compare NT mean results and effect sizes, which are census-based.[36]

In NT Math, for example, the mean differences in scores between males and females are larger in MoE and *other government* schools (five points) than non-government schools (2 points). The effect size of 0.4 in MoE schools indicates that the magnitude of the differences between male and female scores is close to "medium." That difference is even larger in *other government* schools where the effect size is 0.5. In terms of interpretation, an effect size of 0.5 indicates that the average female score exceeds 69% of males' scores. Those numbers contrast sharply with UNRWA and private schools (two point differences). In the former, the effect size of 0.1 is considered small. In the latter, the effect size is also small at 0.2, suggesting that female and male scores are almost the same.

In NT Arabic, the mean differences between males´ and females´ scores are also larger in government than non-government schools (13 and 16 points difference in mean scores, respectively). The large differences also demonstrated by the effect sizes (0.7 and 0.8, respectively) suggest that in *other government* schools, for example, the scores of the average female exceed 79% of males´ scores. These numbers contrast with the gender difference in private and UNRWA schools (8 and 7 points difference in mean scores, respectively).

---

[35] Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates

[36] The effect size is most accurately calculated when the population standard deviation is imputed onto the formula.

These findings suggest student and school factors might affect male and female scores differently, depending on which school they attend. Later in this report we will explore some of the potential factors that affect gender differences in MoE schools.

Table 4: Males (M) and Females (F) Mean Scores and Effect Size (ES) of the Gender Difference in TIMSS, PISA, NAfKE, and NT by Subject and School Type

Subject	Authority (Number of Schools) (N[37])	TIMSS 2011			PISA 2012			NAfKE 2011									NT 2013		
		8th Grade			15 yr. old			5th Grade			9th Grade			11th Grade			10th Grade		
		F	*M*	*ES*	*F*	*M*	*ES*	*F*	*M*	*ES*	*F*	*M*	*ES*	*F*	*M*	*ES*	*F*	*M*	*ES*
Math	MoE (1,703)	412	372	0.4	387	361	0.4	30	26	0.3	33	29	0.3	23	19	0.3	34	29	0.4
	Other Government[38] (33)	--			365	357	0.1	28	25	0.2	35	30	0.4	25	21	0.3	35	30	0.5
	Private (306)	469	464	0.1	469	468	0.1	33	35	.09	43	43	.01	42	26	.80	42	40	0.2
	UNRWA[39] (121)	448	426	0.3	408	380	0.4	29	27	0.1	47	37	0.6	--	--	--	47	45	0.1
Arabic	MoE (1,699)	N/A			428	349	0.9	40	30	0.5	43	29	0.9	47	35	0.7	54	41	0.7
	Other Government (32)				405	339	0.6	44	27	0.9	43	30	0.9	53	35	1.3	57	41	0.8
	Private (309)				495	442	0.8	49	43	0.3	54	41	0.9	56	44	0.6	67	59	0.4
	UNRWA (121)				460	373	1.1	39	34	0.3	56	44	0.8	--	--	--	66	59	0.3
Science	MoE (1,727)	464	407	0.6	422	375	0.6	44	37	0.4	39	29	0.6	33	24	0.6	45	36	0.6
	Other Government (32)				414	382	0.3	48	36	0.6	40	29	0.7	30	23	0.4	48	37	0.7
	Private (307)	513	498	0.1	495	464	0.5	54	52	0.2	54	45	0.6	47	34	0.7	54	49	0.2
	UNRWA (121)	499	465	0.4	444	396	0.7	41	43	0.1	46	36	0.5	--	--	--	56	52	0.2
English	MoE (1,709)	N/A															48	36	0.6
	Other Government (32)																56	37	0.9
	Private (306)																75	64	0.5
	UNRWA (123)																60	54	0.2

[37] "N" represents the overall number of schools in Jordan (assessed through the NT). The numbers differ from the overall number of schools in the Kingdom and for each subject because data collection was conducted at different times, and sometimes enumerators could not reach all schools. For TIMSS, PISA, and NAfKE, representative samples are drawn from the overall number of schools.

[38] Other Government Schools is a category comprised of government schools managed by any of the following institutions: Ministry of Defense, Ministry of Social Affairs, Ministry of Higher Education, or Ministry of Awqaf and Islamic Affairs". Although they follow the same curriculum as MoE schools, they are not under MoE's management and their teachers are not paid, hired, or trained by the MoE.

[39] UNRWA (United Nations Relief and Works Agency) schools ensure all Palestinian refugee children have access to quality education. School children in UNRWA schools follow Jordan's curricula and textbooks. UNRWA supplements the curriculum with its own materials on human rights.

Finally, an investigation of the gender achievement gap in rural and urban areas in Jordan indicated that there were no clear differences in effect sizes between those two areas. That means that the gender achievement gap in urban areas was very similar to or the same as in rural areas.

To determine the differences among directorates, we focused on NT results only, as it would not be accurate or fair to report the magnitude of the gender achievement gap in directorates based on samples of schools. We classified directorates in three categories with regard to gender achievement gap: low, medium, and high -- based on the mean differences in scores between males and females and the effect size in those directorates. Maps 1-3 illustrate those findings by subject.

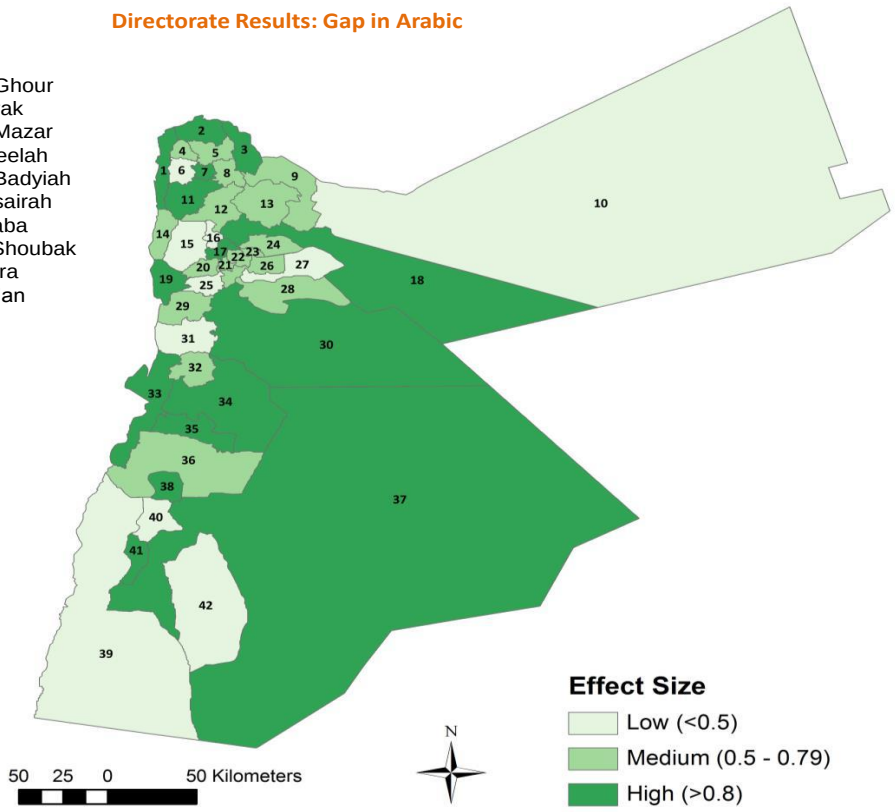
As expected, we found that females outperformed males in all directorates and subjects, except in Sahab (.07), Madaba (.09), Alqaser (.03), and South Badyiah (.5). In those four directorates, math score differences favored males, even though the effect sizes were not large.

We also noticed that there were fewer directorates with large gender achievement gaps in Math and Science than in Arabic and English. However, it is important to emphasize that in a couple of directorates the gender achievement gap was very high in all subjects. Such was the case in Ramtha and Busairah. In Ramtha, where the largest gaps were found, females' average scores were higher than over 90% of males in all subjects. Furthermore, some directorates present very high gender achievement gaps in specific subjects. In Arabic, for example, we found that North Aghwar (1.3) and Aljeezah (1.0) had the highest gender achievement gap. An effect size of 1.3, for example, indicated that the average female score in North Aghwar was higher than 90% of the males in that directorate. For English, the highest gaps could be found in North Mazar (1.2), Petra (1.2), and Al Shoubak (2.2). An effect size of 2.2 in Al Shoubak suggests that the average female score was higher than 98% of males. For Science, North Aghwar (1.6) and Petra (1.3) had the largest gaps. For math, South Shonah (.8), North Mazar (0.8), and South Shonah (0.8) had the largest gender achievement gaps.

Finally, some directorate results indicated there were very small gender achievement gaps in specific subjects. The lowest gender achievement gap in English was observed in Northeast Badyiah (0.1) and Alsalt (0.2). For math, Na'our (.07), Sahab (.07), Al-Mwaqar (0.1), Madaba (0.1), Bani Obeid (0.2), Northwest Badyiah (0.2), Alrusaifah (0.2), Alsalt (.05), South Ghour (0.2), Petra (0.2), Aqaba (0.2). For science, Sahab (0.1), Dheeban (0.2), Alkoorah (0.2), Jerash (.08), Ma'an (0.1), and South Badyiah (.01) had the lowest gaps. Dheeban (0.2), Alsalt (0.2), and Ma'an (0.1) had the lowest gaps in Arabic. Some of the small gaps might be explained by the overall very low scores in those subjects among males and females. However, they may also be indicative of small differences in educational measures adopted by all schools (male and female) in those locations. For a detailed presentation of the effect sizes for each directorate with respect to gender achievement gap in specific subjects, see Appendix B.

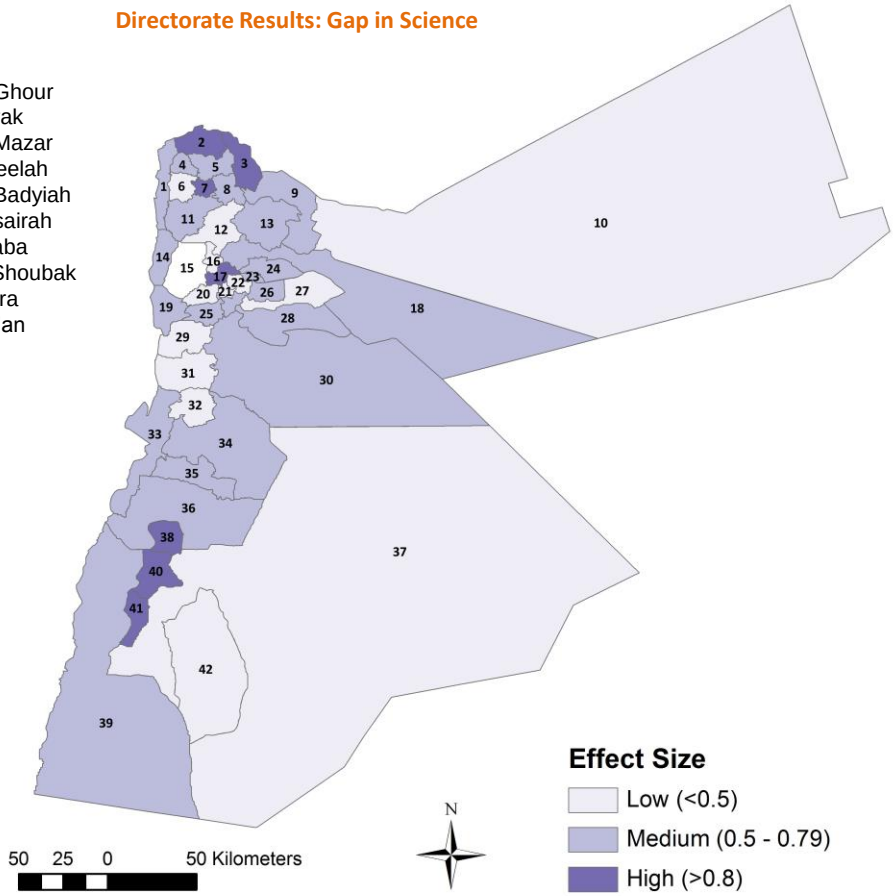
Directorate Results: Gap in Arabic

1 – N. Aghwar
2 – Bani Kenanah
3 – Ramtha
4 – Taibah & Westiah
5 – Irbid
6 – Alkoorah
7 – N. Mazar
8 – Bani Obeid
9 – N.W. Badyiah
10 – N.E. Badyiah
11 – Ajloun
12 – Jerash
13 – Mafraq
14 – Deir Ala
15 – Alsalt
16 – Ain Albasha
17 – Al Jama'ah
18 – Zarqa 2
19 – S. Shonah
20 – Wadi Alseer
21 – Amman
22 – Alqweismeh
23 – Alrusaifah
24 – Zarqa 1
25 – Na'our
26 – Markah
27 – Sahab
28 – Al-Mwaqar
29 – Madaba
30 – Aljeezah
31 – Dheeban
32 – Alqaser
33 – S. Ghour
34 – Karak
35 – S. Mazar
36 – Tafeelah
37 – S. Badyiah
38 – Busairah
39 – Aqaba
40 – Al Shoubak
41 – Petra
42 – Ma'an



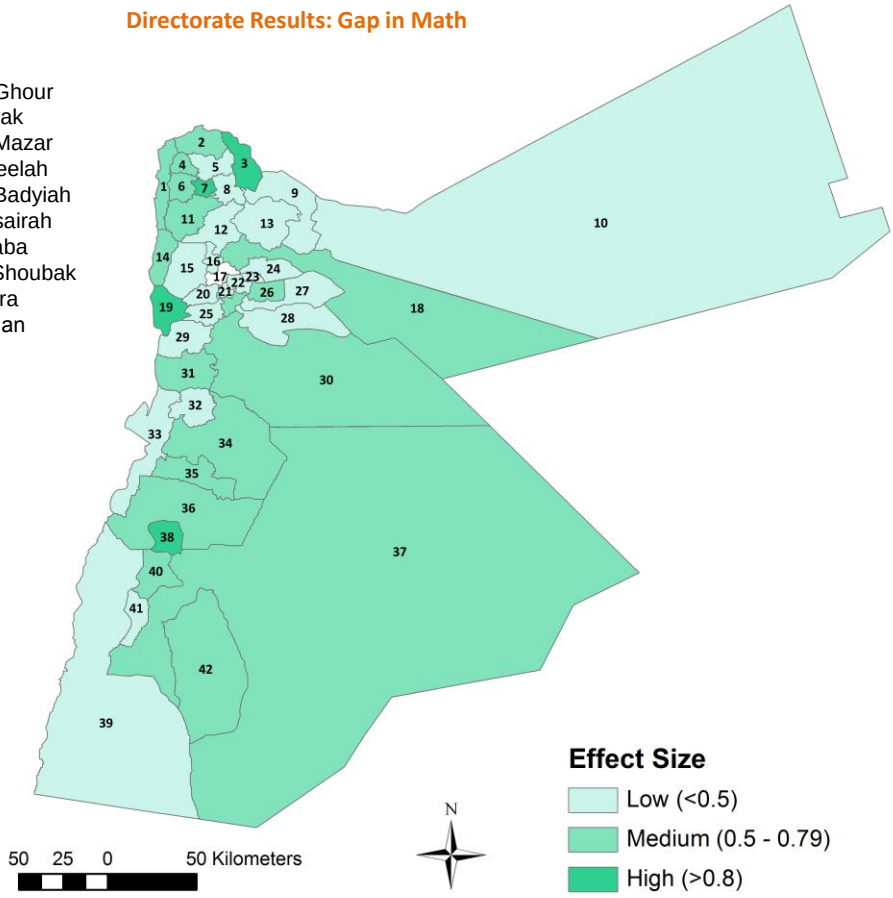
Directorate Results: Gap in Science

1 – N. Aghwar
2 – Bani Kenanah
3 – Ramtha
4 – Taibah & Westiah
5 – Irbid
6 – Alkoorah
7 – N. Mazar
8 – Bani Obeid
9 – N.W. Badyiah
10 – N.E. Badyiah
11 – Ajloun
12 – Jerash
13 – Mafraq
14 – Deir Ala
15 – Alsalt
16 – Ain Albasha
17 – Al Jama'ah
18 – Zarqa 2
19 – S. Shonah
20 – Wadi Alseer
21 – Amman
22 – Alqweismeh
23 – Alrusaifah
24 – Zarqa 1
25 – Na'our
26 – Markah
27 – Sahab
28 – Al-Mwaqar
29 – Madaba
30 – Aljeezah
31 – Dheeban
32 – Alqaser
33 – S. Ghour
34 – Karak
35 – S. Mazar
36 – Tafeelah
37 – S. Badyiah
38 – Busairah
39 – Aqaba
40 – Al Shoubak
41 – Petra
42 – Ma'an



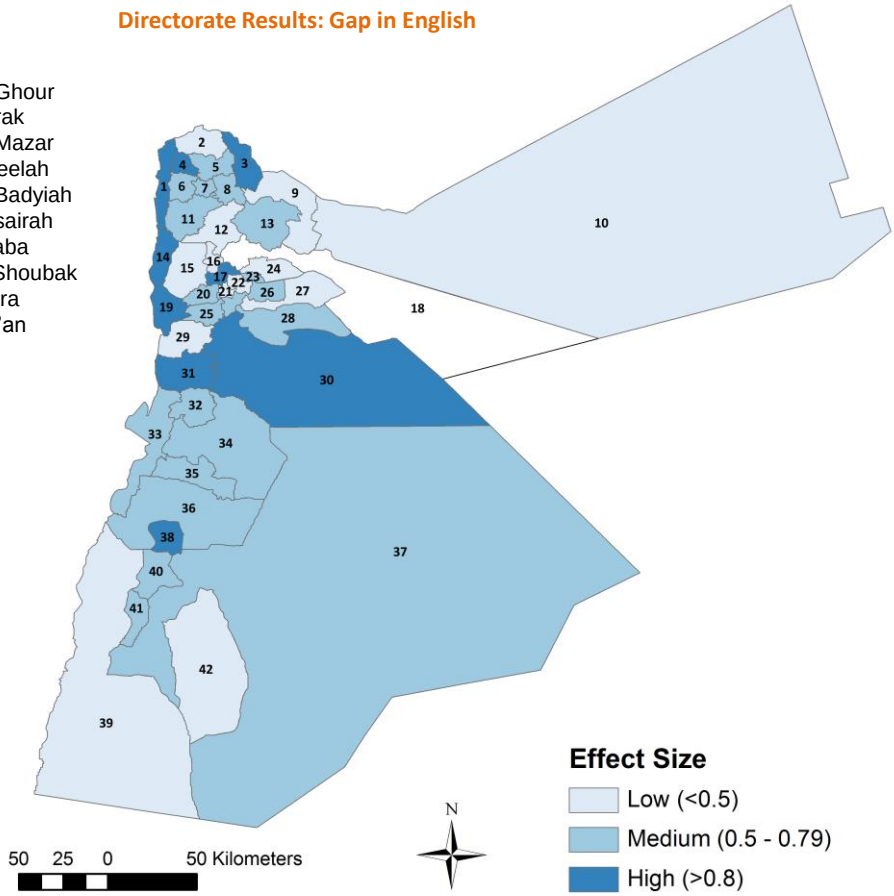
Directorate Results: Gap in Math

1 – N. Aghwar
2 – Bani Kenanah
3 – Ramtha
4 – Taibah & Westiah
5 – Irbid
6 – Alkoorah
7 – N. Mazar
8 – Bani Obeid
9 – N.W. Badyiah
10 – N.E. Badyiah
11 – Ajloun
12 – Jerash
13 – Mafraq
14 – Deir Ala
15 – Alsalt
16 – Ain Albasha
17 – Al Jama'ah
18 – Zarqa 2
19 – S. Shonah
20 – Wadi Alseer
21 – Amman
22 – Alqweismeh
23 – Alrusaifah
24 – Zarqa 1
25 – Na'our
26 – Markah
27 – Sahab
28 – Al-Mwaqar
29 – Madaba
30 – Aljeezah
31 – Dheeban
32 – Alqaser
33 – S. Ghour
34 – Karak
35 – S. Mazar
36 – Tafeelah
37 – S. Badyiah
38 – Busairah
39 – Aqaba
40 – Al Shoubak
41 – Petra
42 – Ma'an



Directorate Results: Gap in English

1 – N. Aghwar
2 – Bani Kenanah
3 – Ramtha
4 – Taibah & Westiah
5 – Irbid
6 – Alkoorah
7 – N. Mazar
8 – Bani Obeid
9 – N.W. Badyiah
10 – N.E. Badyiah
11 – Ajloun
12 – Jerash
13 – Mafraq
14 – Deir Ala
15 – Alsalt
16 – Ain Albasha
17 – Al Jama'ah
18 – Zarqa 2
19 – S. Shonah
20 – Wadi Alseer
21 – Amman
22 – Alqweismeh
23 – Alrusaifah
24 – Zarqa 1
25 – Na'our
26 – Markah
27 – Sahab
28 – Al-Mwaqar
29 – Madaba
30 – Aljeezah
31 – Dheeban
32 – Alqaser
33 – S. Ghour
34 – Karak
35 – S. Mazar
36 – Tafeelah
37 – S. Badyiah
38 – Busairah
39 – Aqaba
40 – Al Shoubak
41 – Petra
42 – Ma'an



Summary of Findings

- Overall, the gender gap was higher in government schools (MoE and other government) than non-government schools (private and UNRWA), particularly when we compared NT results, which are census-based.

- Females outperformed males in all directorates and subjects, except in Sahab (.07), Madaba (.09), Alqasr (.03), and South Badyiah (.5). In those four directorates, math score differences favored males, even though the effect sizes were not large.

- In a couple of directorates the gender gap was very high in all subjects, i.e., in Ramtha and Bsaira.

- North Aghwar (1.3) and Aljeezah (1.3) had the highest gender gap differences in *Arabic* scores.

- North Mazar (1.2), Petra (1.2), and Ashoubak (2.2) had the highest gender gap in *English* scores.

- North Alaghwar (1.6) and Petra (1.3) had the largest gaps in science.

- North Mazar (0.8), South Shounah (0.8), had the largest gender gap in math.

4.2 Factors Associated with Gender Achievement Gap

To understand some of the factors associated with the gender gap in students' achievement, we conducted both quantitative and qualitative analyses, which will be presented below.

The research team chose to utilize two-level Hierarchical Liner Modelling (HLM) as the preferred quantitative technique to identify student- and school-level data that might help explain gender differences in achievement. As explained in the methodology section, HLM was conducted for each assessment in one specific subject in the most recent cycle:

- The model for TIMSS (2011) scores focused on *Science 8th grade* scores.
- The model for PISA (2012) focused on *Math for 15 year-olds, i.e., 10th grade.*
- The model for NAfKE focused on Arabic *in 5th grade.*

Variables Associated with Student Performance and Gender Differences

To select the variables we would introduce into the TIMSS, PISA, or NAfKE HLM models, we examined TIMSS (2011), NAfKE (2011), and PISA (2012) assessment results and background questionnaires, which were designed to gather information about students, teachers, parents, and/or school characteristics. Selected male and female average scores and variables/themes from those questionnaires are presented in Table 5. All themes/variables presented in Table 5 were significantly correlated with students' scores in the aforementioned tests. In addition, there were significant differences between males and females with regards to all those variables/themes, except for SES.

Table 5 shows that students, teachers, and principals in female schools reported behaviors and attitudes that were more conducive to learning than their counterparts in male schools. For example, female students were more likely to do homework, and had higher academic ambitions, experienced less repetition, and skipped classes less often than male students. Conversely, males reported higher general use of computer at home, school, and for socialization (e-mail, social networks, collaborative games, etc). Males were also more likely to receive out-of-school lessons (probably as a remedial intervention) and have less parental involvement in their education than females.

At the school level, differences between male and female schools were also apparent. Female teachers reported higher job satisfaction than their male counterparts. Female schools were also reported to be safer than male schools. In addition, female schools seemed to have more well-prepared teachers, better student-teacher relations, less disrupted teaching, and more supervision of homework. Finally, male principals reported fewer resources in their schools than female principals. Female principals reported closer supervision of teachers, less teacher turnover, and higher parental participation in female schools. With respect to teachers' utilization of student-centered teaching style, the results from NAfKE suggest that females were more likely to utilize student-centered methodologies than males. However, the reverse was found when we analyzed PISA responses. The difference might be explained

by the subjects those teachers taught. PISA focused on math teachers, while NAfKE focused on Arabic teachers. In TIMSS, female teachers reported a higher use of questioning to elicit reasoning and explanations from students.

These two distinct realities between males and females presented in Table 5 provided an interesting insight into male and female behaviors, attitudes, and their school characteristics. However, it was necessary to conduct HLM to determine the extent to which student and school characteristics, including gender, could explain male and female performance in TIMSS, PISA, and NAfKE, when controlling for all other factors.

Table 5: Variables and Themes Associated with Academic Performance and Gender Differences

Themes/Variables	Assessments	Females	Males
Student-Related			
TIMSS 2011 average science scores (SE)	TIMSS (0-1000)	464 (9.8)	407 (7.5)
PISA 2012 Average math scores (SE)	PISA (0-1000)	387 (2.4)	361 (5.3)
NAfKE 2011 (Arabic Average Scores among 5th Graders)	NAfKE (0-100)	40 (0.7)	30 (0.8)
Participation in out-of–school lessons—tutoring, studying with family member or partner	NAfKE	30.1%	41.3%
Do homework	NAfKE	66.8%	50.6%
Use of computer	PISA	48.9%	52.2%
	TIMSS	58.6%	63.1%
	NAfKE	32.6%	35.0%
Sense of belonging (feel like an outsider at school).	PISA (Scores 0-15)	M=3.1 (.59)	M=2.8 (.62)
Grade repetition	PISA	5.2%	12.7%
Skips classes	PISA	32.4%	52.3%
Academic ambitions (desires university degree or higher)	TIMSS	67.1%	61.5%
Student motivation	NAfKE	54.8%	37.0%
Parental involvement in child's education	NAfKE (Scores 0-1) (Involvement in school activities)	M=0.57 (.50)	M=0.33 (.47)
	TIMSS (homework)	57.4%	61.3%
School-Related			
SES (socio-economic status): mean and standard deviation	TIMSS	9.5 (0.0)	9.5 (0.0)
	PISA	0.56 (0.0)	0.48 (0.0)
	NAfKE	11.9 (0.1)	11.6 (0.1)
Student-centered teaching style (composite scores)	PISA (Scale 0-4)	M=2.7 (.55)	M=2.8 (.62)
	NAfKE (SCALE 0-1)	M=0.62 (.49)	M=0.32 (.47)
Teacher uses questioning to elicit reasons and explanations	TIMSS	90.9%	72.1%
Lack of qualified math teachers	PISA	44.1%	47.6%
Teachers are not adequately prepared for classes	PISA	28.4%	45.6%
Teacher turnover	NAfKE	10.1%	27.1%
Higher job satisfaction among teachers	NAfKE	85.0%	58.9%
	TIMSS	82.1%	60.9%
Disruptive students limit ability to teach	TIMSS	78.2%	87.2%
School safety (student feels safe when s/he is in school; Students are not made fun of or called names at school; students are respectful of the teachers; school's security policies and practices are sufficient) (composite score transformed into dichotomous variable)	TIMSS	40.4%	70.3%
Inadequate school facilities and resources	NAfKE	49.0%	64.6%
Teachers monitor homework	TIMSS	87.2%	74.5%
Poor student-teacher relations	PISA	24%	30.2%
	TIMSS	19.4%	39.7%
Principals' monitoring teachers' activities	TIMSS	95.6%	83.8%
School extracurricular activities: Music/Theater	PISA	70.9%	28.5%
School extracurricular activities: Art	PISA	62.5%	37.7%

Results of HLM Model for TIMSS, NAfKE, and PISA results

The process used to fit multiple HLM models for TIMSS (science), PISA (math) and NAfKE (Arabic) scores was similar. We conducted a two level analysis, at the student and school levels. Below we will present the results for each HLM model separately.

TIMSS 2011 (SCIENCE)

Model 0, also known as the null model, presented in Table 6, represents the level of variance attributed to Level 1 (students) and Level 2 (schools). Using the information (for Levels 1 and 2) estimated in the null model, we were able to calculate intra-class correlation coefficients (ICC), which represent the proportion of variance in student scores between schools[40]. In the case of TIMSS, ICC revealed that 28.8% of the total variance in TIMSS science scores was attributable to school effects. We consider this school level proportion of the variance to be quite high. In simple terms the quality of schools made a difference regarding student performance, and the variation among the schools was large in Jordan. Student achievement rates depended significantly on which schools they attended. Model 1 demonstrates the simple effect of school gender on the outcome variable (student scores). The results showed that school gender is an important predictor of achievement. In fact, 27%[41] of the true variance between schools in TIMSS science achievement was accounted for by school gender. In other words, school gender is significantly and positively related to mean achievement. Female schools score, on average, 56.8 points higher than male schools in science.

Model 2 determines the effects of gender and SES at the school level. Those two factors are shown to affect students´ science scores in Jordan. School SES played an important role in students' achievement when controlling for gender. More specifically, the higher the school SES was, the higher were the students' average scores when controlling for gender effects. This correlation held true for both females' and males' schools. Thus, gender remains a significant predictor of student achievement, even when we control for SES.

The final model demonstrates that both student and school level variables contributed to students' achievement. At the student level, parental control over homework, students' use of computer, academic ambitions, and school safety have all had an effect on students' science achievement in TIMSS. The coefficients for each independent variable reflect their relative effect when controlling for other variables in the model. Further, all slope coefficients are positive, which suggests that on average, and holding all other variables in the model constant: 1) students who had high academic ambitions, i.e., intended to go to university and/or beyond, scored 37.7 points higher than students who did not intend to study beyond high-school; 2) students who felt safe[42] in schools scored 27.7 points higher than

[40] For the detailed statistics used for the computation and significance of the Null Models, see Appendix C.

[41] For the detailed explanation used for the computation of Model 1 variance, see Appendix C.

[42] Students who felt safe reported: feeling safe when s/he are at school; not being made fun of or called names at school; being respectful of the teachers; being in a school with sufficient security policies and practices

students who felt unsafe; 3) students whose parents made sure s/he set time aside for homework scored 26.9 points higher than students whose parents did not engage in that behavior; and 4) students who made frequent use of computers at home scored 13.5 points higher than students who did not use computers (or use them infrequently).

At the school level, gender and SES were the only significant predictors of outcome when controlling for all other variables. On average, and holding SES constant, female schools scored 48.5 points higher than male schools. In addition, schools with higher SES scored higher than schools with lower SES when holding gender constant. In comparison with the null model, the final student model explained approximately 9%[43] of the variance at the student level and 54% of the variance at the school level.

Table 6: Student and School Effects on Students Science Achievement in TIMSS 2011.

TIMSS 2011 Variables	Model 0	Model 1	Model 2	Full Model[44]
Student level				
Parental control over homework (BSBG11C)				26.9**
Use of computer (BSBG10A)				13.5**
Academic Ambitions BSBG07				37.7**
School Safety				27.7**
School level				
Gender		56.8**	58.4**	48.5**
SES			43.1**	36.1**
Level-1 Variance (student)	7797.3	7801.7	7807.9	7060.3
Level-2 Variance (school)	3148.9	2307.5	1586.9	1442.4

**p<.001

In addition, we believe that qualitative data deepened our understanding of the results obtained in the HLM model. Regarding academic expectations, qualitative data clearly suggests that females had higher academic aspirations than males because they associated education with high personal and financial returns. As it was noted in focus group discussions, "a university certificate means freedom for girls" (Focus group discussions, students, Irbid Elementary School for Females, UNRWA, Irbid, January, 2014). A "university degree may help [us] get a respectful career in the future" (Focus group discussions, students, Kufr-rahta Basic Schools for Females, MoE, Irbid, January, 2014). Parents and teachers also believed that providing higher education for females was a more "necessary" investment than providing that kind of education for males. When parents in FGDs were asked whether males and females would probably get a better income if they obtained a university degree, only 32.2% thought males would earn

[43] For detailed explanation used for the computation of Model 1 variance, see Appendix C.

[44] All TIMSS variables presented in Table 5 were entered into the model progressively to assess their contribution to students' TIMSS scores. However, for the sake of brevity, we only present the significant student and school-level variables in the final model.

better incomes with a university degree. However, 62.6% of parents stated that females with a university degree would earn higher incomes. Teachers also believed that higher education has a higher effect on females' lives than males'. While 50% of male teachers (n=54) stated that university degrees were important for males to get good jobs, 68.5% of those same teachers believed that university degrees are important for females. Although female teachers reported overall higher confidence in the returns of a university degree for both males and females, they also believed that the returns for females would be higher than for males (83.1% and 70.8%, respectively). As one principal explained, "a female student believes that having a certificate has social value. She has a goal to achieve. On the other hand, male students have many other opportunities, such as the military." (Focus group notes, principal, Natfeh Basic School for Males, MoE, Irbid, January, 2014).

Since we also wanted to ascertain whether a safe school environment played a role in student performance in science, we had to examine the differences between males' and females' schools with regard that variable. Focus group discussions and questionnaire results reinforced the belief that some student and school characteristics were more prevalent in male schools than female schools. For example, we found that 56% of male students (n=56) reported suffering some sort of bullying in school in the past year. Among females (n=66), the percentage was much lower -- only 11%. During discussions, students confirmed that "violence is present in all forms and it is caused by mismanagement of school, bad parenting at home, sometimes excessive racism, lack of acceptance of the other." (Focus group notes, students, Al-Jbuaiha Secondary, Amman, 12/30/2013). Female and male teachers also reported different levels of violence in their schools. 68% of male teachers (n=54) reported they had witnessed physical violence among students in their schools during the past year. That contrasts with only 22% of female teachers who witnessed the same problem in female schools. Further, male teachers reported a higher incidence of students' verbal or physical threats against them (24% and 6%, respectively). The percentage of female teachers who had been verbally or physically threated in female schools was much lower (0% and 1.5%, respectively). Parents (n=115), principals (n=12), and supervisors (n=43) all believed that social discrimination and violence within the home were some of causes of the violence they witnessed in schools.

Although not all variables presented in Table 5 were statistically significant when entered into an HLM model, it is important to highlight that some should be taken into consideration in all discussions about gender achievement gap in Jordan. In fact, some themes were mentioned and elaborated fully by focus group participants. One such variable was teachers' motivation (or professional satisfaction). A simple tabulation of the TIMSS background questionnaire results indicated that 17.7% of female teachers (n=113) were dissatisfied with teaching. In contrast, 40.6% of male teachers shared that same opinion (n=101). Teachers who participated in focus groups and filled out the evaluation questionnaire reported a very similar trend. A full 60% of males reported being dissatisfied with their jobs as teachers (n=54), while only 20% (n=65) of female teachers were not satisfied. This is a very dramatic gap between female and male teachers.

Teachers' professional dissatisfaction might reflect the way in which they teach, but how students perceive teachers might be just as important. We found that only 40% of male students (n=56) believed that their teachers cared about how well they did in schools. However, that percentage was much

higher for females (74.4%, n=66). It is essential to understand how teachers' motivation and satisfaction ultimately affect students' ambitions and desire to succeed in schools. If we relate male teachers' professional dissatisfaction with their low expectations for males (as they believe a higher degree carries a much lower value for males than females) we might conclude that the education system for males, as currently set up, is much less conducive for learning than for females. Unless there are more incentives for males to perform at a higher level in schools (as teachers or as students), it is unlikely that they will feel the need to invest time and energy in their academic pursuits. As one teacher in a female school pointed out:

> "Male schools do not offer students quality education like female schools, since male teachers depend on the private lessons to increase their income. They are committed to another job. Thus, the teacher`s performance and his effort inside the classroom is poor and insufficient. They are not committed or interested in the lesson plan, or even in fulfilling the learning outcomes. The administration in male schools has no interest in following up the teacher technically. As a result, teachers become uncommitted during school hours. That, in turn, affects negatively the completion of the curricula according to the pre-set plan and the performance of students." (Focus group notes, teachers, Iskan Al Jamaa' School for Females, Amman, 12/30/2013)

PISA (Math) 2012

Similar to the analyses presented above, Model 0, also known as the null model, presented in Table 7, represents the level of variance attributed to Level 1 (students) and Level 2 (schools). The ICC for the PISA null model, reveals that 23.5%[45] of the total variance in PISA math scores is attributable to school effects. This proportion of variance is lower than the one found in TIMSS, but it is still considered to be high. In simple terms, schools make a difference in student performance in math, and the variation among schools is large in Jordan. Students' math achievement results are directly related with the schools that they attend.

Model 1 demonstrates the simple effect of school gender on the outcome variable (student scores). Once again, the results showed that school gender was an important predictor of achievement. In fact, 14.3% of the true variance between schools in PISA math achievement was accounted for by school gender. In other words, school gender was significantly and positively related to mean achievement. In math, female schools scored, on average, 26 points higher than male schools.

Model 2 determines the effect of gender and SES at the school level. Those two factors clearly affected students´ Math scores in Jordan. School SES played an important role in students' achievement when controlling for gender. More specifically, the higher the school's SES was, the higher were the students' average scores, when controlling for gender effects. This correlation held true for both female and male

[45] For the detailed statistics used for the computation and significance of the Null Models, see Appendix C.

schools. Thus, gender remains a significant predictor of student achievement, even when we control for SES.

The final model demonstrates that, at the student level, computer use, a sense of belonging, and students' history of repetition all had an effect on students' math achievement in PISA. The coefficients for each independent variable reflect their relative effect when controlling for other variables in the model. Two slope coefficients were positive, which suggests that on average, and holding all other variables in the model constant: 1) students who used computers frequently for drawing, painting, using graphics program, and for networking scored 8.3 points higher than students who did not use computers (or used them infrequently); and 2) students who had a stronger sense of belonging within their schools (i.e., were not left out of school activities or did not feel lonely or "out-of-place") scored 12.9 points higher than students who felt the opposite. The negative slope observed in grade repetition suggests that on average, and holding all other variables in the model constant, students who had previously repeated a grade scored 59.9 points lower than students who had not.

At the school level, gender and SES were the only significant predictors of the outcome when controlling for all the other variables. On average, and holding SES constant, female schools scored 20.7 points higher than male schools. In addition, schools with a higher SES scored higher than schools with a lower SES, when holding gender constant. In comparison with the null model, the final student model explained approximately 12% of the variance at the student level and 36% of the variance at the school level.

Table 7: Student and School Variables Effects on Students Math Achievement in PISA 2012.

PISA 2012 Variables	Model 0	Model 1	Model 2	Full Model[46]
Student level				
Computer use				8.3**
Sense of belonging				12.9**
Grade repetition				-59.9**
School level				
Gender		26.1**	27.9**	20.7**
SES			28.7*	10.9**
Level-1 Variance (student)	3839.4	3839.2	3840.5	3380.5
Level-2 Variance (school)	1177.8	1009.3	860.7	753.2

**p<.001

[46] All NAfKE variables presented in Table 5 were entered into the model in a progressive manner in order to assess their contribution to students' Arabic math scores. However, for the sake of brevity, we only present the significant school-level variables in the final model. Because there were no significant variables at the student level, those results were omitted.

NAfKE (Arabic) 2011

Like the previous assessments presented above, Model 0, also known as the null model, presented in Table 8, represents the level of variance attributed to Level 1 (students) and Level 2 (schools). The ICC for the NAfKE model revealed that 30% of the total variance in 5th graders' Arabic scores was attributable to school effects. We believe that this proportion of variance is quite high. In simple terms, schools clearly made a difference in student performance in Arabic, and the variation among schools was large. In sum, students' achievement results in Arabic depended directly on which school they attended.

Model 1 demonstrates the simple effect of school gender on the outcome variable (student Arabic scores). Once again, the results showed that school gender was an important predictor of achievement. In fact, 30% of the true variance between schools in NAfKE Arabic achievement was accounted for by school gender. In other words, school gender was significantly and positively related to mean achievement. And in Arabic, female schools scored 9.3 points on average higher than male schools.

The NAfKE full model results presented in Table 8 suggest that only school gender and SES were significant predictors of Arabic scores when controlling for other school and student factors in the model. The fact that other variables (as presented in Table 5) were not significant, while holding gender and SES constant, does not mean they had no effect on student assessment results when they were analyzed independently. In fact, those variables may just be confounded with gender and SES. In simple terms, they were significantly correlated with achievement (outcome), but when they were added to the HLM model with other significant and correlated predictors, they became insignificant. We observed that female schools were expected to score an average 7.9 points higher than male schools in NAfKE Arabic assessment for 5th graders, controlling for all other factors. As expected, school SES played an important role in students' achievement when controlling for gender. More specifically, the higher the school SES was, the higher the students' average scores when controlling for gender effects. This correlation held true for both females' and males' schools. In sum, the final model demonstrated that, compared with the null model, it explained approximately 48% of the variance at the school level.

Table 8: Student and School Variables Effects on Students Arabic Achievement in NAfKE 2011

NAfKE 2011 Variables	Model 0	Model 1	Model 2 (Final)
School level			
School gender		9.3**	7.9**
SES			2.6**
Level-1 Variance (student)	280.1	279.6	280.0
Level-2 Variance (school)	68.5	44.6	47.9

**p<.001

5. CONCLUSION AND POLICY OPTIONS

The objective of this study was to determine whether there is an actual gender achievement gap in student achievement in Jordan and, if so, to identify the possible reasons for such a gap, if it exists. The results presented above confirm that the gender achievement gap is real and that females have been outperforming males in TIMSS, NAfKE, PISA, and NT in all subjects and grades for several assessment cycles. In fact, as mentioned earlier, the latest PISA results (2012) show that Jordan ranks "Number1"in gender gap among PISA participating countries. Furthermore, the gender achievement gap in favor of females exists in all Directorates and is more pronounced in government schools than non-government schools. Moreover, as students move up through the education system and reach higher grades, there is a tendency for the gender achievement gap to increase. However, despite the pervasiveness of the gender gap phenomenon in Jordan, we cannot say that the problem is increasing consistently overtime.

With respect to the factors that might explain this gender achievement gap, we can conclude that there are both student- and school-level factors responsible for the differences between the sexes. At the student level, parental control over homework, students' use of computers, academic ambitions, school safety, a sense of belonging, and students' history of repetition were shown to all have a clear effect on students' assessment scores in Jordan. At the school level, school gender and SES consistently appeared as significant factors that explained differences in students' performance. Female schools tended to outperform male schools, and similarly, schools with higher SES outperformed schools with lower SES. Focus group discussions also shed light on how those variables – combined with socio-cultural factors in Jordanian society -- contributed to males' and females' different performances in assessment tests. In addition, we learned that the violent environment in male schools, males' relative lack of academic ambition and incentives, male teachers overall lack of motivation to teach, and their relationship with students were key elements in explaining gender differences. In addition, focus group discussions gave us insight into the cultural beliefs and practices that might have an effect on males' and females' academic performance. Females' more "controlled" environment and more restricted social life may have increased their time for academic pursuits. In addition, women seemed to perceive higher financial and social returns from education than males. Those perceptions as well as the real labor opportunities for males will need to change in order to affect males' academic performance.

Despite the findings highlighting the gender achievement gap and its underlying causes, it is important to emphasize that Jordanian students (both males and females) have been performing below the average (center-point) in national and international assessments. That is particularly true in math and science. Therefore, we recommend that the MoE consider promoting the overall improvement of education for all students and not focus solely on males when it develops its educational policies for the Kingdom. More specifically, action plans should not reallocate resources from female schools to male schools. In fact, while the MoE should maintain and enhance female schools on an ongoing basis, it must heavily invest in male schools to accelerate the improvement of males' performance. If followed, the options presented below will have a very strong chance of improving the educational system, and more importantly, improved learning for all students and decrease the gender achievement gap.

1. ***Invest in teachers***. It will be necessary to (re)shape teaching profession requirements and career path options in order to attract talented individuals of both sexes and to stimulate current teachers to perform at high levels. Unless the teaching profession becomes more rewarding, not just financially, but also by offering a better work environment and career path, there will be no incentives for talented individuals to enter the profession or for current teachers to perform at the desired levels aimed by ERfKE II's stated levels or any educational reform project. Currently, male teachers are forced to hold multiple jobs to fulfill their social and financial obligations and that affects the time they can spend on preparing lessons and meeting the academic needs of male students. To improve the status of the teaching profession to attract and retain committed individuals, particularly males, we recommend several key measures, including.[47] a) Engaging in serious dialogue that will promote strategies to provide financial incentives to all teachers so both males and females can embrace the teaching profession as their primary (or sole) source of income. Financial incentives must be combined with accountability systems to ensure results. Teachers should be ranked according to their performance in the classroom; b) Enhancing the status of the teaching profession by utilizing the media to change teacher´s image in society; c) Offering scaffolding for low-skilled teachers through scripted lesson plans and coaching on curriculum; d) Strengthening pre- and in-service training and finding appropriate times to conduct those trainings; e) Promoting peer-led learning for teachers; and f) activating the role of supervisors as coaches. Although that has been a key strategy under the School and Directorate Development Program (SDDP), much remains to be done with regards to increasing the number of supervisors and training them effectively. Further, supervisors have little incentive to supervise male schools.

2. ***Invest in principals***. A strategic plan to select principals should also be put in place. Ideally, principals should be selected from a pool of excellent teachers or supervisors in the system and be trained accordingly. Best principals should be assigned to most challenging male schools and receive adequate compensation for working in difficult environments. An accountability system with incentives mechanisms should be in place to encourage leaders to perform at their best. Further, principals should share administrative tasks with a vice-principal (or equivalent), so s/he can supervise teachers to improve instruction.

3. ***Improve overall safety in schools***. Safety is a major concern in male schools. Educational environments where students, teachers, and administrators feel unsafe decrease their motivation to learn and work. Therefore, it will be essential to address this issue in a multifaceted manner in order to:

(a) Increase awareness about current school-based initiatives to deter school violence in Jordan. Currently, some male schools have been experimenting with strategies to involve parents and

[47] These suggestions were presented in detail in the McKinsey *Report How the World Best-Performing School Systems Come Out on Top* (2207) by Mona Mourshed, as essential measures undertaken by high-performing school systems around the world.

communities in school life in an attempt to make schools safer and to improve the overall quality of education. Those experiences have been yielding promising results[48] by emphasizing family involvement and educating parents about values of a male student. Some schools in Jordan have held workshops that adopted an Islamic approach to school violence and vandalism where they educate both the parents and students on the values of school preservation and respect to the teacher. The MoE might benefit from such local initiatives by promoting a national forum for shared experiences about improving safety in schools;

(b) Activate councils at the central and directorate levels to search for innovative solutions and exchange information regularly on how to tackle the safety problem. Under ERfKE II Component 1, the MoE established educational councils at the school cluster and directorate levels. In addition, parents´ and students´ councils have also been established. However, not all councils are in fact active. One key strategy adopted by the councils should be to listen to and include children's views, perspectives, and experiences before developing ways to deter violence. A reward system could be implemented to recognize successful initiative to decrease violence. The committees would suggest and approve a general framework for the strategies and mechanisms that are expected to contribute to the improvement of the overall safety of schools. The MoE would have to visit schools often to ensure councils are active.

(c) Ensure there is appropriate legislation that will safeguard the rights of victims (teachers, administrators, and students);

(d) Create an accountability system in cooperation with schools, communities, and law enforcement;

(e) Involve relevant institutions (such as the judicial system, social services, and/or Islamic religious authorities) in forming a partnership with schools and supporting families and communities when violence extends beyond the school premises;

(f) Coordinate efforts with existing programs (funded locally or by international agencies) to avoid duplication of efforts and to enhance the effects of the initiatives;

(g) Equip teachers, principals, and supervisors with relevant competencies and decision-making mechanisms to deal with violence in schools. Both capacity building and development of mechanisms to deal with violence should be included in the school development and school action plans.

[48] Alaref, J., Linnemann, H., Quota, M., & Woolcock, M. (2014). Jordan Education Case Study, for MENA Regional Report. The World Bank.

(h) Activate the role of school counselors to offer psychological support to students and parents and to generate reflection about the roots of violence and how students and parents may react to it.

(i) Increase awareness about violence by utilizing the media, publications inside the school, and university radio.

4. ***Make education more relevant to the needs of students***. Based on the findings presented in this report it is clear that parents, teachers, and students perceive the returns of education for males and females differently, due to existing socio-cultural and economic factors. However, we strongly believe that education is a critical variable for enabling males and females to gain the skills necessary to meet the demands of the labor market, no matter what profession they choose. Therefore, it becomes necessary to adapt teaching methodologies and streamline the curriculum in a way that emphasizes the direct application of education to "real life" necessities and skills that do not fall under gender stereotypical lines. It is clear that the traditional notions of teaching, even when they are "student-centered," may not be sufficient to raise students' interest in learning and thus their performance. In a technology-driven world it means better utilization of ICT in the classroom to enhance students' familiarity with that tool, and application of more innovative strategies to capture the attention and interest of students, particularly males. Close relationships with businesses can enhance schools' understanding of the private sector expectations and also improve students' awareness about the importance of education and motivation, and their likely direct relationships with future earnings.

5. ***Increase students´ motivation to learn.*** This study has shown that there are significant differences between female and male learning environments, which might negatively affect males´ motivation to attend school and learn. Some measures to change that situation might include: a) improve the physical learning environment of male schools by ensuring schools are clean, organized, and have the necessary learning materials and resources; b) provide extra-curricular activities for male students (such as sports, army sponsored programs, "values" and "morals" classes); c) create remedial programs to support failing students; d) reward high-performing students.

6. ***Initiate interventions in directorates with large gender performance gaps***. Although the gender achievement gap is a phenomenon observed across Jordan, some locations identified in this report are in need of immediate actions to ameliorate the problem. Education programs currently implemented in those areas might expand their scope of work to address the gender inequity problems in student achievement. For example, school principals, supervisors, and MoE staff from the central Gender Department could assist male schools in developing specific action plans that contain initiatives and strategies designed to decrease the gender achievement gap.

7. ***Create initiatives to involve parents and communities in children's education.*** Many parents do not visit their children's schools and are not aware of their children's school performance in national assessments. It is recommended that the MoE invest in communication campaigns and

strategies to invite parents and community members to visit the school. In addition to sponsoring *open school days* for community activities, it will be necessary to develop mechanisms to share school-level assessment results with both parents and the communities at large in order to increase school accountability.

As pointed out at the beginning of the report, while women´s higher performance on standardized tests and at universities reflect strong improvements in educational development in Jordan (and other countries), the low performance of males in school as well as the low participation of women in the economic and political arenas are genuine reasons for concern. In fact, in a male dominated society, the stagnation of male educational attainment not only impairs future earning opportunities, but it also reinforces gender stereotypes, as less educated men tend to have stricter views about gender roles and may resist and resent women in leadership positions[49]. Therefore, it is evident that reducing the gender achievement gap will require changes in the overall quality of the education system and targeted interventions for males. Furthermore, changes will have to take into account deep-seated beliefs about the teaching profession, the ways in which students learn, and the roles of men and women within the Jordanian society. In the final analysis, the leaders of Jordan's public education system will need to engage in long-term planning with respect to governance, incentives, and funding structures – and not take political shortcuts. They will also need to be realistic and pragmatic in developing their implementation strategies and consult with all relevant stakeholders in order to come up with a plan that will have the greatest chance of success.

[49] Barker, G., Verma, R., Crownover, J., Segundo, M., Fonseca, V., Contreras, J.M., Heilman, B., & Awlak, P. (2012). Boys and education in the global south: Emerging vulnerabilities and new opportunities for promoting changes in gender norms. Journal of Boyhood Studies, 6(1-2), 137-150.

APPENDIX A

Schools were selected from the pool of schools that had participated in TIMSS 2011, PISA 2012, and NAfKE 2011. Each school should have participated in at least two of the assessments, with higher preference given to schools that participated in all three studies. A special index was also created to capture this requirement. The distribution of the schools according to the requirement is shown in the table below (N: NAfKE, T:TIMSS, P:PISA). A sample of 12 specific schools was selected to fulfill the Sample Selection Strategy (Prepared by MEP/NCHRD) as shown in the table:

Index	Frequency	Percent	Valid Percent	Cumulative Percent
N	195	31.7	31.8	31.8
T	77	12.5	12.5	44.3
P	165	26.8	26.9	71.2
N&T	109	17.7	17.8	88.9
N&P	27	4.4	4.4	93.3
T&P	12	2.0	2.0	95.3
N&T&P	29	4.7	4.7	100.0
Total	614	99.8	100.0	

APPENDIX B

Subject	High Gender Gap = Effect Size ≥ .8	Medium Gender Gap= Effect Size ≥ .5	Low Gender Gap= Effect Size < .5
English	Aljeezah (1.0) Dheeban (.9) Taibah & Westiah (.9) **Ramtha (1.7)** North Aghwar (1.3) Deir Ala (.9) South Shonah (.9) **Busairah (.9)**	Markah (0.6) Wadi Alseer (0.5) Na'our (0.5) Al- Mwaqar (0.5) Irbid (0.5) Banie Obeid (0.6) North Mazar (0.66) Alkoourah (0.6) Ajloun (0.6) Mafraq (0.5) Alrusaifah (0.5); Karak (0.5); South Mazar (0.5); Alqaser (0.7) South Ghour (0.7) Tafeelah (0.6); Petra (0.5) Al Shoubak (0.7) South Badyiah (0.5)	Amman (0.4) Al Jama'ah (0.4) Alqweismeh (0.3) Sahab (0.4) Madaba (0.4) Bani Kenanah (0.4) Jerash (0.4) Northeast Badyiah (0.1) Northwest Badyiah (0.4) Zarqa1 (0.4) Alsalt (0.2) Ain Albasha (0.3) Ma'an (0.3) Aqaba (0.3)
Math ***Males outperform Females**	**Ramtha (1.5)** **Busairah (1.0)** **South Shonah (0.8)** **North Mazar (0.8)**	**South Badyiah (.5) Markah (0.5) Aljeezah (0.5) Dheeban (0.7) Taibah & Westiah (0.5) Alkoourah (0.6) Bani Kenanah (0.5) North Aghwar (0.5) Ajloun (0.6) Zarqa 2 (0.6) Deir Alla (0.6) Karak (0.6) South Mazar (0.6) Tafeelah (0.5) Ma'an (0.5) Al Shoubak (0.6)	**Sahab (.07) **Madaba (.09) **Alqaser (.03) Amman (0.3) Al Jama'ah (0.3) Alqweismeh (0.3) Wadi Alseer (0.4) Na'our (0.07) Al-Mwaqar (0.1) Madaba (0.1) Irbid (0.3) Bani Obeid (0.2) Jerash (0.3) Mafraq (0.3) Northeast Badyia (0.2) Northwest Badyia (0.4) Zarqa 1 (0.3); Alrusaifah (0.2) Alsalt (0.05); Ain Albasha (0.3) South Ghour (0.2); Petra (0.2) Aqaba (0.2)
Science	North Mazar (1.2) **Ramtha (1.9)** **Busairah (1.2)** Petra (1.2) Al Shoubak (2.2)	Markah (0.7) Na'our (0.5) Aljeezah (0.7) Al-Mwaqar (0.7) Irbid (0.5)	Amman (0.4) Al Jama'ah (0.4) Alqwaismeh (0.3) Wadi Assa (0.4) Sahab (0.10)

	Bani Kenanah (0.8)	Bani Obeid (0.6) Taibah & Westiah (0.5) North Aghwar (0.7) Ajloun (0.6) Mafraq (0.7) Northwest Badyiah (0.6) Zarqa 1 (0.5) Zarqa 2 (0.6) Alrusaifah (0.5) Deir Alla (0.6) South Shonah (0.7) Karak (0.6) South Mazar (0.5) South Ghour (0.6) Tafeelah (0.5) Aqaba (0.5)	Madaba (0.4) Dheeban (0.2) Alkoourah (0.2) Jerash (0.08) Northeast Badyiah (0.3) Alsalt Ain Albasha (0.3) Alqaser (0.4) Ma'an (0.1) South Badyiah (0.01)
Arabic	North Mazar (1.0) **Ramtha (1.8)** North Aghwar (1.6) Ajloun (0.9) **Busairah (1.0)** Petra (1.3) South Badyiah (0.9) Bani Kananah (0.8) Aljeezah (0.8) Zarqa2 (0.8) South Shonah (0.8) Karak (0.8) South Mazar (0.8) South Ghour (0.8)	Amman (0.6) Alqweismeh (0.5) Markah (0.7) Wadi Alseer (0.6) Al-Mwaqar (0.7) Madaba (0.7) Irbid (0.6) Bani Obeid (0.6) Taibah & Westiah (0.5) Jerash (0.6) Mafraq (0.6) Ajloun (0.5) Zarqa1 (0.6) Alrusaifah (0.6) Deir Alla (0.5) Alqaser (0.7) Tafeelah (0.5)	Al Jama'ah (0.4) Na'our (0.4) Sahab (0.3) Dheeban (0.2) Alkoorah (0.3) Northeast Badyia (0.3) Alsalt (0.2) Ain Albasha (0.4) Ma'an (0.1) Al Shoubak (0.4) Aqaba (0.3)

APPENDIX C

Table 6: Results from the one-way ANOVA Models for TIMSS, PISA, and NAfKE

TIMSS SCIENCE				
Fixed Effect	*Coefficient*	*se*		
Average School Mean	*432.4*	*7.7*		
Random Effect	*Variance*	*d.f.*	X^2	*p-value*
Level 2 (school)	3148.9	179	2654.9	<0.001
Level 1 (student)	7797.3			
ICC	28.8%	3148.9/(3148.9 +7797.3)=0.288		
Variance in Model 1	(3148.9-2307.5)/3148.9=0.267 Note: Random ANOVA provides the base in this application because it represents the total parameter variance in the school means that is potentially explainable by alternative level-2 models.			
Variance in Final Model	(7797.3 – 7060.3)/7797.3 = 0.09. Note: Random ANOVA provides the appropriate base in this application because it represents the total within-school variance that can be explained by any Level-1 and Level-2 data, respectively.			
PISA MATH				
Fixed Effect	*Coefficient*	*se*		
Average School Mean	371.4	3.9		
Random Effect	*Variance*	*d.f.*	X^2	*p-value*
Level 2 (school)	1177.8	202	2196.8	<0.001
Level 1 (student)	3839.4			
ICC	23.5%	1177.8/(1177.8 + 5017.2)		
Variance in Model 1	(1177.8 – 1009.3)/1177.8=0.14 Note: Random ANOVA provides the base in this application because it represents the total parameter variance in the school means that is potentially explainable by alternative level-2 models.			
Variance in Final Model	(3839.4-3380.5)/3839.4=0.12 (1177.8—753.2)/1177.8=0.36 Note: Random ANOVA provides the appropriate base in this application because it represents the total within-school variance that can be explained by any Level-1 and Level-2 data respectively.			
NAfKE ARABIC				
Fixed Effect	*Coefficient*	*se*		
Average School Mean	35.0	0.94		
Random Effect	Variance	*d.f.*	X^2	p-value
Level 2 (school)	68.5	113	371.5	<0.001
Level 1 (student)	280.1			
ICC	19.6%	68.5/(68.5 + 280.1)		
Variance in Model 1	(68.5-47.9)/68.5=0.30 Note: Random ANOVA provides the appropriate base in this application because it represents the total within-school variance that can be explained by any Level-2 data.			
Variance in Model 2 (Final)	(68.5-35.4)/68.5=0.48 Note: Random ANOVA provides the appropriate base in this application because it represents the total within-school variance that can be explained by any Level-1 and Level-2 data respectively.			