Commissioned Paper
January 2017

Examining Gender Differences in the Mathematical Literacy of 15-Year-Olds and the Numeracy Skills of the Age Cohorts as Adults

Authors:

Alka Arora

Emily Pawlowski

Suggested Citation: Arora, A. and Pawlowksi, E. (2017). *Examining Gender Differences in the Mathematical Literacy of 15-Year-Olds and the Numeracy Skills of the Age Cohorts as Adults.* Retrieved [insert date], from [insert website]. Washington, DC.

This project has been funded by the American Institutes for Research through a contract with the National Center for Education Statistics (NCES) of the U.S. Department of Education. This report is based on PIAAC data released in October 2013. The views expressed in this paper do not necessarily reflect the views or policies of the American Institutes for Research, National Center for Education Statistics, or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply their endorsement the U.S. Government.

AIR-PIAAC Contact:

Jaleh Soroui (AIR-PIAAC Director)
Saida Mamedova (Senior Research Analyst)
PIAACgateway.com
piaac@air.org

Author Contact:

American Institutes for Research. Washington, D.C.
Alka Arora at aarora@air.org;
and Emily Pawlowksi at epawlowski@air.org.

Examining Gender Differences in the Mathematical Literacy of 15-Year-Olds and the Numeracy Skills of the Age Cohorts as Adults

Executive Summary

Patterns of gender disparities in science, technology, engineering, and mathematics (STEM) fields are seen at various stages, from early education to secondary school through college and into the workforce. These disparities have often been documented in national and international large-scale school assessments and in labor force studies. This study uses data from the two assessments—mathematical literacy in the Program for International Student Assessment (PISA) and numeracy in Program for the International Assessment of Adult Competencies (PIAAC) —to look at the skills and characteristics of a group of 15-year-old students and their age cohort as 23-to-25-year-old adults. **Combining PISA and PIAAC allows one to see the progression of gender differences in mathematics skills from the 15-year-old students in PISA to the cohort of 23-25 years old young adults in PIAAC**.

- In general, there is a fairly close correlation between countries' mathematics performance in PISA 2003 and in numeracy in PIAAC 2012, when looking at the relevant age cohort in PIAAC (23- to 25-year-olds).
- The gender gap in mathematics performance of the cohort of 15-year-olds in PISA 2003 either stayed the same in PIAAC 2012 (when those in the cohort were 23 to 25 years old) or increased. Approximately half of the countries showed an increase in the gender gap, with Finland and United States showing the largest increase.
- Within the total PIAAC population, the size of the gender gap in numeracy increases as age increases. The 16 to 24 age group shows the least number of significant differences between males and females within countries.
- In most countries that participated in PISA 2003, male students were more engaged in and had more positive attitudes toward learning mathematics than females, although most of these gender differences were small.
- In all but one country, more females than males ages 23-25 had completed a university degree. However, many more males than females earned a degree in the STEM-related areas of science, engineering, mathematics, and computing. More females than males choose non-STEM areas and the females who did choose STEM areas more often chose the areas of education sciences or health and welfare.
- Female adults in 10 out of 16 countries in the study used their numeracy skills at home less frequently than males did. Females in 8 of the 16 countries in the sample used their numeracy skills at work less often than males; the Netherlands had the highest gender difference in adults' use of numeracy skill at work.
- In most countries, there was no gender difference in adults' readiness to learn new ideas and information. The United States and Japan were the only two countries in which females showed slightly less readiness to learn new ideas or information than their male counterparts.

These findings suggest that there is a still a long way to go toward gender equity in the STEM fields. Educators at various levels need to understand these differences and work with their female students to improve their attitudes and engagement with STEM fields. It is also important for colleges and universities to create resources and policies to encourage female students to choose and complete their major area of study in the STEM fields.

Section I: Introduction

The Importance of STEM Fields

Science, technology, engineering, and mathematics (STEM) fields are a source of innovation and are critical to the national economy, and identifying and developing talent in these fields is seen as vital to creating new jobs, improving our quality of life, and maintaining our position as a global leader (National Science Board 2010). On an individual level, having an understanding of mathematics is necessary for life in modern society. A growing number of problems and situations encountered in daily life require some level of understanding of mathematical reasoning and tools, as people from all walks of life are now presented with an increasing amount and a wider range of information of a quantitative nature in a range of contexts. The importance of mathematics skills in the workplace is also increasing, and numeracy-related skills have been shown to be a key factor in labor market participation. Strong mathematics skills are also needed for certain career fields and for post-secondary education in some areas, such as engineering or medicine. For example, mathematics is essential in daily life for such activities as cooking, financial planning, building things, and examining health risk factors (OECD 2013c; PIAAC Numeracy Expert Group 2009).

Gender Disparities in STEM

An important step in developing talent in STEM fields is to develop skills and potential across all demographic groups. The abilities and talents of students traditionally underrepresented in STEM may go unrecognized and undeveloped and these individuals may face barriers to achievement. Because these students are never identified or given an opportunity to realize their potential, they constitute a considerable source of untapped talent in STEM. (National Science Board 2010). One of these underrepresented groups that has untapped talent and potential in the STEM fields are women. Additionally, attracting and retaining more women in STEM fields can lead to increased innovation, creativity, and competitiveness. Women will bring different and diverse perspectives which, for example, can lead to better designed products, services, and solutions that are more likely to represent all users (Hill et al. 2010).

Research using NAEP data, from 2012, show that there were no significant gender gaps in mathematics achievement at ages 9 or 14 in the U.S. However, a significant difference in mathematics scores was found at age 17, with male students scoring higher than female students (National Center for Education Statistics 2013). Gender differences in academic course-taking at the high school level in the U.S. have also been found, with data from the 2009 National Assessment of Educational Progress (NAEP) High School Transcript Study (HSTS) indicating that more males than females earned credits in physics, engineering, and computer/information science (Cunningham, Hoyer, and Sparks 2015).

Looking at post-secondary education, although women earn the majority of new bachelor's degrees, they earn a much smaller percentage of degrees in STEM fields, such as engineering or physics, than men in the U.S. (Snyder and Dillow 2012). Additionally, higher percentages of female than male students in the U.S. who pursue associate's and bachelor's degrees switch from majors in STEM fields to majors in non-STEM fields (Chen and Ho 2012). Patterns of gender disparities in STEM fields are also seen at the graduate level in the U.S., where females earn a smaller percentage of doctorates in STEM fields than males do (Hill et al. 2010). Females in the workforce also tend to use their numeracy skills at work less than their male counterparts do, as measured by an index constructed from responses to questions asking about frequency of mathematical tasks of varying levels of complexity, from calculating budgets to using advanced statistics (Lindemann 2015). Women are vastly underrepresented in many STEM professions, particularly in engineering and physics (Hill et al. 2010). This gender disparity is problematic because the underrepresentation of women in STEM occupations contributes to the earnings gap between men and women, as those working in STEM fields tend to have higher earnings (Beede et al. 2011; Hill et al. 2010; Shauman 2006).

Attitudes Towards STEM Fields

Positive attitudes towards mathematics and science have been found to be associated with higher mathematics and science achievement (Mullis et al. 2012). Gender differences in attitudes and motivations towards mathematics may change as students progress in their schooling, with females forming less positive attitudes than males as they continue in their schooling. For example, internationally, more fourth-grade females than males felt it was important to do well in mathematics and science. Among eighth-grade students, few gender differences were seen in this attitude. Meanwhile, looking at students in their final year of secondary school, more males than females reported that it was important to do well in mathematics and science (Mullis et al. 2000). Data from the 2009 NAEP grade 12 student questionnaires in mathematics and science also indicate that U.S. males and females have different attitudes toward the STEM fields, with fewer females than males reporting that they like mathematics or science or that mathematics or science is one of their favorite subjects (Cunningham, Hoyer, and Sparks 2015). Previous reports have suggested that females' lack of self-confidence in their mathematics and science ability and their anxiety toward mathematics contribute to their underrepresentation among high performers in mathematics and science. This may be because individuals with more self-confidence are more likely to allow themselves the freedom to engage in the trial-and-error processes that are essential to building these types of knowledge (OECD 2015). These results indicate that one of the factors contributing to gender disparities in STEM fields may be their attitudes toward mathematics.

The metaphor of a "leaky pipeline" carrying students from their early education to secondary school through college and into the workforce is used to explain the pattern of gender disparities and under-representation of women in STEM careers (Blickenstaff 2005). Individuals leave the pipeline at various places, from when they choose a major to when they select a career outside of a STEM field. The research described above shows that more women than men leave this STEM pipeline at many different points, as women leave or participate less in the STEM fields at different stages in their education and careers, from high school to college to graduate school to the workforce. The attitudes of males and females towards mathematics is also shown to differ as they progress through their education and into the workforce and this may also contribute to more females than males leaving STEM fields

The current research looks more closely at the skill levels of males and females; specifically, at whether there is a gender gap in the mathematics skill of 15-year-olds and adults and whether this gap increases among this age cohort over time, from PISA 2003 to PIAAC 2012, as students complete their education and enter the workforce. The research will look at this pattern across countries and will also examine whether some of the disparities described above in attitudes, areas of study, and employment characteristics are also observed among these cohorts. Data from the 2003 Program for International Student Assessment (PISA) and the 2012 Program for the International Assessment of Adult Competencies (PIAAC), both of which are large-scale assessments of key competencies, are used to look at the skills and characteristics of a group of 15-year-old students and their age cohort as 23-to-25-year-old adults.

Section II: Research Questions

Given the importance of STEM fields and the pattern of gender differences in performance and outcomes in STEM seen in the research described above, and based on the similarities between the content and cohorts assessed in PISA and PIAAC (which are described in more detail below), this study uses data from the two assessments—mathematical literacy in PISA and numeracy in PIAAC—to explore the following research questions:

1. How similar or different are the performance of countries in PISA and PIAAC?
2. Are there any significant differences in performance by gender across the participating countries in the two assessments?
3. How are these gender differences in performance different among students in PISA and adults in PIAAC?
4. Are there any gender differences in attitude and engagement toward learning mathematics among students in PISA?
5. Are there any gender differences in highest level of education and major area of study among adults in PIAAC?
6. Are there any gender differences in usage of numeracy skills at home and at work among adults in PIAAC?
7. Are there any significant gender differences in readiness to learn new ideas/information among adults in PIAAC?

Section III: Data and Methodology

PISA and PIAAC

PISA

PISA is a large-scale, cyclical, direct assessment conducted in schools; it is coordinated by the Organization for Economic Cooperation and Development (OECD).The direct assessment in PISA includes measures of reading, mathematics, and science literacy. It was first conducted in 2000 and has been conducted every 3 years, with the major domain of study rotating between mathematics, science, and reading in each cycle. In 2003, the major domain of study was mathematics, and 41 OECD and partner countries participated. Each country administered the PISA assessment to a sample of about 4,000 to 5,000 students who are 15 years old.[1]

PIAAC

PIAAC is a large-scale, cyclical, direct assessment conducted in households; like PISA, it is coordinated by the OECD. The direct assessment in PIAAC includes measures of literacy, numeracy, and problem solving in technology-rich environments. PIAAC was first conducted in 2011–12 and involved 24 OECD and partner countries. Each country administered PIAAC to a sample of about 5,000 adults aged 16–65.[2]

Comparing PISA and PIAAC

Comparing PISA Mathematics and PIAAC Numeracy

Both PISA and PIAAC assess literacy, numeracy/mathematics, and problem solving. The cycles of PISA conducted between 2000 and 2009 included students that were part of the cohorts that participated in PIAAC 2012. Looking at these cycles, reading literacy, mathematical literacy, and scientific literacy are part of every PISA cycle, but the problem-solving assessment was included in PISA only in 2003. Although reading literacy in PISA and literacy in PIAAC may also be comparable and have available data for some cohorts for both assessments, due to the importance of STEM in the workforce and economy, this study will focus on the domain of numeracy/mathematics.

PISA defines mathematical literacy as

> *"the capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgements and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen."*

PIAAC defines numeracy as

> *"the ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life."*

[1] PISA samples students between the ages of 15 years 3 months and 16 years 2 months enrolled in grade 7 or above, regardless of the type of institution attend and whether they attend full time or part time.
[2] The PIAAC sample includes noninstitutionalized adults ages 16–65 residing in the country regardless of citizenship, nationality, or language.

Table 1. Comparison of PISA and PIAAC: numeracy

	PISA Mathematics	**PIAAC Numeracy**
Definition	The capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgements and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen.	The ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life.
Content	Quantity Space and shape Change and relationships	Quantity and number Dimension and shape Pattern, relationships, change Data and chance
Cognitive processes	Reproduction (simple mathematical operations) Connections (bringing together ideas to solve straightforward problems) Reflection (wider mathematical thinking)	Identify, locate or access Act upon and use (order, count, estimate, compute, measure, model) Interpret, evaluate and analyse Communicate
Contexts	Personal Educational and occupational Public Scientific	Everyday life Work-related Community and society Education and training

Source: The Survey of Adult Skills: Reader's Companion

As shown in table 1, both PISA and PIAAC cover similar content areas, such as "quantity" in PISA and "quantity and number" in PIAAC. The contexts (the settings, situations, or circumstances in which the items take place) covered in PISA and PIAAC, such as "public" in PISA and "community and society," also show significant overlap. The cognitive processes used by respondents are also similar in the two assessments, although they do not match up as exactly as do the content and context classifications (OECD 2013b). A more detailed comparison of PISA and PIAAC, and the cognitive domains they assess, can be found in chapter 6 of *The Survey of Adult Skills: Reader's Companion* (OECD 2013b).

Population

Table 2. Age of PISA cohorts in 2011-12 and focus of PISA assessment

PISA Cohort	Main focus of PISA	Age of PISA cohort at time of PIAAC 2012
PISA 2000	Reading	26-27
PISA 2003	Mathematics	23-24
PISA 2006	Science	20-21
PISA 2009	Reading	17-18

As shown in table 2, several PISA cohorts are included in the population assessed in PIAAC in 2011-12, as respondents to PIAAC ages 17-27 were part of cohorts that participated in PISA. However, there are differences in the coverage of these cohorts in PISA and PIAAC. PISA tested 15-year-olds enrolled in educational institutions at grade 7 or above, but not those who were not enrolled in an educational institution or who were enrolled in a grade below grade 7. In contrast, the target population for PIAAC is the entire noninstitutionalized, resident population. Additionally, both emigration and immigration would have changed the composition of each PISA cohort as they have aged between administrations of PISA and PIAAC. For example, those who migrated between the ages of 15 and when PIAAC was administered in 2011-12 could be part of the PISA sample of their native country but could be part of the PIAAC sample of the country they immigrated to.

Comparing results from PISA and PIAAC can provide insight into questions of skills development over time, since there are similarities in how skills are conceptualized and defined in the two assessments. Both assessments have complementary goals, with PISA seeking "to identify ways in which students can learn better, teachers can teach better, and schools can operate more effectively" and PIAAC focusing "on how adults develop their skills, how they use those skills, and what benefits they gain from using them" (OECD 2013a). Previous research has found a similar pattern of performance across countries in literacy and numeracy in PISA and PIAAC (Lundetræ 2014; OECD 2013a).

However, a psychometric link between PISA and PIAAC has not been established, and the two assessments do not share any common items, so caution must be exercised when comparing their results. Although similar cohorts are assessed in PISA and PIAAC, the overlap between the target populations of the age cohorts in the two assessments is not complete (i.e. not all adults included in the PIAAC sample would have been included in the PISA sample as 15-year-olds). Additionally, while there is a close relationship between the concepts of numeracy in PIAAC and mathematical literacy in PISA, the measurement scales are not the same. Therefore, this study takes special care in comparing results across these two assessments.

This study uses the cohort that participated in 2003 PISA, in which mathematics was the focus of the assessment. This year was selected because it provides the most valid and extensive PISA mathematical literacy data on a cohort that also took part in PIAAC and provides additional background information related to mathematics, including information on the students' attitudes towards mathematics. A 3-year age band is used in the analysis of the PIAAC cohort to increase the sample size and therefore the reliability of estimates. Therefore, the study primarily focuses on 23 to 25-year-olds in the PIAAC 2012 cohort. Additionally, using this age cohort allows this analysis to look at a key period in the "leaky pipeline" as the cohort completes their education and enters the workforce. However, for one specific analysis, the entire PIAAC population (16- to 65-year-olds) is included in order to look more broadly at patterns of performance across the lifespan and the progression of gender difference over age.

Countries selected

Countries included in this study were selected on the basis of comparability and public availability of data from both PISA 2003 and PIAAC 2012.[3] Of the 24 countries that participated in PIAAC 2012, Cyprus

---

[3] Researchers had access to the restricted-use data file for the United States and the data file for Australia in which a more detailed age variable was available.

and Estonia were the only two countries that did not participate in PISA 2003, so their data were not included. The response rate for the United Kingdom was too low to ensure comparability for PISA 2003, so its results were not included. Some countries participated in both assessments, but did not assess comparable populations. For instance, Belgium assessed only its Flemish population in PIAAC 2012 and the Russian sample did not include adults from the Moscow municipal region in PIAAC 2012, while the entire populations of both countries were included in the sample for PISA 2003. As a result, data from these countries were not included. Finally, Austria, Canada, and Germany were not included in the analyses due to limitations of the age variables in the publicly available PIAAC data. In total, data from 16 countries were included.

Analysis

The OECD's PIAAC International Data Explorer (IDE) and the National Center for Education Statistics' (NCES's) PISA IDE, both of which are online data tools, were used in combination with the IEA's International Association for the Evaluation of Educational Achievement's (IEA) International Database Analyzer (IDB Analyzer), SPSS software, and the PIAAC SPSS micro-data files for data analysis. These tools were used to conduct analyses taking into account the plausible values and sampling weights. Additionally, where noted, data on the characteristics of PISA students from previous PISA publications were included in this paper.

Analyses were conducted to examine the overall average mathematics scores in PISA 2003 and the average numeracy scores of the cohort in PIAAC to answer Research Question 1. Then, the gender differences in these scores were examined to answer Research Question 2. Because proficiency scores on PISA[4] and PIAAC[5] are reported on different scales, the results from the two surveys could not be compared directly. Thus, effect sizes (Cohen's *d*[6]) for each assessment were calculated to compare gender differences while accounting for the different scales of the assessments to answer Research Question 3. This measure of effect size expresses the size of the gender difference as the share of a standard deviation; an effect size of 0.2 is considered to be small, an effect size of 0.5 is considered to be moderate, and an effect size of 0.8 is considered to be large (Cohen 1992). All effect sizes reported in this study use the gender difference of the male average minus the female average, with a "positive gender effect" referring to a gender difference favoring males and a "negative gender effect" referring to a gender difference favoring females. Also examining Research Question 2, the gender differences in average numeracy scores were examined in the total PIAAC population (16-65) by 10-year age band.

This study also includes information on various characteristics of the PIAAC cohort by gender, including educational attainment and area of study, which are reported as percentage distributions to answer Research Question 5. Gender differences in adult skill use and learner characteristics, as measured by PIAAC indices, were also analyzed to answer Research Questions 6 and 7. Effect sizes were again used to

---

[4] PISA proficiency results are represented on a scale that was constructed to have an OECD average in mathematics of 500 score points, with a standard deviation of 100 score points.
[5] PIAAC proficiency results are represented on a 500-point scale. The OECD average in numeracy is 269 score points, with a standard deviation of 52.6 score points.
[6] Cohen's d = (Mean1 − Mean2)/Pooled Standard Deviation.

look at gender differences across these various indices, so there would be comparable measures across the various indices and scales.

Section IV: Results

Overall mathematics performance of countries in PISA 2003 and PIAAC 2012

This section will address the first research question. Table 3 presents the overall average scores for the 16 countries that participated in PISA 2003 and PIAAC 2012. In general, countries performed very similarly in both assessments. Relatively low performing countries in both assessments include Italy, Poland, Spain, and the United States, whereas relatively high performing countries include the Czech Republic, Denmark, Finland, Japan, the Netherlands, and Sweden.

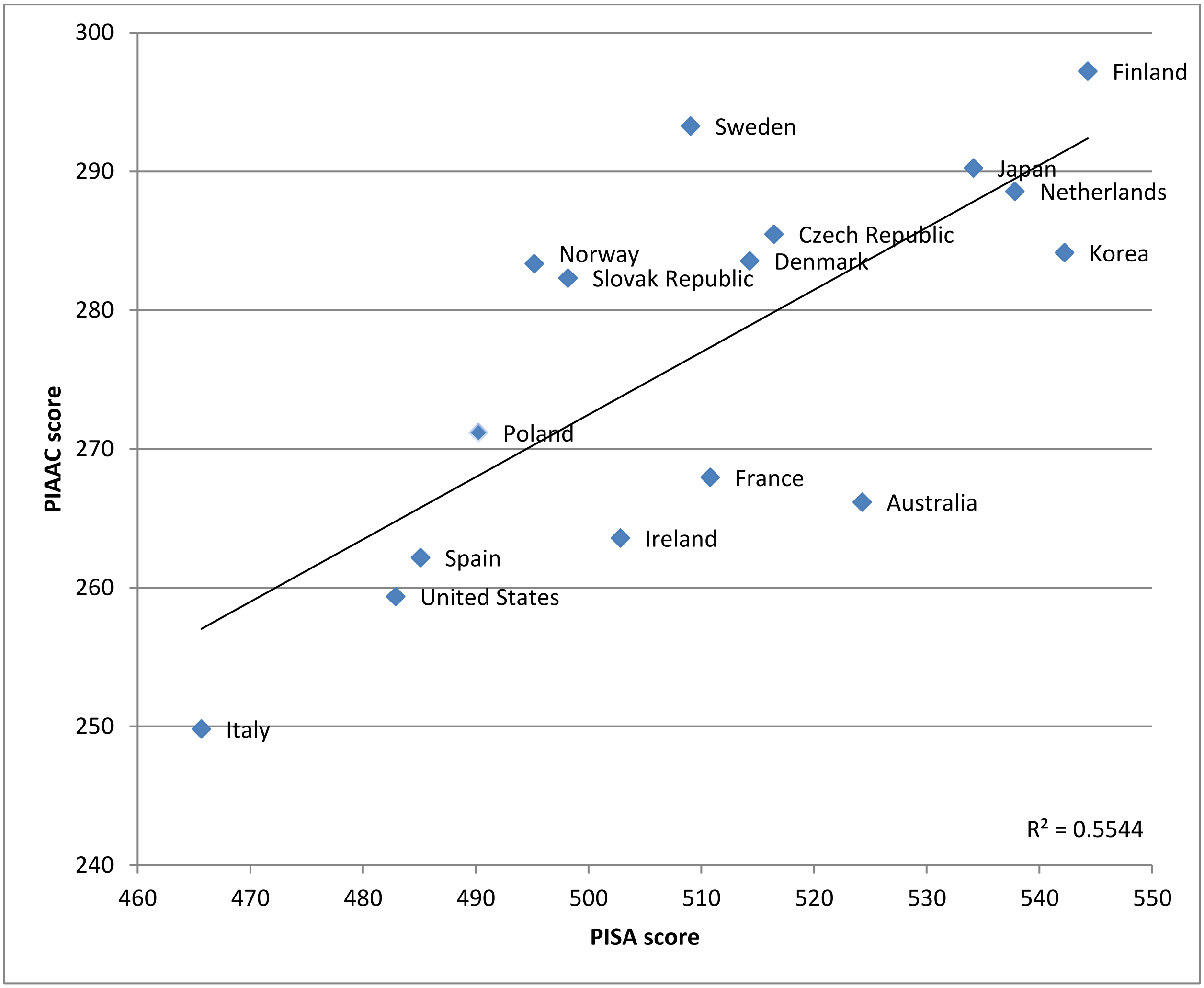
Table 2 . Average mathematics scores of 15-year-olds in PISA 2003 and numeracy scores of 23- to 25-year-olds in PIAAC 2012, by country

	Average Scores			
Country	PISA		PIAAC	
Australia	524	(2.1)	266	(3.7)
Czech Republic	516	(3.5)	285	(3.3)
Denmark	514	(2.7)	284	(3.8)
Finland	544	(1.9)	297	(3.4)
France	511	(2.5)	268	(2.9)
Ireland	503	(2.4)	264	(4.0)
Italy	466	(3.1)	250	(4.6)
Japan	534	(4.0)	290	(3.5)
Republic of Korea	542	(3.2)	284	(2.7)
Netherlands	538	(3.1)	289	(3.1)
Norway	495	(2.4)	283	(3.6)
Poland	490	(2.5)	271	(1.4)
Slovak Republic	498	(3.3)	282	(3.8)
Spain	485	(2.4)	262	(3.4)
Sweden	509	(2.6)	293	(2.9)
United States	483	(2.9)	259	(3.2)

Figure 1 displays a scatterplot of the relationship between average mathematics scores in PISA 2003 and numeracy scores in PIAAC 2012 and also shows best fit line. Figure 1 confirms that there is a fairly close correlation (R^2 = 0.55) between countries' mathematics performance in PISA 2003 and in numeracy in PIAAC 2012, when looking at the relevant age cohort in PIAAC (23- to 25-year-olds). However, there are some outliers. For example, the cohort in Australia performed relatively lower in PIAAC it did in PISA 2003. In contrast, the cohort in Sweden and Norway and the Slovak Republic did relatively better in PIAAC than in PISA 2003. Norway is actually one of the countries that performed below average in PISA

2003 and above average in PIAAC 2012, while Australia performed above average in PISA 2003, but at the OECD average in PIAAC 2012.

Figure 1. Average mathematics scores of 15-year-olds in PISA 2003 and numeracy scores of 23- to 25-year-olds in PIAAC 2012, by country



Gender gap in mathematics achievement in PISA and PIAAC

This section will address the second and third research question. Table 4 presents the gender gap in mathematics achievement in terms of average scale score difference as well as in terms of Cohen's *d*. Gender differences are shown for PISA 2003 and for the same cohort in PIAAC 2012. A table showing the number of males and females that are included in the samples of both assessments and the average mathematics scores in PISA and numeracy scores in PIAAC by gender is included in Appendix A.

Statistically significant results can be found despite small effect sizes (and little practical significance) if there is a large enough sample size; while large effect sizes do not guarantee statistical significance when small sample sizes are used (Lantz 2013; Sullivan and Feinn 2012). Therefore, because of the larger sample size in PISA, the chances of gender differences being statistically significant in PISA is greater than in PIAAC when looking at 23- to 25-year olds, and a greater number of findings of statistical significance in PISA may purely be the artifact of the larger sample size. Therefore, as discussed in the analysis section, the current study uses effect size to compare gender differences. In general, the results confirm that males performed better than females in both PISA and PIAAC. However, the size of the gender effect is miniscule in PISA 2003, with only the Republic of Korea showing a positive gender effect on mathematical literacy of 0.25 (which is equal to ¼ of a standard deviation).

Several countries showed an increased gender effect on numeracy from PISA to PIAAC. Finland and United States are the only two countries which had a medium size gender effect on numeracy performance in PIAAC, with the effect size being more than half of the standard deviation in PIAAC. France, the Netherlands, Norway, and Spain also showed a small gender effect in PIAAC, in the range of 0.2 to 0.5.

Table 4. Gender differences in mathematics scores of 15-year-olds in PISA 2003 and numeracy scores of 23- to 25-year-olds in PIAAC 2012, by country

Country	PISA 2003 (age 15)			PIAAC (ages 23-25)		
	Difference in mathematics scale score (males-females)	Standard deviation of total sample	Effect size (Cohen's *d*)	Difference in numeracy scale score (males-females)	Standard deviation of total sample	Effect size (Cohen's *d*)
Australia	5.34	95.42	0.06	3.13	51.66	0.06
Czech Republic	14.97*	95.94	0.16	4.93	40.16	0.12
Denmark	16.58*	91.32	0.18	7.88	50.97	0.15
Finland	7.41*	83.68	0.09	27.77*	49.82	0.56
France	8.51*	91.7	0.09	10.79*	50.51	0.21
Ireland	14.81*	85.26	0.17	1.66	47.62	0.03
Italy	17.83*	95.69	0.19	-13.07	47.62	-0.27
Japan	8.42	100.54	0.08	8.22	42.69	0.19
Rep. of Korea	23.41*	92.38	0.25	11.96*	38.93	0.31
Netherlands	5.12	92.52	0.06	16.56*	47.39	0.35
Norway	6.22	92.04	0.07	12.88	54.27	0.24
Poland	5.59	90.24	0.06	2.88	47.83	0.06
Slovak Republic	18.66*	93.31	0.20	8.16	47.99	0.17
Spain	8.86*	88.47	0.10	9.51	47.35	0.20
Sweden	6.53*	94.75	0.07	3.76	49.83	0.08
United States	6.25*	95.25	0.07	27.43*	48.36	0.57

No gender effect (Cohen's *d* ≤.2)

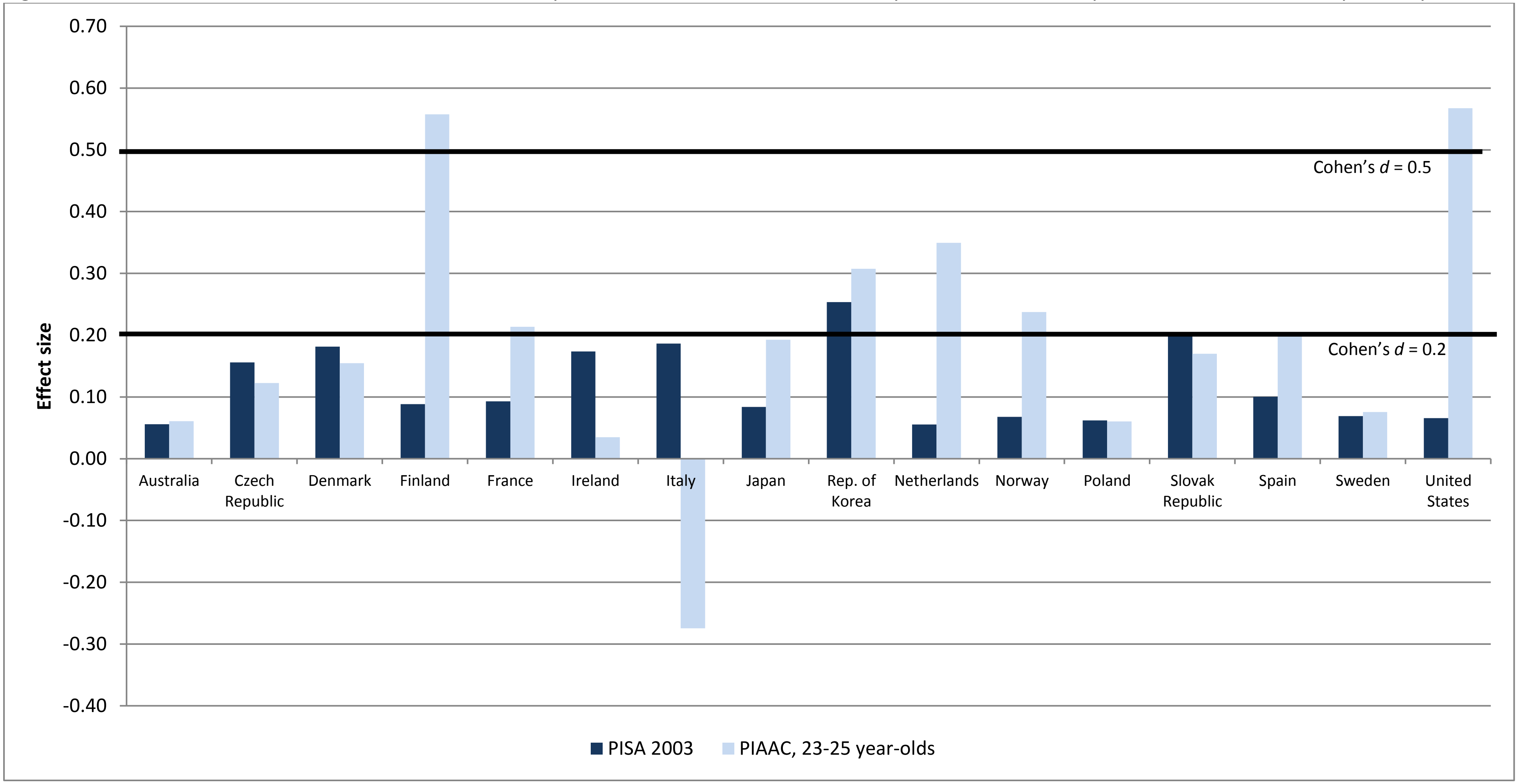
Small gender effect (.2 ≤ Cohen's d ≤ .5)

Medium gender effect (Cohen's *d* ≥. 5)

* Male scores are statistically significantly different (p < .05) than Female scores.

Figure 2 presents a graphical display of the information shown in table 3. The data show that in most of the countries included in this study, the gender effect in PIAAC 2012 either stayed the same or became larger than the gender effect in PISA 2003, with Finland and the United States showing the largest change between the two assessments in favor of males and Italy showing the largest change in favor of females.

Figure 2. Gender differences in mathematics scores of 15-year-olds in PISA 2003 and numeracy scores of 23- to 25-year-olds in PIAAC 2012, by country



In addition to looking at gender differences among the cohorts that participated in PISA, data from PIAAC also allows one to look at the gender differences in performance across the lifespan. Table 5 presents the gender differences in PIAAC 2012 within different age bands. The total PIAAC population (ages 16 to 65) is divided into five 10-year age bands. Looking at the gender differences in performance across these different age bands allows one to see the overall pattern by age (i.e. whether there is a larger gender difference in performance among older or younger adults) and also allows one to see if the pattern of gender differences seen in 23- to 25-year-olds is consistently found among the other age bands. In general, the size of the gender gap in numeracy increases as age increases. The first column, which displays gender differences in the 16 to 24 age group, shows the least number of significant differences between males and females within countries; Finland has the largest gender gap in this age band. The Slovak Republic has only one gender gap that is significant in the 35 to 44 age band. Poland is the only country with no significant gender difference in any age group.

Table 5. Gender differences in numeracy scores in PIAAC 2012, by 10-year age band and country

	Differences in numeracy scores, by gender									
Country	16- to 24-year-olds		25- to 34-year-olds		35- to 44-year-olds		45- to 54-year-olds		55- to 65-year-olds	
Australia	5.79		9.40	*	14.00	*	19.56	*	19.59	*
Czech Republic	5.77		5.59		13.36	*	14.11	*	3.13	
Denmark	2.09		12.37	*	15.01	*	8.97	*	11.67	*
Finland	12.81	*	4.89		7.59		11.70	*	10.49	*
France	9.74	*	11.78	*	9.12	*	10.92	*	11.07	*
Ireland	10.83	*	10.78	*	15.00	*	10.47	*	15.09	*
Italy	-0.96		10.91	*	12.03	*	12.25	*	12.24	*
Japan	8.59	*	8.64	*	15.98	*	13.41	*	14.52	*
Republic of Korea	3.55		5.54	*	9.96	*	11.41	*	21.17	*
Netherlands	8.82	*	12.58	*	21.19	*	18.58	*	19.63	*
Norway	9.51	*	9.03	*	19.32	*	17.55	*	17.43	*
Poland	0.19		4.86		5.82		-3.28		-1.16	
Slovak Republic	1.23		2.41		7.65	*	1.17		-2.08	
Spain	6.49	*	10.40	*	12.99	*	10.95	*	16.42	*
Sweden	8.90	*	16.32	*	12.25	*	11.02	*	18.37	*
United States	8.23		14.93	*	16.27	*	12.54	*	18.72	*

* Male scores are statistically significantly different ($p < .05$) than Female scores.

Gender differences in students' attitude and engagement toward learning mathematics

This section will address the fourth research question. PISA contains several indices[7] of students' attitudes towards mathematics that are constructed based on their responses to a series of related questions. These indices provide information on students' motivations, interests, and beliefs, which can

[7]PISA indices were constructed using IRT scaling with a mean of zero and a standard deviation of 1. Each item used in calculating the indices used a 4-point Likert scale.

influence outcomes and performance. Looking at these indices can indicate the types of gender differences that can be seen in 15-year-old students.

These indices are important indicators of future learning and skill development in mathematics. For instance, students who have high levels of interest and low levels of anxiety related to mathematics are more likely to develop the skills necessary to learn STEM subjects effectively (and not avoid mathematics and potentially miss educational and career opportunities) (OECD 2004).

The PISA indices that were examined in this study are

- *instrumental motivation in mathematics* (to what extent students are encouraged to learn by external rewards, such as good job prospects);
- *interest in and enjoyment of mathematics* (students' interest in mathematics as a subject as well as their enjoyment of learning mathematic);
- *anxiety in mathematics* (to what extent students feel helpless and under emotional stress when dealing with mathematics);
- *self-efficacy in mathematics* (to what extent students believe in their own ability to handle learning situations in mathematics effectively and to overcome difficulties); and
- *self-concept in mathematics* (students' belief in their own mathematical competence).

Additional information about the questions included in the construction of these indices can be found in appendix B.

Table 6 presents gender differences in these attitudinal indices as measured in terms of effect sizes, with a positive gender effect indicating that males have higher values and a negative gender effect indicating that females have higher values. Most countries show a small to medium size gender effect on the indices. All of the differences show a positive gender effect (with the exception of the index of anxiety in mathematics, where there is a negative gender effect).

On the Instrumental motivation and anxiety indices, 13 of the 16 countries show a small positive gender effect favoring males. On the self-efficacy index, 14 countries show a positive gender effect and in fact, Finland and the Netherlands show a medium-size gender effect. Fourteen of the 16 countries show a positive gender effect for the self-concept index, with the effect being larger in the Netherlands. In contrast, the data from only 8 countries— Australia, the Czech Republic, Denmark, Finland, France, Japan, the Netherlands, and Norway—show that males are more interested in and enjoy learning mathematics more. In the remaining 8 countries, females have a similar interest and enjoyment in learning mathematics.

A few countries show interesting patterns in these attitudinal indices. Poland is the only country with no significant gender effect in any of the five indices. Finland and the Netherlands are the only countries that have medium-size gender effects (Finland for the self-efficacy index and the Netherlands for both the self-efficacy index and the self-concept index). In the United States, a gender effect was found for only two of the five indices, and the effect sizes are small. U.S. females are more anxious and have less confidence while doing mathematics than their male counterparts. In Italy too, the gender effect is small and was seen for only two of the five indices (self-efficacy and instrumental motivation).

Table 6. Gender differences in learner characteristics of 15-year-olds in PISA 2003, by country

Country	Gender differences (measured in terms of effect size)				
	Instrumental motivation in mathematics	Interest in and enjoyment of mathematics	Anxiety in mathematics	Self-efficacy in mathematics	Self-concept in mathematics
Australia	0.24	0.23	-0.31	0.37	0.34
Czech Republic	0.26	0.26	–0.26	0.42	0.36
Denmark	0.43	0.29	–0.38	0.45	0.48
Finland	0.36	0.34	-0.39	0.56	0.45
France	0.35	0.24	–0.39	0.31	0.37
Ireland	0.32	0.04	–0.28	0.30	0.23
Italy	0.23	0.11	-0.17	0.36	0.14
Japan	0.31	0.26	-0.26	0.31	0.36
Rep. of Korea	0.20	0.16	–0.14	0.20	0.26
Netherlands	0.50	0.34	–0.38	0.59	0.55
Norway	0.23	0.25	–0.36	0.37	0.42
Poland	0.05	0.11	–0.03	0.17	0.18
Slovak Republic	0.23	0.17	–0.25	0.33	0.30
Spain	0.09	0.03	–0.34	0.28	0.25
Sweden	0.32	0.19	–0.30	0.27	0.35
United States	0.10	0.16	-0.23	0.19	0.27

Source: OECD (2004), *Learning for Tomorrow's World: First Results From PISA 2003.*

No gender effect (Cohen's *d* ≤ .2)

Small gender effect (.2 ≤ Cohen's d ≤ .5)

Medium gender effect (Cohen's *d* ≥ .5)

Gender differences in adults' highest level of education and their major area of study

This section will address the fifth research question. Figure 3 presents the percentage of 23- to 25-year-old adults who attained a university degree by gender, as reported in PIAAC 2012. The figure shows that in all countries except Japan, an equal or higher percentage of females than males attained a university degree. Table 7 and 8 show the areas of study in which these adults attained their education, with table 7 displaying STEM fields. The table shows that females were underrepresented in the area of "engineering, manufacturing, and construction" in all countries and underrepresented in all countries except Australia and Finland in the area of "science, mathematics, and computing." In contrast, higher percentages of females than males attained their education in "health and welfare" and in "education training and education science."

Figure 3. Percentage of 23- to 25-year-olds in PIAAC 2012 attaining a university degree, by country

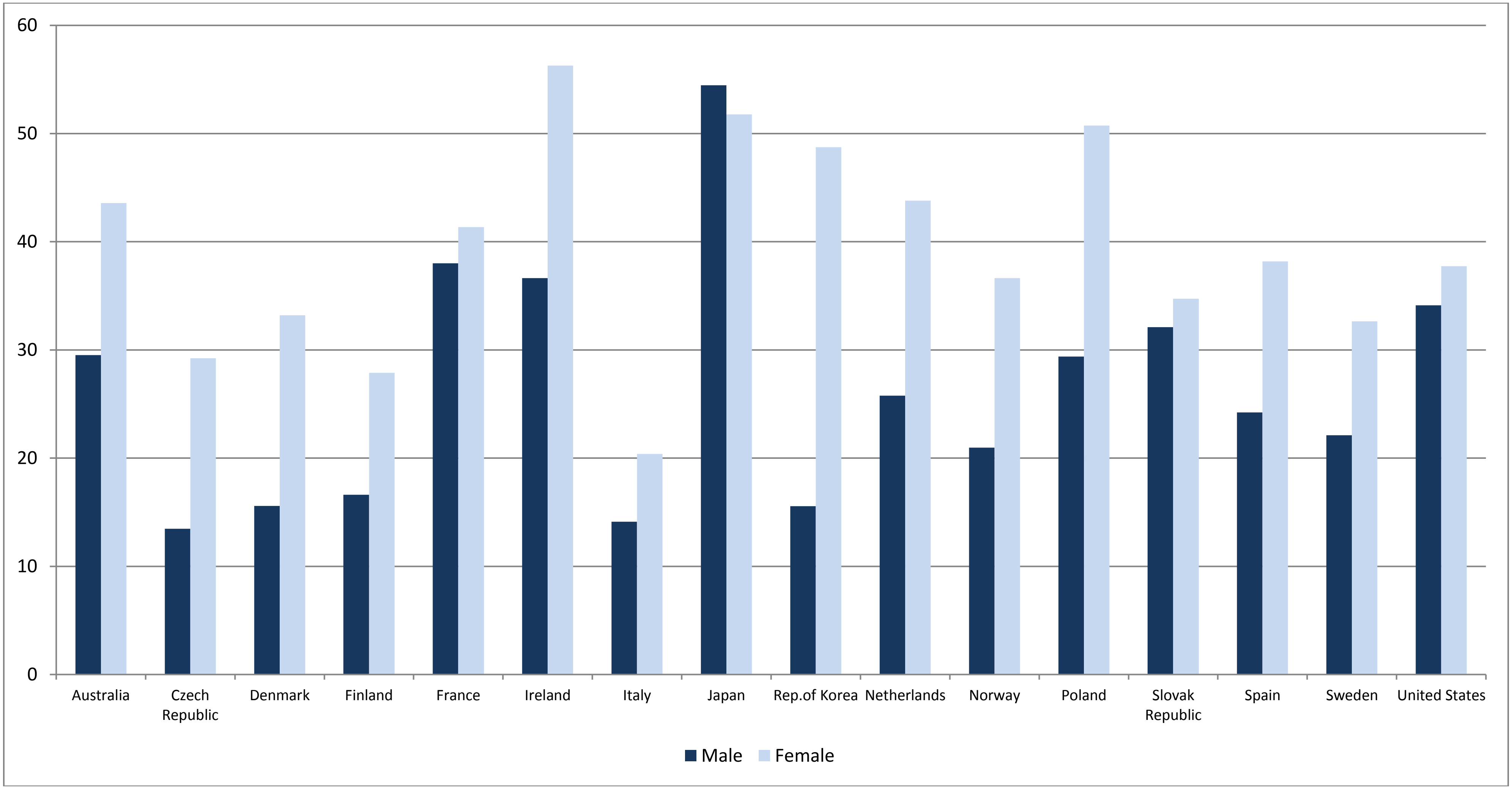


Table 7. Percentage distribution of area of study of 23- to 25-year-olds in PIAAC 2012, in STEM areas by gender and country

Country	Teacher training and education science		Science, mathematics and computing		Engineering, manufacturing and construction		Agriculture and veterinary		Health and welfare	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Australia	#	6	3	3	29	2	1	#	3	15
Czech Republic	#	6	5	2	49	7	3	9	1	5
Denmark	1	2	16	3	22	3	3	4	1	11
Finland	1	3	2	2	32	6	1	1	1	10
France	1	2	11	4	33	4	6	4	2	16
Ireland	3	5	10	8	13	2	1	2	3	11
Italy	1	3	22	19	10	1	#	#	2	6
Japan	6	8	7	4	21	4	1	2	2	10
Rep. of Korea	#	7	10	13	16	9	1	#	1	4
Netherlands	1	11	9	5	20	2	3	4	8	25
Norway	1	7	10	5	36	9	3	2	2	24
Poland	3	10	10	9	27	7	3	1	2	8
Slovak Republic	3	6	18	4	31	6	3	3	#	6
Spain	4	9	12	11	20	2	1	#	2	13
Sweden	2	5	7	6	34	9	2	2	5	14
United States	4	14	13	6	9	#	2	1	4	19

Rounds to zero.

Table 8. Percentage distribution of area of study of 23- to 25-year-olds in PIAAC 2012, in non-STEM areas by gender and country

Country	General programmes		Humanities, languages and arts		Social sciences, business and law		Services		Did not report due to low level of educational attainment	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Australia	25	19	5	7	15	23	5	11	13	14
Czech Republic	7	10	3	5	12	36	11	9	10	11
Denmark	17	27	2	9	13	22	3	9	24	11
Finland	40	33	3	5	6	15	2	15	11	10
France	8	8	3	6	14	27	12	16	12	13
Ireland	4	5	7	14	13	22	5	9	43	22
Italy	3	3	10	29	10	18	9	5	34	16
Japan	32	27	5	7	13	15	4	11	9	12
Rep. of Korea	59	29	3	15	5	14	4	7	1	2
Netherlands	9	4	1	5	27	24	3	4	19	15
Norway	17	19	7	7	11	11	4	5	9	12
Poland	19	15	5	10	10	22	9	14	12	3
Slovak Republic	8	19	3	9	14	20	13	17	9	10
Spain	5	6	1	9	10	18	1	1	44	31
Sweden	17	14	7	16	15	22	8	7	4	6
United States	11	4	8	10	20	17	4	8	25	21

Gender differences in adults' use of skills at home and at work

This section will address the sixth and seventh research question. Based on questions in the background questionnaire, PIAAC developed several skill use indices[8] about the frequency with which respondents perform specific tasks in their everyday life at home or at work (OECD 2013b). The use of skills is an important indicator because, although formal education may be the primary source of learning, using skills at home or at work may be important for continuing skill development and maintenance and for preventing skill loss. This study looks at the indices of *numeracy skill use at home* and *numeracy skill use at work*, which are based on questions about the use of mathematical skills of varying levels of complexity.

The PIAAC background questionnaire also included questions about respondents' learning characteristics and approaches. The responses to these questions were used to create an index of *readiness to learn*, which is also examined in this study.

Additional information about the questions included in the construction of these indices can be found in appendix B.

Table 9 presents the gender effect for these PIAAC indices. The first column presents the effect size for the index of *numeracy skill use at home*, the second column shows the effect size for *numeracy skill use at work*, and the last column shows the effect size for the index of *readiness to learn*. For all three indices, the effect size is positive, favoring males. The results for numeracy skill use at home and on readiness to learn look very different. The data in the first column show that females use numeracy skills at home less frequently than males; 9 of the 16 countries show a small gender effect on numeracy skill at home and one country, Finland, shows a medium-size gender effect. For numeracy skill use at work, 7 of the 16 countries show a small gender effect and one country, the Netherlands, shows a medium-size gender effect. Only two countries, Japan and the United States, show a gender effect in the index of readiness to learn, and the effect size is small. The United States is the only country with a gender effect in all three indices.

[8] PIAAC indices were constructed using IRT scaling with an international average of 2, and a standard deviation of 1. Each item that is used in calculating the indices uses a five-point Likert scale.

Table 9. Gender differences in skill use and learner characteristics of 23- to 25-year-olds in PIAAC 2012, by country

Country	Gender differences (measured in terms of effect size)		
	Numeracy skill use		
	At home	At work	Readiness to learn
Australia	0.08	0.26	0.04
Czech Republic	-0.04	-0.03	0.14
Denmark	0.39	-0.01	0.16
Finland	0.62	0.33	0.05
France	0.24	0.32	0.01
Ireland	0.36	0.00	0.07
Italy	0.16	-0.01	-0.10
Japan	0.18	0.38	0.42
Rep. of Korea	0.46	0.13	0.18
Netherlands	0.29	0.53	0.06
Norway	0.25	0.38	-0.16
Poland	0.04	-0.08	0.02
Slovak Republic	0.09	-0.03	0.06
Spain	0.28	0.00	-0.14
Sweden	0.41	0.20	-0.06
United States	0.30	0.32	0.31

No gender effect (Cohen's $d \leq .2$)

Small gender effect ($.2 \leq$ Cohen's $d \leq .5$)

Medium gender effect (Cohen's $d \geq .5$)

Section V: Conclusions and Policy Implications

Key Findings

The mathematical literacy performance of the cohort of 15-year-olds in PISA 2003 is very similar to the performance of the same cohort as 23- to 25-year-olds in PIAAC 2012. Looking across countries, there is a strong correlation between the performance of the cohort on both assessments.

In most countries, the gender gap in numeracy performance increased as the age of the respondents increased. Finland was the only country that showed a bigger gap in the youngest age group (16-24) than in the oldest age group (55-65), and Poland was the only country that showed no gender difference in any age group.

The gender effect in the cohort of 15-year-olds in PISA 2003 either stayed the same in PIAAC 2012 (when those in the cohort were 23 to 25 years old) or increased. Approximately half of the countries showed an increase in the gender effect, with Finland and United States showing the largest increase (of 0.5).

In most countries that participated in PISA 2003, males were more engaged in and had more positive attitudes toward learning mathematics than females, although most of these gender effects were small. These results are similar to those described in other analyses of the PISA data (OECD 2015).

In all countries but Japan, more females than males ages 23-25 had completed a university degree. However, many more males than females earned a degree in the STEM-related areas of science, engineering, mathematics, and computing. More females than males choose non-STEM areas and the females who did choose STEM areas more often chose the areas of education sciences or health and welfare. These findings confirm those from a previous study done in the United States indicating that more males than females earn high school credits in these STEM areas (Cunningham, Hoyer, and Sparks 2015).

With the exceptions of Australia, the Czech Republic, Italy, Japan, Poland, and the Slovak Republic, female adults in the countries participating in PIAAC 2012 used their numeracy skills at home less frequently than males did. Finland had the highest effect size of gender on the use of numeracy skills at home among all countries in the study. Females in 8 of the 16 countries in the sample used their numeracy skills at work less often than males; the Netherlands had the highest gender effect on adults' use of numeracy skill at work, with the difference being about half of the standard deviation.

In most countries, there was no gender effect in adults' readiness to learn new ideas and information. The United States and Japan were the only two countries in which females showed slightly less readiness to learn new ideas or information than their male counterparts.

Selected Policy Implications

This paper adds to the existing body of research that shows the underachievement and underrepresentation of females in STEM fields (e.g., see Blickenstaff 2005). Using two large-scale international assessments, this study shows that in most countries, females are not only behind males in their mathematical achievement, but that they also have less positive attitudes toward mathematics. They have weaker self-confidence in their mathematical ability, and they feel more anxious while doing mathematics. Additionally, even though more females than males are getting university degrees, they are still less likely to choose their major area of study in the areas of science and engineering. Moreover, as also seen in previous research (e.g., NCES 2013), these gender differences increase as females move through higher grades and as they age.

These findings suggest that there is a still a long way to go toward gender equity in the STEM fields. Educators at various levels need to understand these differences and work with their female students to improve their attitudes and engagement with STEM fields, which may improve their performance in these areas (OECD 2015). These teachers may need to be trained to recognize and address biases they may have about the performance of females in STEM courses in order to help these females develop their talents. Teachers can also help to improve female students' confidence and attitudes towards mathematics by providing positive reinforcement when they have accomplishments and by providing them with low-stakes opportunities to practice their scientific and mathematical problem solving skills. Another strategy to promote female achievement in STEM courses is for teachers to use certain methods of teaching mathematics, such as cognitive-activation strategies, that may be particularly beneficial for girls (OECD 2015).

It is also important for colleges and universities to create resources and policies to encourage female students to choose and complete their major area of study in the STEM fields. One way to provide to encourage females to complete STEM majors and to enter into careers in the STEM fields is to strengthen career services and provide programs such as job-shadowing opportunities, which would expose these students to the range of career opportunities that are available in STEM. One example of this type of program are the National Boys' Days and Girls' Days that are organized in several countries, such as Belgium and Germany, where students visit universities or businesses and learn about degrees or occupations in areas in which their gender is underrepresented (OECD 2015). Other methods that have been used to encourage greater representation of females in STEM fields in higher education include measures of planning specific initiatives, monitoring and setting specific targets towards gender equity, such as those taken in France, where Mission for Parity in Higher Education and Research. Other common policies include granting special awards to females in STEM fields (European Commission 2010).

Once women are pursuing majors in STEM fields or are in STEM occupations, certain policies may be useful to promote their retention and advancement in these areas. These initiatives include offering financial support, providing mentorship, or improving work-life balance. For example, in the U.S., the National Science Foundation has a program that provides research grants to projects that have a goal to increase the presence of women in academic careers in STEM fields. Another example in the U.S. is that the Department of Energy provides female undergraduates in STEM with mentors in relevant areas. An example of an effort in Australia that seeks to reduce gender disparities in STEM by promoting work-life balance is that the Australian Research Council provides paid maternity leave for fellowships and also has selection criteria that helps those whose careers have been interrupted because of family responsibilities (OECD 2015). Finally, the White House's Educate to Innovate Campaign works through private-public partnerships, with one of its goals being broadening the participation of underrepresented groups in STEM, including women. This initiative aims to achieve this goal by providing female students with experiences and hands on opportunities in STEM fields and appointing female role models to lead the initiative (The White House).

Directions for Future Research

Further investigation into countries such as Finland and Poland is needed to provide more insight into what may be contributing to the patterns found in these countries (e.g. why the gender effect in mathematics is increasing in Finland while Poland shows no gender effect). Additional data sources, such as national labor data, can be used to provide more information on these cohorts, and other variables in the PIAAC data set, such as those relating to employment and occupation, can be examined. Additionally, with the upcoming release of the U.S. PIAAC National Supplement data, which oversampled young adults ages 16-34, more in-depth analysis of the U.S. population will be possible.

Future research should also look at additional cohorts that were included in both PISA and PIAAC—for example, the 15-year-olds in PISA 2006 who were 21 to 23 years old when they were assessed in

PIAAC—to determine whether the patterns found in this study among 23- to 25-year-olds can also be found in other age groups. This research would provide more information as to whether the pattern of gender differences found in this study are relevant just to the specific cohort or whether they are consistently found as a cohort ages.

Additionally, future research could look at gender differences in performance in other STEM areas in large-scale assessments or other domains of PISA and PIAAC. This could include looking at the problem solving in technology-rich environments domain in PIAAC or looking at the science or problem solving domains in PISA. The Trends in International Mathematics and Science Study (TIMSS) also provides information on the performance of fourth- and eighth-grade students in science and mathematics in an international context. Therefore, one could use the TIMSS data to look at gender differences in these age groups, although its framework is not so closely related to the framework of PISA and PIAAC, since it focuses on measuring how well students have learned the curriculum rather than application of knowledge to real world situations like PISA and PIAAC do.

Another direction for future research would be to look more closely at the mathematical literacy items in PISA and the numeracy items in PIAAC in order to study the specific content areas and determine if gender differences in performance vary across these areas.

References

Chen, X., and Ho, P. (2012). *STEM in Postsecondary Education: Entrance, Attrition, and Course taking Among 2003-2004 Beginning Postsecondary Students*. Washington, DC: National Center for Education Statistics.

Blickenstaff, J.C. (2005). Women and Science Careers: Leaky Pipeline or Gender Filter? *Gender and Education*, *17*(4), 369-386.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, *112* (1), 155-159.

Cunningham, B., Hoyer, K.M., and Sparks, D. (2015). *Gender Differences in Science, Technology, Engineering, and Mathematics (STEM) Interest, Credits Earned, and NAEP Performance in the 12th Grade* (NCES 2015-075). Washington, DC: National Center for Education Statistics.

European Commission (2010), *Gender differences in educational outcomes: Study on Measures Taken and the Current Situation in Europe*, Education, Audio-visual and Culture Executive Agency.

Hill, C., Corbett, C., and St. Rose, A. (2010). *Why So Few? Women in Science, Technology, Engineering, and Mathematics*. Washington, DC: American Association of University Women.

Lantz, B. (2013). The large sample size fallacy. *Scandinavian journal of caring sciences*, *27*(2), 487-492.

Lindemann, D.J. (2015). Gender and Numeracy Skill Use: Cross-National Revelations From PIAAC. Retrieved from https://piaac.squarespace.com/s/Lindemann_PIAAC.pdf.

Lundetræ, K., Sulkunen, S., Gabrielsen, E., and Malin, A. (2014). A Comparison of PIAAC and PISA Results. *Associations Between Age and Cognitive Foundation Skills in the Nordic Countries*, p. 171.

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. International Association for the Evaluation of Educational Achievement. Herengracht 487, Amsterdam, 1017 BT, The Netherlands.

Mullis, I. V., Martin, M. O., Fierros, E. G., Goldberg, A. L., & Stemler, S. E. (2000). *Gender differences in achievement: IEA's third international mathematics and science study (TIMSS)*. TIMSS International Study Center, Boston College.

National Center for Education Statistics. (2013). *The Nation's Report Card: Trends in Academic Progress 2012* (NCES 2013-456). Washington, DC: Author.

National Science Board. (2010). *Preparing the Next Generation of STEM Innovators: Identifying and Developing Our Nation's Human Capital.* Arlington, VA: National Science Foundation

Organization for Economic Cooperation and Development (OECD). (2004). *Learning for Tomorrow's World: First Results From PISA 2003*. Paris, France: Author.

Organization for Economic Cooperation and Development (OECD). (2013a). *OECD Skills Outlook 2013: First Results From the Survey of Adult Skills*. OECD Publishing.

Organization for Economic Cooperation and Development (OECD). (2013b). *The Survey of Adult Skills: Reader's Companion*. OECD Publishing.

Organization for Economic Cooperation and Development (OECD). (2013c). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy.* OECD Publishing.

Organization for Economic Cooperation and Development (OECD). (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, and Confidence*. OECD Publishing.

PIAAC Numeracy Expert Group. (2009). *PIAAC Numeracy: A Conceptual Framework* (No. 35). OECD Publishing.

Shauman, K.A. (2006). Occupational Sex Segregation and the Earnings of Occupations: What Causes the Link Among College-Educated Workers? *Social Science Research*, *35*, 577-619.

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size—or Why the *P* Value Is Not Enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. http://doi.org/10.4300/JGME-D-12-00156.1

Snyder, T. D., & Dillow, S. A. (2012). *Digest of education statistics 2011*. National Center for Education Statistics.

The White House. (n. d.). *Educate to innovate*. Retrieved June 24, 2015, from https://www.whitehouse.gov/issues/education/k-12/educate-innovate

Appendix A: Total Number and Average Scores in PISA and PIAAC, by Gender

Table A-1. Total number and average mathematics scores of 15-year-olds in PISA 2003 and total number and average numeracy scores of 23- to 25-year-olds in PIAAC 2012, by gender and country

Country	PISA 2003						PIAAC					
	Total number of males	Total number of females	Average mathematics scores				Total number of males	Total number of females	Average numeracy scores			
			Male		Female				Male		Female	
Australia[1]	6,340	6,220	522	(2.70)	527	(3.00)	160	160	268	(5.90)	265	(5.60)
Czech Republic	3,240	3,080	509	(4.40)	524	(4.30)	200	220	288	(4.30)	283	(4.80)
Denmark	2,080	2,140	506	(3.00)	523	(3.40)	140	140	287	(5.20)	280	(4.50)
Finland	2,870	2,930	541	(2.10)	548	(2.50)	130	130	311	(4.30)	283	(5.80)
France	2,030	2,270	507	(2.90)	515	(3.60)	150	170	274	(4.00)	263	(3.60)
Ireland	1,970	1,910	495	(3.40)	510	(3.00)	120	160	264	(5.80)	263	(5.30)
Italy[2]	5,620	6,020	457	(3.80)	475	(4.60)	90	90	244	(7.30)	257	(5.00)
Japan	2,300	2,400	530	(4.00)	539	(5.80)	120	120	294	(4.30)	286	(4.90)
Rep. of Korea	3,210	2,230	528	(5.30)	552	(4.40)	140	160	290	(3.80)	278	(3.60)
Netherlands	2,020	1,980	535	(3.50)	540	(4.10)	130	120	297	(4.20)	280	(5.30)
Norway	2,050	2,010	492	(2.90)	498	(2.80)	150	130	289	(5.00)	276	(5.10)
Poland[3]	2,180	2,200	487	(2.90)	493	(3.00)	1,020	950	273	(1.80)	270	(2.00)
Slovak Republic	3,730	3,610	489	(3.60)	507	(3.90)	170	170	286	(4.90)	278	(4.30)
Spain[2]	5,240	5,550	481	(2.20)	490	(3.40)	150	150	267	(3.80)	258	(4.50)
Sweden	2,340	2,280	506	(3.10)	512	(3.00)	130	110	295	(4.20)	291	(4.90)
United States	2,740	2,720	480	(3.20)	486	(3.30)	140	170	274	(5.70)	246	(3.90)

1. Australia drew a PISA sample that is substantially larger than the minimum requirements for two reasons: (1) students who participate in PISA are invited to take part in the Longitudinal Surveys of Australian Youth (LSAY); and (2) smaller states, territories, and indigenous students are oversampled.

2. In addition to Australia, Spain and Italy, oversampled in PISA to allow for regional comparisons.

3. Poland oversampled individuals ages 16 to 29 in PIAAC.

NOTE: The total number of males and females for each assessment in each country is rounded to the nearest 10.

Appendix B: PISA and PIAAC Indices

PISA Indices

Instrumental motivation in mathematics

The PISA index of instrumental motivation in mathematics was derived from students' reported agreement with the following statements:

- making an effort in mathematics is worth it because it will help me in the work that I want to do later on;
- learning mathematics is important because it will help me with the subjects that I want to study further on in school;
- mathematics is an important subject for me because I need it for what I want to study later on; and
- I will learn many things in mathematics that will help me get a job.

Interest in and enjoyment of mathematics

The PISA index of interest in and enjoyment of mathematics was derived from students' reported agreement with the following statements:

- I enjoy reading about mathematics;
- I look forward to my mathematics lessons;
- I do mathematics because I enjoy it; and
- I am interested in the things I learn in mathematics.

Anxiety in mathematics

The PISA index of anxiety in mathematics was derived from students' reported agreement with the following statements:

- I often worry that it will be difficult for me in mathematics classes;
- I get very tense when I have to do mathematics homework;
- I get very nervous doing mathematics problems;
- I feel helpless when doing a mathematics problem; and
- I worry that I will get poor <marks> in mathematics.

Self-efficacy in mathematics

The PISA index of self-efficacy in mathematics was derived from students' reported level of confidence with the following calculations:

- using a <train timetable>, how long it would take to get from Zedville to Zedtown;
- calculating how much cheaper a TV would be after a 30 per cent discount;
- calculating how many square metres of tiles you need to cover a floor;
- understanding graphs presented in newspapers; solving an equation like $3x + 5 = 17$;
- finding the actual distance between two places on a map with a 1:10,000 scale;
- solving an equation like $2(x+3) = (x + 3)(x - 3)$; and
- calculating the petrol consumption rate of a car.

Self-concept in mathematics

The PISA index of self-concept in mathematics is derived from students' level of agreement with the following statements:

- I am just not good at mathematics;
- I get good <marks> in mathematics;
- I learn mathematics quickly;
- I have always believed that mathematics is one of my best subjects; and
- in my mathematics class, I understand even the most difficult work.

PIAAC Indices

Numeracy skill use

The PIAAC indices of numeracy skill use were derived from respondent's responses to how often they usually do the following activities in their current/previous job or outside their work/in everyday life:

- calculate prices, costs or budgets;
- use or calculate fractions, decimals or percentages;
- use a calculator - either hand-held or computer-based;
- prepare charts, graphs or tables;
- use simple algebra or formulas: and
- use more advanced math or statistics such as calculus, complex algebra, trigonometry or use of regression techniques.

Index of readiness to learn

The PIAAC index of readiness to learn was derived from respondents' reported agreement with the following statements:

- when I hear or read about new ideas, I try to relate them to real life situations to which they might apply;
- I like learning new things;
- when I come across something new, I try to relate it to what I already know;
- I like to get to the bottom of difficult things;
- I like to figure out how different ideas fit together; and
- if I don't understand something, I look for additional information to make it clearer.