International Assessments of Student Achievement and Public Confidence in Education: Evidence from a Cross-National Study

Oren Pizmony-Levy
Teachers College, Columbia University

Peter Bjorklund Jr.
University of California, San Diego

Corresponding author:
Oren Pizmony-Levy
Department of International and Transcultural Studies
Teachers College, Columbia University
New York, NY 10027
Email: pizmony-levy@tc.columbia.edu

Word count (including tables and bibliography): 10,327

Biographical notes

Oren Pizmony-Levy is an Assistant Professor in the Department of International and Transcultural Studies. He holds a PhD in sociology and comparative and international education from Indiana University-Bloomington. His research and teaching focus on the intersection between education and social movements, such as the accountability movement and its role in the emergence of international large-scale assessments.

Peter Bjorklund Jr. is PhD student in the Department of Education Studies at the University of California, San Diego. His research focuses on educational change and the analysis of trust, social networks, teacher-student relationships, and teacher identity**.**

International Assessments of Student Achievement and Public Confidence in Education: Evidence from a Cross-National Study

Abstract

One of the overarching goals of international large-scale assessments (ILSA) is to inform public discourse about the quality of education different countries. To fulfill this function, the Organisation for Economic Co-operation and Development (OECD), for example, raises awareness of the Program for International Student Assessment (PISA) results through different forms of traditional and social media (e.g., press releases and other activities under the slogan PISA Day). Scholars have responded to the rapid growth of ILSA by examining public discourse through newspapers articles, policy documents, and other outlets. However, we know very little about whether and to what extent the general public is actually affected by PISA results. This article expands the range of stakeholders that engage with PISA by exploring public opinion. Specifically, the study uses data regarding public trust in education from the 2011 wave of the International Social Survey Program (ISSP). Drawing on survey data from 30 countries and Hierarchical Linear Models (HLM), the study shows that PISA ranking have a significant effect on public perceptions. After taking into account gross domestic product (GDP), we find that in high performing countries the general public expresses higher levels of confidence in the education system. We discuss these patterns in the context of growing politization of education policy making and the use of ILSA as evidence.

Keywords: Public opinion; trust; OECD, PISA; education system; cross-national

Introduction

Following the rapid growth of country participation in international large-scale assessments (ILSA), there is a great interest in the implications of these assessments for education policy (Meyer & Benavot, 2013; Wiseman, 2010). After all, in their current configuration, ILSA are designed to inform the work of policy makers (Howie & Plomp, 2005; Pizmony-Levy, 2014) and to "unsettle complacency in education policy terms within participating nations" (Lingard, 2016, 615). Indeed, leaders of international organizations responsible for ILSA—the Organisation for Economic Co-operation and Development (OECD) and the International Association for the Evaluation of Educational Achievement (IEA)—have recognized the importance of infusing ILSA reports into the public sphere as a mechanism for policy change (Schleicher, 2006; Wagemaker, 2004).

Scholars often draw on media coverage of ILSA reports as a "window" into the impact of ILSA on education policy and to the extent to which various stakeholders interpret these reports to advance their causes (Martens & Niemann, 2013; Pons, 2011; Takayama, 2007, 2008; Yasukawa, Hamilton, & Evans, 2017). However, we know very little about public reactions to ILSA (Fladmoe, 2012; Morgan & Poppe, 2012; Pizmony-Levy, 2017a). For our purpose, we define "public opinion" as opinions concerning social or governmental matters rather than on private matters (such as one's favorite TV show; Clawson & Oxley, 2013). It comprises of opinions, attitudes, preferences, beliefs, and values. These opinions are held by individuals in society, but they become public through technical practices such as standardized representative surveys (Perrin & McFarland, 2011).

This article focuses on one specific aspect of public opinion: confidence in social institutions. Also known as social contract trust, institutional confidence is the expectation of

appropriate behavior based on the norms of that institution. The public, for example, generally trusts the education system in a given country will do its utmost to educate children and prepare them to be contributing members of society. Institutional confidence is the basis of "taken for granted" aspects of social interaction, and a necessary ingredient for cooperative actions and foundation for social capital (Coleman 1988, 1990).

Understanding public opinion towards education is important for at least three reasons. First, social institutions—such as education—are dependent upon public support for legitimacy (Jacobsen, Saultz, & Snyder, 2013; Meyer & Rowan, 1977). Public confidence/trust is necessary for political leaders to initiate and implement policies that involve allocation of resources to attain societal goals (see review: Chanley, Rudolph & Rahn, 2000). Second, public opinion is an important factor in the development of policy. In their Advocacy Collation Framework, Sabatier and Jenkins-Smith (1999) argue that changes in public opinion represent external events that are critical prerequisites to major policy change. The premise of this argument is that public opinion can alter general spending priorities and perceived seriousness of various problems (Berkman & Plutzer, 2005). The relative stability of public opinion is also seen as a key mechanism in the path dependency of social policy because it signals to policy makers what constituents expect (Brooks & Manza 2006; Burstein, 1998). Third, the overall goal of ILSA is to inform educational stakeholders (Howie & Plomp, 2005). The conceptual framework for the Program for International Student Assessment (PISA), for example, called for "solid internationally comparable evidence of educational outcomes" that would address information gaps among policy makers and the general public (OECD, 1999). The question remains whether the release of PISA results affected public opinion towards education. If national performance on ILSA affects confidence and trust in education, it could erode support for public spending in education

and generate legitimacy (and manufacture demand) for educational reforms that challenge public education.

This study investigates the relationship between national performance on PISA and public confidence in education across 30 countries in 2011. Our analyses examine a set of four different measures of performance for each PISA domain (mathematics, science, and reading). This set of measures includes country average score, relative position of the country average score to the OECD average, country ranking, and ranking-based grouping (top 10, 11-20, and below 20). Using Hierarchical Linear Models (HLM) with data from the International Social Survey Program (ISSP), this study demonstrates that performance on PISA is important for public confidence in education. We conclude with a call for more research on the link between ILSA and public opinion.

Literature Review

Public confidence in education

The literature on confidence in education is limited (e.g., Coleman, 1990; Fladmoe, 2012; Guppy & Davies, 1999; Jacobson, 2009; Klugman & Xu, 2008; Loveless, 1997; Newton & Norris, 2000; Lipset & Schneider, 1987; Lyons & Lowery, 1986; Weiler, 1983). While there are slight differences between these studies in terms of focus and time period, they generally converge on one key point. They all provide evidence of a steady erosion of public confidence in education. Jacobsen (2009) and Klugman and Xu (2008), for example, show a dramatic reduction in the share of respondents in the United States who have a "great deal" of confidence in public schools over time.

There is also consensus among scholars regarding the importance or implications of studying confidence in education. For example, Loveless (1997), Guppy and Davies (1999), and

Jacobsen (2009) all posit that confidence in education has a strong impact on public behavior. Jacobsen and her colleagues (2013) contend that public schools are particularly reliant on public support as they derive legitimacy and funding from the electorate. Coleman (1990) also contends that decline in confidence in public education has led to voters defeating bond measures and tax increases that constrain public schools' ability to function and lead to the increased use of private schools. Finally, Coleman (1990) suggests that when trust is withdrawn from one arena (e.g. public schools) it is frequently placed in another (e.g. private schools or charter schools). Moreover, dissatisfaction or lack of trust with public services often leads to citizen action which often manifests itself as "exit" or "voice" (Jacobsen, 2009; James & Moseley, 2014).

Exit refers to families or individuals opting to leave a school and seeking educational alternatives such as private schooling or charter schools (Jacobsen, 2009). Performance indicators can prove integral to decision-making and choice to exit one institution and enter another (James & John, 2007). Voice, on the other hand, describes the process in which individuals act together to express their dissatisfaction with a service or institution through lobbying, petitioning, or protest (Jacobsen, 2009; James & John, 2007; James & Moseley, 2014). The notions of exit and voice are important to mention when discussing confidence in education systems. Lack of confidence in public education can lead parents to "exit" their students and enroll them in privatized education systems, potentially damaging the public education system. (Jacobsen, 2009). Lack of confidence in the education system can also bring about "voice" as people take a stand and advocate for changes to be made to alter the system (Jacobsen, 2009). Whether or not parents "exit" their students or "voice" their concerns are generally thought to be related to information of their education system and this usually comes in the form of performance data (more below).

Another reason confidence in education is compelling and meaningful to study lies in the notion that public confidence has long served as a thermostat for driving reform (Klugman & Xu, 2008; Guppy & Davies, 1999; Loveless, 1997). According to Klugman and Xu (2008), declining levels of confidence in education in the United States has led to support or lack of support for various institutional and organizational reforms conceived to address school failures. Similarly, in their study of confidence in education in Canada, Guppy and Davies (1999) attribute support for educational reform initiatives to a lack of confidence in public schooling. Here, confidence in education is likely seen as a measure of school quality and performance; reform is regarded as a solution to improve these areas. It should be noted that literature on public confidence/trust in education is also limited by the fact that most if not all research published in English on this topic is situated one country, generally the United States, Canada, or the UK (an exception is Fladmoe, 2012).[1] This study hopes to add to the literature by looking at a multinational sample and roles the variables mentioned above play in this sample.

Scholars have documented the role of individual-level characteristics on public confidence in education. Many agree that there is an inverse relationship between education attainment and confidence in education (Clawson & Oxley, 2012; Klugman & Xu, 2008; Guppy & Davies, 1999). However, there are mixed findings with respect to other variables. Klugman and Xu (2008) find that people with higher incomes and who live in urban settings are all less likely to express confidence in education. Bali (2016) had similar findings, showing people who live in urban communities had lower confidence in schools than people in non-urban communities. However, Guppy and Davies (1999) found no significant relationship between confidence and income, age, or having children at home. Loveless (1997) and Bali (2016)

[1] Our literature review is limited to studies published in English. It is very likely that additional studies on public opinion towards education are published in other languages.

contradict Guppy and Davies (1999) by asserting that people with children in school tend to have more confidence in the schools. Lastly, employment status has been shown to have a relationship to confidence in education. People who are unemployed tend to have more confidence in education than people who employed. Thus we will control for age, gender, level of education, income, community, having children in school, and employment status. We posit that beyond these individual factors information about performance of schools will also impact public confidence in education.

Cues, performance data, PISA, and the media

Information is integral to democratic governance and enabling citizens to exert some level of control over the government and public institutions—like education (James, 2010; James & John, 2007). That being said, citizens rarely have complete knowledge about political issues or the performance of public institutions, nor do they have much of an incentive to become informed (James, 2010; James & John, 2007).[2] Frequently issues regarding the performance of public institutions—like schools—are complex and people do not have the time nor the bandwidth to fully explore data about these institutions (James, 2010, 2011; James & John, 2007; James & Moseley, 2014; Soroka, 2006). As a result, citizens tend to use information shortcuts or *cues* to "economize on information" and develop their views, thus overcoming the demands associated with being fully informed about the performance of public institutions (James, 2010, 2011; James & John, 2007; James & Moseley, 2014).

Citizens receive cues about the performance public institutions from personal experience, word-of-mouth, and the media (Coleman, 1990; James & Moseley, 2014). These cues often

[2] Citizens vary in their motivations to become informed on political issues or the performance of public institutions. Parents of school-age children, for example, may be more interested than others in the performance of schools and the overall education system when they make decisions about their children's education.

come in the form of summary performance data or data that reports on the absolute or relative performance of an institution. Studies of media effects have consistently pointed to the fact that sources and presentation of performance data affect its reception by the population (James, 2010). People will be less inclined to trust information and performance that comes from politicians or public institutions because they believe these actors are more likely to misrepresent the data (James, 2010). As a result, people often use various media outlets as arbiters of performance data regarding public institutions (Colman, 1990; James & Moseley, 2014). And more and more the media is using the results of ILSA like PISA to evaluate the performance of education systems.

Over the last decade PISA become increasingly popular in the media. It fuels education debates, and some argue that it has an impact on governance of the education system in many countries (Pons, 2011; Takayama, 2007, 2008). Several scholars have examined the discourse around PISA results and how it is received by the media (Dixon et al., 2013; Figazzolo, 2009; Stack, 2007; Steiner-Khamsi, 2003; Takayama, 2010). They have found that the media tend to endorse results with little critique (Stack, 2007), be influenced by shifts in political and social conditions (Takayama, 2007, 2008), focus on negative reactions (Dixon et al., 2013), provide little background about PISA results and what they mean (Pizmony-Levy, 2017b), and focus on international rankings more than anything else (Pizmony-Levy, 2017b; Pons, 2011). Some scholars have indicated that increased media attention to PISA results has influenced public opinion regarding education, but few have offered empirical evidence for such arguments (Fladmoe, 2012). Pizmony-Levy (2017a) found that people in the United States and Israel tend to be misinformed about PISA results. Moreover, recent studies examining PISA's relation to

public discourse are limited by the fact that they draw on qualitative analysis of samples of newspaper articles and thus are limited in their generalizability.

Frequently, it matters less about what data, like the PISA results, are reported rather than how it is reported. Studies have found that information cues about relative performance are more helpful than information about absolute performance because relative performance gives people a benchmark to make their judgements while absolute performance can seem abstract (James, 2010). We will examine this by including various presentations of ranking PISA results in our models. The effects of public performance data on citizens' perceptions have not been fully addressed by the literature and need more investigation (James, 2010; James & Moseley, 2014). Moreover, scant research has investigated how performance information affects confidence in public institutions at the international level.

Theoretical Framework

We draw on two different theories for our framework: Newton and Norris' (2000) institutional performance model and the James and Moseley's (2014) nascent performance information theory. Newton and Norris' (2000) institutional performance model posits that people form trust in public institutions because institutions perform well or that people lose trust them when they perform poorly. They set this in contrast to two competing theories on how people come to trust public institutions. One alternate theory is the socio-psychological explanation: the theory that people's trust in institutions is a part of their personality—i.e., they have a trusting disposition. The other is the socio-cultural model of trust that contends that peoples' trust in institutions is a product of their life experience and socialization. In other words, these theories focus on individual-level characteristics to explain variation in confidence in public institutions. In this study we address these theories by including several individual-level variables in the models.

The incipient performance information theory suggests that public performance information about institutions will have an effect on confidence and the behavior of citizens towards those institutions (James & Moseley, 2014). The two differ slightly in that the latter is focused more on the impact of public performance information while the former has a broader understanding of how people garner information about performance. Moreover, the institutional performance model is set in contrast to other theories about how people form trust. Both theories provide a useful lens to examine the relationship between PISA results and public confidence in education systems. We expect to find that people's confidence in education will have a positive relationship with their country performance on PISA.

For the purpose of this study we use country performance on PISA as a proxy for the potential sentiments spread by policy makers and the media about the education system. We assume that in high preforming countries, overall, policy makers and the media will have less opportunity to interpret the results in a negative fashion. Similarly, we assume that in low preforming countries, overall, policy makers and the media will have more opportunity to use PISA results to scandalize the education system. As for countries that show an unexceptional performance (neither high nor low), we assume that policy makers and the media will have more flexibility in interpreting and framing the results. We recognize that public opinion towards education is not directly shaped by PISA performance; rather it is indirectly influenced by PISA performance as the scores and rankings create the opportunity to shape the portrayal of the education system which in turn could affect public opinion.

Data & Methods

Sample

Data for this study came from the International Social Survey Programme (ISSP), a collaboration conducting standardized, comparable, nationally representative surveys (ISSP, 2017). Founded in 1984, the ISSP comprises of research organizations such as universities, academic organizations, and survey agencies; each organization represents one nation and funds all of its own costs (Smith, 2009). The ISSP does not provide financial support for conducting national surveys and has no central funds for organizational or methodological tasks (ISSP, 2017). By now, the ISSP has delivered a series of cross-national surveys of adult respondents aged 18 and older on myriad social science topics, including environment, family and changing gender roles, religion, role of government, social inequality, and work orientation.

The 2011 ISSP module ("Health and Health Care") included measures of respondents' evaluation of health care system in their country, personal health, and access to health insurance. For the purpose of comparison, the survey also included measures of respondents' evaluation of the education system. The 2011 ISSP module was conducted in 32 countries. The average response rate was 54.3% (Gendall, 2013). We restrict our analysis to the 30 countries that participated in the 2009 cycle of PISA. After listwise deletion, our final sample included 42,331 individuals from Australia, Belgium (Flanders and Wallonia), Bulgaria, Chile, China, Chinese Taipei, Croatia, Czech Republic, Denmark, Finland, France, Germany (East and West), Israel (Jewish and Arab descent), Italy, Japan, Korea, Lithuania, Netherlands, Norway, Poland, Portugal, Russian Federation, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, and the United States.

Nation-level data are drawn from the OECD Report titled *PISA 2009 Results: What Students Know and Can Do*. The report presents detailed information about student performance in reading, mathematics and science, including average scores, ranking, differences within countries (e.g., based on gender), and differences between countries. Following the public release of the report in December 2010, media outlets in various countries published news stories about student performance in PISA (see discussion above). We use PISA because researchers posit that the source of performance information of public institutions, like schools, has an effect on how it is received. Public performance information that comes from independent agents or bodies seem to bridge partisan divides and have similar effects on citizens of all political leanings (James, 2010; James & Moseley, 2014). One example of an ostensibly independent agent is the OECD and it has made large strides in creating performance measurement systems, like PISA, that are intended for consumption by citizens and policy makers alike (James & John, 2007). Nation-level data are also drawn from reports by the World Bank (WB) and the International Monetary Fund (IMF), which are publicly available sources that provide aggregate nation-level data about economic development.

Dependent variable

The ISSP dataset contained one measure that assessed confidence in the education system. Respondents were asked, "In general, how much confidence do you have in the educational system in [country]." Respondents were asked to rate their confidence on a five-point scale ranging from 1 "complete confidence" to 5 "no confidence at all". For ease of interpretation, we reversed the coding of this variable so that a higher score indicates higher confidence.

Country-level independent variables

The central hypothesis of this paper concerns the relationship between student performance in PISA 2009 and confidence in education. We decided to use PISA because the number of countries and systems participating in PISA is higher than in any other ILSA. Moreover, the fact that PISA is conducted by the OECD gives it more importance among policy makers and the media. We decided to use the 2009 cycle because it was the last ILSA released before the administration of the ISSP module in 2011.

For each of the PISA domains—reading, math, and science—we generated four measures of performance. The first measure is country average score, which is reported on a scale of 0 to 1000 (Most scores fall between 200 and 800). In addition to scores, we also generated measures that frame student performance in a comparative or relational perspective. The second measure is the relative position of the country average score to the OECD average. This information was coded as a trichotomous variable: (1) statistically significant above the OECD average, (2) not statistically different than the OECD average; (3) statistically significant below the OECD average. The next two measures focus on rank order of countries. Scholars noted that rankings are often what is reported and discussed in the media (Pons, 2011; Pizmony-Levy, 2017b), therefore they may have a larger impact on confidence than test scores. The third measure is the country ranking. The fourth measure is simplified country ranking. This information was coded as a trichotomous variable: (1) top ten countries in terms of ranking, (2) countries ranked 11-20, and (3) countries ranked 21 or below.

Country-level controls

We examine country-level controls that previous research links with trust/confidence in institutions. Because previous research suggests that economic development is important to the creation of trust (Delhey & Newton, 2005; Freitag & Bühlmann, 2009; Paxton, 2007), we control

for Gross Domestic Product (GDP) per capita, logged. The countries in the sample do vary in their GDP per capita, but this variation is lower than we would find in the large population of countries.

Scholars have noted a difference in institutional trust in more collectivist societies as opposed to more individualistic societies. For example, Helgesen (2006) found there were distinctly higher levels of trust in Denmark, Sweden, and Finland compared to South Korea and Japan. Huff and Kelly (2003) find that people from collectivist Asian countries tend to have lower levels of general trust compared to people in the individualistic U.S. Thus, we use a dummy variable for the four Asian countries included in the ISSP data: China, Chinese Taipei, Japan, and Korea (1 = yes, 0 = no).

Individual-level controls

Models control for individual characteristics that previous research identifies as relevant for confidence in public institutions. Gender is measured as a dummy variable (1 = woman, 0 = man). Age and age-squared are measured in years. Education is based on self-reported highest completed education level; we use a series of three dummies: less than secondary degree (reference category), secondary degree, and university degree or higher. Labor market status is captured in a series of three dummies: employed, unemployed, and not in labor force (i.e., students, apprentice or trainee, permanently sick or with disability, retired, and domestic work). We include relative income based on country-specific z-scores, which allow for international comparison without currency conversion. Parental status is measured as a dummy variable (1 = at least one child in the household, 0 = no children in the household). Community type is measured as a series of three dummies: urban, suburb and town, and rural. Taken together, these characteristics capture different aspects of the socio-psychological and socio-cultural expiations

(Newton & Norris, 2000) for the formation of confidence in public institutions. Table 1 presents definitions, metrics, and descriptive statistics for all variables in the study.

Table 1 about here

Analysis

The ISSP dataset is hierarchically organized with individuals nested within countries, and information at the both the individual level and the country level is used to determine trust in the educational system. The clustering of individuals within countries violates assumptions of the ordinary least square regression (OLS). Therefore, we use HLM technique that simultaneously estimates individual and country-level effects. Specifically, we estimate random intercept models, which can be expressed as two equations (Raudenbush & Bryk, 2002). First, the level of trust in educational system for the i^{th} individual in the j^{th} country is a function of country intercepts (β_{0j}), a set of individual-level fixed effects (βX_{ij}), and an error term (r_{ij}):

Equation #1: $Y_{ij} = \beta_{0j} + \beta X_{ij} + r_{ij}$

Second, each country intercept (β_{0j}) is estimated as a function of an intercept (γ_{00}), a set of country-level variables (γ_{Cj}), and an error term (u_{0j}):

Equation #2: $\beta_{0j} = \gamma_{00} + \gamma_{Cj} + u_{0j}$

We examine a series of models for each PISA domain (reading, math, and science). Our strategy is to comprehensively examine the effects of *different* forms of student performance in PISA on confidence in educational system. All models include controls for individual-level and country-level variables. We estimated the models in Stata V14, using the xtmixed command.

Results

Before discussing multilevel models, we consider country-level descriptive patterns (Table 2). Finland shows the highest level of confidence in educational system (M=3.93; SD=.65), where close to four-fifths of respondents (78.9 percent) say they have a great deal of confidence or complete confidence (the second highest and highest response categories on the item respectively) in the educational system (63.2 percent and 15.7 percent, respectively). Other countries with high levels of confidence in educational system are Norway, Denmark, China, and Switzerland. Chinese Taipei shows the lowest level of confidence in educational system (M=2.52; SD=.99), where less than one-fifth of respondents (16.2 percent) say they have a great deal of confidence or complete confidence in the educational system (14.4 percent and 1.8 percent, respectively). Other countries with low levels of confidence in educational system are Chile, Japan, Korea, and Poland.

Table 2 about here

In addition to descriptive patterns, we also examine bivariate correlations between country-level confidence in educational system and performance in PISA 2009. Across all three PISA domains, the correlation between attitudes and country average PISA score is positive and significant (math: $r = .33$, $p<.05$; science: $r = .29$, $p<.05$; reading: $r = .32$, $p<.05$). We found a similar pattern—but reversed in direction—for the correlation between attitudes and country ranking in PISA (math: $r = -.35$, $p<.05$; science: $r = -.22$, $p<.05$; reading: $r = -.31$, $p<.05$). In other words, countries that perform well on PISA also show high levels of confidence in educational system.[3]

[3] We also examine bivariate correlations between country-level variation in confidence in educational system and performance in PISA 2009. Across all three PISA domains, the correlation between the variation in attitudes (i.e., standard deviation) and country average PISA

HLM Models with Performance in Mathematics

Table 3 presents HLM models of confidence in educational system and tests the effect of national performance on PISA in mathematics across all 30 countries. We initially estimated the unconditional model, which does not include any independent variables (not reported). The model illustrated that confidence in educational system vary significantly across the 30 countries. The intraclass correlation coefficient (ICC), suggested that 13.2 percent of the variation in confidence in educational system is associated with differences between countries. This finding substantiates the application of HLM models.

Table 3 about here

Model 1 introduces all the individual-level variables. Except for income, all variables have a significant relationship to confidence in the education system.[4] Across countries, women show more confidence in educational system. Although the relationship is significant, the coefficient is relatively small (b=.028, p<.01). Age shows a U-shaped and statistically significant relationship with confidence in educational system (see negative coefficient for age and the positive coefficient for age squared). The inflection point in the estimates is reached at 45. Alternative specifications of the variable age (e.g., age groups) produce similar patterns. Parents show more confidence in educational system (b=.081, p<.001) than people without children. Respondents with higher levels of education show less confidence in educational system.

---

score is negative and significant (math: r = -.43, p<.05; science: r = -.54, p<.05; reading: r = -.51, p<.05). Performance in PISA 2009 is not only associated with the level of confidence in educational systems, but also with the level of agreement among the general public. Countries that perform well on PISA also show high levels of agreement among the general public with respect to confidence in educational system. Perhaps this is related to interpretation work (by the media and policy makers) that is possible in countries that are ranked in the middle or the bottom of the PISA rankings table, but not in countries that perform well.

[4] In other model specifications (available upon request) we found that income has no effect on confidence in educational system. For example, we found that the bivariate coefficient is small and non-significant (b=.002, p=.674).

Compared to respondents who are employed, those who are unemployed or are not in the labor force show less confidence in educational system. Finally, respondents in urban communities show less confidence in educational system. These patterns hold even after we control for country-level variables (Models 2-6).

Model 2 introduces the country-level control variables. The coefficient for economic development, as reflected in GDP per capita (logged), is not statistically significant. That is, respondents in more advanced economies are not statistically different than their counterparts in less advanced economies. This pattern is not surprising given that all of the countries in the analysis are industrialized nations with fairly advanced economies. The coefficient for Asian countries is not statistically significant. However, the size of the coefficient increases and becomes significant once we control for national performance on PISA (Models 3-6).

Next, we assess the relationship between PISA results and public confidence in education. Specifically, we explore national performance in mathematics, a domain that is considered to be most suitable for international comparison (Husén, 1967). Looking at performance in the form of country average score (Model 3), the results show that performance on PISA in mathematics is associated with a significant increase in confidence in educational system ($b=.009$, $p<.001$). The effect size is relatively large (.80; see Tymms [2004] for discussion on the calculation of effect sizes in multilevel models).

Model 4 includes a series of dummy variables that capture the country average score in relation to the OECD average score. This specification is important because the average score of OECD countries—the original target of PISA—is a common reference by which the media and policy makers assess performance on PISA. Compared to respondents in countries performing above the OECD average (reference category), respondents in countries performing at the OECD

average show less confidence in educational system (b=-.263, p<.06). A similar pattern is evident for respondents in countries performing below the OECD average show (b=-.552, p<.001). In additional analysis (available upon request) we found significant difference between respondents in countries performing at the OECD average and respondents in countries performing below the OECD average (b=-.289, p<.05).[5]

Model 5 introduces national performance on PISA in the form of country ranking. Similar to the patterns discussed earlier, lower ranking on PISA in mathematics is associated with a significant decrease in confidence in educational system (b=-.026, p<.001). Finally, Model 6 introduces a simplified country ranking with a series of dummy variables. Respondents in top ten countries (the reference category) show the highest level of confidence in educational system. The differences between respondents in top ten countries and respondents in countries ranked between 11-20 is marginally significant (b=-.389, p<.08). The differences between respondents in top ten countries and respondents in countries ranked 21 or below is statically significant (b=-.699, p<.001). In additional analysis (available upon request) we found significant difference between respondents in top twenty countries and countries ranked 21 or below (-.310, p<.05).

HLM Models with Performance in Science and Reading

Table 4 presents HLM models of confidence in educational system and tests the effect of national performance on PISA in science (Models 1-4) and reading (Models 5-8) across all 30 countries. As noted, all models include controls for individual-level and country-level variables.

[5] This coefficient is easily calculated by subtracting the coefficient for the third category (below OECD average) from the coefficient for the second category (no difference from the OECD average). For example: (-.552) – (-.263) = .289.

For brevity, however, we remove the coefficients for these controls from the table (complete tables are available upon request).

Table 4 about here

Overall, the results for science and reading echo the patterns discussed earlier for mathematics. This is not surprising given that the correlations between country average scores are very strong (for mathematics and science: $r=.94$, $p<.001$; for mathematics and reading: $r=.90$, $p<.001$). Country average score in science (Model 1) and reading (Model 5) is positively associated with confidence in educational system. The effect size for science is modest (.54) and the effect size for reading is large (.71). The coefficients for the dummy variables that capture the country average score in relation to the OECD average score (Models 2 and 6) are also positively related, but they are not statistical significance for science and they are marginally significant for reading. A lower ranking on PISA in science (Model 3) and reading (Model 7) is associated with a significant decrease in confidence in educational system. Also, respondents in top ten countries show the highest level of confidence in educational system (Models 4 and 8).

In summary, the HLM models show clear pattern between country performance on PISA and public confidence in national education systems. Respondents in high performing countries—measured in multiple ways—express higher levels of confidence in their educational system than countries with lower scores and rankings. All models account for alternative explanations by including controls at the individual-level and country-level.

Supplemental Analyses

We conducted several methodological checks on our results (available upon request). First, we used multilevel mixed-effects ordinal logistic regression (with melogit command in Stata) to ensure that our findings were not affected by the categorical nature of the outcome variable

(Long, 1997). Patterns were consistent across all models, suggesting that our results are not an artifact of using HLM technique as opposed to a categorical technique.

Second, we introduced and additional control—confidence in health care system—to test the socio-psychological explanation (Newton & Norris, 2000) for formation of trust in public institutions. Here, we use the additional control variable as a proxy for trusting disposition or personality. The global bivariate correlation between confidence in health care system and confidence in educational system is positive and significant ($r=.55$, $p<.001$).[6] The link between performance in PISA and confidence in educational system holds even after we account for confidence in health care system. Across all models, the coefficient for confidence in health care system was positive and significant.

Finally, we estimated the models without Asian countries because previous research suggests that these countries often show low levels of confidence in institutions and high performance in PISA. In models with performance in mathematics, patterns were consistent in terms of coefficient direction and significance. In models with performance in science and reading, patterns were consistent in terms of coefficient direction but not in terms of significance.

Discussion

The past two decades witnessed an immense growth in ILSA, especially in subjects such as mathematics, science, and reading. In addition to the growth in the sheer number of assessments, scholars have documented a growing number of countries that subject their educational systems to the scrutiny of Trends in International Mathematics and Science Study (TIMSS), Progress in

[6] More than half of respondents (56.6 percent) reported the same level of confidence for both public institutions. Almost equal share of respondents reported having more confidence in educational system (22.7 percent) or more confidence in health care system (20.7 percent). The strength of the correlation vary across countries, with Russian Federation having the strongest correlation ($r=.73$, $p<.001$) and Switzerland having the weakest correlation ($r=.36$, $p<.001$).

International Reading Literacy Study (PIRLS), and PISA. For example, 55 countries participated in the most recent cycle of TIMSS and 76 countries participated in the most recent cycle of PISA (both cycles were completed in 2015). Once released, results from ILSA often appear on the front pages of newspapers around the world, often in a form of an international league table. In many countries, results from ILSA ignite a public discussion about weaknesses, strengths, and opportunities in the educational system.

In response to the emergence of the ILSA phenomenon, scholars have examined the consequences of these assessments for education policy. Drawing on news stories and policy documents, scholars demonstrated the ways in which policy makers and pundits (i.e., media outlets) engage with and respond to results of TIMSS, PIRLS, and PISA. In this article, we posit that in order to fully understand the political consequences of ILSA, scholars should also consider the perspective of the general public. This belief directed us to the main research question that guided this study: is there a relationship between national performance on PISA and public confidence in education?

The theoretical framework for this study is informed by two lines of research. The first line examined public confidence in education. This line of work, which is based on samples from the United States and Canada, identified important individual-level control variables. The second line used Newton and Norris' (2000) institutional performance model coupled with James and Moseley's (2014) nascent institutional performance theory. Both expounded the hypothesis that people gain or lose confidence in institutions based on their performance as opposed to other factors. Using these hypotheses to frame our study we argued that respondents in high preforming countries will express higher level of trust in their respective educational system.

In order to test this hypothesis, we used high quality survey data from the health and health care module of the ISSP, which was conducted in 2011 in 30 countries. The ISSP dataset is hierarchically organized with individuals nested within countries, thus we used HLM technique. For each PISA domain (mathematics, science, and reading), we examined a series of models with different forms of student performance in PISA.

We conclude that there is initial evidence that national performance on PISA has a significant positive relationship to public confidence in education. Respondents in countries that perform well on PISA—whether measured in country average score, country average score relative to OECD average, ranking, and grouping by ranking (top ten, 11-20, and 21 and below)—show higher levels of confidence in education. While results are robust for national performance in mathematics (Table 3), the results for performance in science and reading are somewhat weaker (Table 4). All the coefficients are in the right direction, but not all of them reach statistical significance.

One possible explanation for this pattern is to consider the results of PISA, specifically low or disappointing results, as disrupting the "taken for granted" aspects of an educational system. Indeed, scholars have used the term "shock" to describe policy reactions to PISA results (Lingard, 2016). We speculate that in low preforming countries, the media and policy makers have more opportunity to interpret the quantitative results and to scandalize the educational system (Steiner-Khamsi, 2003). In turn, the exposure to this kind of public discourse raises questions on the assumed or believed quality of education in a given country. Further, ILSA results point to what is "educationally possible" in other contexts and may shape expectations by policy makers and the general public.

The implications of these results are of importance to scholarship on education policy. They suggest that policy makers may use ILSA results to "shake" the educational system up by manufacturing a sense of crisis among the general public (Berliner & Biddle, 1996; Lingard, 2016). In turn, this "shake up" will create a policy window for introducing solutions for low performance in ILSA (e.g., privatization, school choice, new teacher training and evaluation) and for the decreasing confidence in education and school (e.g., test-based accountability, national evaluation systems).

Following previous research on public confidence in education, and with accordance with socio-psychological and socio-cultural models (Newton & Norris, 2000) for trust formation, we find significant relationships for key individual-level predictors. Women and parents show higher level of confidence in education across 30 countries. Research from the US General Social Survey suggests that men show higher level of confidence in most social institutions and the government (Smith, 2008). One exception is education where there is no significant difference. As for parents, the pattern might be explained by their regular contact with the educational system. Respondents with higher levels of education and respondents living in urban communities show lower level of confidence in education. Importantly, individual-level predictors account for less than 1% of the variation in confidence in education between countries.

Every study has limitations and our study is no different. Three limitations should be considered. First, the ISSP data do not include any information about media consumption habits. Nor does our country-level data include any information about the volume of public discourse (e.g., number of news articles published following the release of PISA 2010). Because educational attainment is often associated with media consumption and political engagement, we

believe that by controlling for education we address this issue to some extent. Still, future research should investigate the link between public discourse regarding ILSA (e.g., volume and content) and public opinion (for example, see: Fladmoe, 2012). Second, our study is based on survey data collected after more than a decade of ILSA results have been infused to the public discourse. Most of the countries included in the analysis began participating in ILSA in the mid-1990s (as part of the 1995 Third International Mathematics and Science Study) and in the early cycles of PISA. It is likely that ILSA results had significant effect on trust in education in the early 2000s. In other words, it is possible that measures of trust in 2011 are shaped by ILSA results published beforehand. Future research could address this limitation by using survey data from earlier periods. Third, our study is based on one time-point (2011). Therefore, the design precludes any interpretation regarding the causal effect of change in ILSA results (e.g., improving, declining, or no change) on transformation in trust in education. It is noteworthy that ILSA results from the past two decades are consistent and show little change in country performance; thus, future research should focus on countries that demonstrated change in their ILSA results. One related limitation is the issue of reverse causality. Although trust in education is specified as the outcome variable, one could argue that trust in education explain cross-national variation in achievement (see for example: Bryk and Schneider, 2002). Future research could address this limitation using time-series analysis or population-based experimental surveys (Mutz, 2011).

The relationship between performance data—such as PISA—and public opinion has many levels of complexity that need to be considered. Future research in this area could develop in three ways. Quantitative research should explore changes in public opinion over time, before and after engagement with ILSA. Other quantitative research should extend beyond trust in

education to include policy preferences, such attitudes toward public spending on education or introduction of policies inspired by neoliberal and New Public Management approaches. Qualitative research should examine the ways in which the general public engages with and makes sense of ILSA results (e.g., what does the public discourse about these assessments mean to the general public?). Other qualitative research should investigate how people engage and consume different types of presentations of performance data (e.g., ranking tables versus grouping of countries, and quality measures versus equity measures).

In conclusion, our study provides evidence, for the first-time to our knowledge, that PISA results have a significant relationship to public trust in education. It contributes to the growing scholarly literature about ILSA by including often neglected stakeholders—members of the general public—who are not only the sponsors of ILSA, but also the benefactors of this policy tool. By doing so, this study suggests further benefits of bringing the public back to educational research in general and to international and comparative education specifically.

References

Bali, V. A. (2016). Evolving Trends in Public Opinion on the Quality of Local Schools. *Educational Policy*, *30*(5), 688-720.

Berkman, M. B., & Plutzer, E. (2005). *Ten thousand democracies: Politics and public opinion in America's school districts*. Georgetown University Press.

Berliner, D. C., & Biddle, B. J. (1996). The manufactured crisis: Myths, fraud, and the attack on America's public schools. *NASSP Bulletin*, *80*(576), 119-121.

Bieber, T., & Martens, K. (2011). The OECD PISA study as a soft power in education? Lessons from Switzerland and the US. *European Journal of Education*, 46(1), 101-116.

Brooks, C., & Manza, J. (2006). Social policy responsiveness in developed democracies. *American Sociological Review*, 71(3), 474-494.

Bryk, A., & Schneider, B. (2002). Trust in schools: A core resource for improvement. Russell Sage Foundation.

Burstein, P. (1998). Bringing the public back in: should sociologists consider the impact of public opinion on public policy? *Social Forces*, 77(1), 27-62.

Chanley, V. A., Rudolph, T. J., & Rahn, W. M. (2000). The origins and consequences of public trust in government: A time series analysis. *Public opinion quarterly*, *64*(3), 239-256.

Clawson, R. A., & Oxley, Z. M. (2012). Public opinion: Democratic ideals, democratic practice. London: Sage Publications.

Coleman, J. (1990). Foundations of social theory. *Cambridge, Mass.: Belknap Press.*

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*, S95-S120.

Delhey, J., & Newton, K. (2005). Predicting cross-national levels of social trust: global pattern or Nordic exceptionalism? *European Sociological Review*, *21*(4), 311-327.

Dixon, R., Arndt, C., Mullers, M., Vakkuri, J., Engblom-Pelkkala, K., & Hood, C. (2013). A lever for improvement or a magnet for blame? Press and political responses to international educational rankings in four EU countries. *Public Administration*, 91(2), 484-505.

Dobbins, M., & Martens, K. (2012). Towards an education approach à la finlandaise? French education policy after PISA. *Journal of Education Policy*, 27(1), 23-43.

Figazzolo, L. (2009). Impact of PISA 2006 on the education policy debate. *Education International*.

Fladmoe, A. (2012). The nature of public opinion on education in Norway, Sweden and Finland–measuring the degree of political polarization at the mass level. *Scandinavian Journal of Educational Research*, *56*(5), 457-479.

Freitag, M., & Bühlmann, M. (2009). Crafting trust: The role of political institutions in a comparative perspective. *Comparative Political Studies*, *42*(12), 1537-1566.

Gendall, P. (2013). International Social Survey Programme, Study Monitoring 2011, Health. Dunedin: University of Otago.

Grek, S. (2009). Governing by numbers: The PISA 'effect'in Europe. *Journal of Education Policy*, *24*(1), 23-37.

Guppy, N., & Davies, S. (1999). Understanding Canadians' declining confidence in public education. *Canadian Journal of Education/Revue canadienne de l'education*, 265-280.

Helgesen, G. (2006). Towards a More Solid East-West Understanding. *Politics, Culture and Self: East Asian and North European Attitudes*, (40), 254-272.

Howie, S., & Plomp, T. (2005). International comparative studies of education and large-scale change. In International handbook of educational policy (pp. 75-99). Springer Netherlands. Retrieved from https://www.nfer.ac.uk/publications/SEF01/SEF01.pdf

Huff, L., & Kelley, L. (2003). Levels of organizational trust in individualist versus collectivist societies: A seven-nation study. *Organization Science*, *14*(1), 81-90.

International Social Survey Programme (ISSP), (2017). Official Website. Retrieved from http://www.issp.org/about-issp/

Jacobsen, R. (2009). The voice of the people in education policy. In: Sykes, G., Schneider, B., & Plank, D. N. (Eds.). (2009). *Handbook of education policy research*. Routledge.

Jacobsen, R., Saultz, A., & Snyder, J. W. (2013). When accountability strategies collide: Do policy changes that raise accountability standards also erode public satisfaction?. *Educational Policy*, *27*(2), 360-389.

James, O. (2010). Performance measures and democracy: Information effects on citizens in field and laboratory experiments. *Journal of Public Administration Research and Theory*, *21*(3), 399-418.

James, O. (2011). Managing citizens' expectations of public service performance: evidence from observation and experimentation in local government. *Public Administration*,*89*(4), 1419-1435.

James, O., & John, P. (2007). Public management at the ballot box: Performance information and electoral support for incumbent English local governments. *Journal of Public Administration Research and Theory*, *17*(4), 567-580.

James, O., & Moseley, A. (2014). Does performance information about public services affect citizens' perceptions, satisfaction, and voice behavior? Field experiments with absolute and relative performance information. *Public Administration*, *92*(2), 493-511.

Klugman, J., & Xu, J. (2008). Racial differences in public confidence in education: 1974–2002. *Social Science Quarterly*, *89*(1), 155-176.

Lingard, B. (2016). Rationales for and reception of the OECD's PISA. *Educação & Sociedade*, 37(136), 609-627. [English version]

Lipset, S. M., & Schneider, W. (1987). The confidence gap during the Reagan years, 1981-1987. *Political Science Quarterly*, 1-23.

Long, S.J. (1997). Regression models for categorical and limited dependent variables. Advanced quantitative techniques in the social sciences. SAGE Publications.

Loveless, T. (1997). The structure of public confidence in education. *American Journal of Education*, 127-159.

Lyons, W. E., & Lowery, D. (1986). The organization of political space and citizen responses to dissatisfaction in urban communities: An integrative model. *The Journal of Politics*, *48*(02), 321-346.

Meyer, H. D., & Benavot, A. (Eds.). (2013). *PISA, power, and policy: The emergence of global educational governance*. Symposium Books.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340-363.

Morgan, S. L., & Poppe, E. S. T. (2012). The consequences of international comparisons for public support of K–12 education evidence from a national survey experiment. *Educational Researcher*, *41*(7), 262-268.

Mutz, D. C. (2011). *Population-based survey experiments*. Princeton University Press.

Newton, K., & Norris, P. (2000). Confidence in public institutions. In: Pharr, S., & Putnam, R.D. (Eds.). (2000). *Disaffected democracies. What's troubling the trilateral countries*. Princeton, NJ: Princeton University Press.

Organisation for Economic Co-operation and Development (OECD) (1999). Measuring Student Knowledge and Skills: A New Framework for Assessment. Paris: OECD Publications Services. Retrieved from https://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/33693997.pdf

Organisation for Economic Co-operation and Development (OECD). (2010). *PISA 2009 results: What students know and can do: student performance in reading, mathematics and science (Volume I)*. OECD, Paris, France.

Paxton, P. (2007). Association memberships and generalized trust: A multilevel model across 31 countries. Social Forces, 86(1), 47-76.

Perrin, A. J., & McFarland, K. (2011). Social theory and public opinion. *Annual Review of Sociology*, 37, 87-107.

Pizmony-Levy, O. (2017a). Big Comparisons, Little Knowledge: Public Engagement with PISA in the United States and Israel. In Wiseman, A. and Stevens Taylor, C. (Eds.). *The OECD's Impact on Education Worldwide (XX-XX125-156)*. Emerald Group Publishing.

Pizmony-Levy, O. (2017b). Media coverage of PISA 2012: Evidence from pilot study in 20 countries. Research report.

Pons, X. (2011). What do we really learn from PISA? The sociology of its reception in three European countries (2001–2008) 1. *European Journal of Education*, *46*(4), 540-548.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Advanced Quantitative Techniques in the Social Sciences. SAGE Publications.

Sabatier, P. & Jenkins--Smith, H.C. (1999). The advocacy coalition framework: An Assessment. P. 117-168. In: P.A. Sabatier (Ed.) *Theories of the Policy Process*. Westview Press.

Schleicher, A. (2006). Chapter 14: How can international organizations work with the media to manage the results of cross-national studies? In: *Cross-National Studies of the Quality of Education: Planning Their Design and Managing Their Impact (*pp. 265-275). International Institute for Educational Planning (IIEP) UNESCO.

Smith, T.W. (2008). Trends in Confidence in Institutions, 1973-2006. General Social Survey Social Change Report No. 54. National Opinion Research Center, University of Chicago.

Smith, T. (2009). The ISSP: History, Organization and Members, Working Principles and Outcomes: An Historical-Sociological Account. In (Haller, M., et al, Eds.), *The International Social Survey Programme 1984–2009* (pp. 2–27). Milton Park: Routledge.

Soroka, S. N. (2006). Good news and bad news: Asymmetric responses to economic information. *Journal of Politics*, *68*(2), 372-385.

Stack, M. (2007). Representing school success and failure: Media coverage of international tests. *Policy Futures in Education*, 5(1), 100-110.

Steiner-Khamsi, G. (2003). The politics of league tables. *Journal of Social Science Education*, 2(1): 1-5.

Steiner-Khamsi, G. (2010). The politics and economics of comparison. *Comparative Education Review*, 54(3), 323-342.

Takayama, K. (2007). A nation at risk crosses the Pacific: Transnational borrowing of the US crisis discourse in the debate on education reform in Japan. *Comparative Education Review*, 51(4), 423-446.

Takayama, K. (2008). The politics of international league tables: PISA in Japan's achievement crisis debate. *Comparative Education*, 44(4), 387-407.

Takayama, K. (2010). Politics of externalization in reflexive times: Reinventing Japanese education reform discourses through "Finnish PISA success". *Comparative Education Review*, 54(1), 51-75.

Tymms, P. (2004). Effect sizes in multilevel models. In: Schagen, I. & Elliot, K. (Eds.), *But what does it mean? The use of effect sizes in educational research* (pp. 55-66). National Foundation for Educational Research. Retrieved from https://www.nfer.ac.uk/publications/SEF01/SEF01.pdf

Wagemaker, H. (2004). IEA: International studies, impact and transition. Proceedings of the IEA International Research Conference. Nicosia, Cyprus.

Weiler, H. N. (1983). Education, public confidence and the legitimacy of the modern state: Is there a crisis' somewhere?. *Journal of Curriculum Studies*, *15*(2), 125-142.

Wiseman, A. W. (Ed.). (2010). *The impact of international achievement studies on national education policymaking*. Emerald Group Publishing.

Table 1: Metrics, Definitions, and Descriptive Stats (n=42,331)

Variable	Definition and description	Mean	SD
Dependent variable			
Confidence in educational system	In general, how much confidence do you have in the educational system in [country]? Reversed coded: 1=no confidence at all, 2=very little confidence, 3=some confidence, 4=a great deal of confidence, 5=complete confidence	3.21	.93
Individual-level variables			
Female	1=female, 0=male	.54	
Age	Years	47.26	16.69
Age squared	Years	2512.05	1650.68
Parent (yes)	1=at least one child in the household, 0=no children in the household	.27	
Education			
Less than secondary school	1=yes, 0=no	.33	
Secondary school	1=yes, 0=no	.38	
University or higher	1=yes, 0=no	.29	
Income	Relative income based on country-specific z-scores	.04	1.04
Employment status			
Employed	1=yes, 0=no	.58	
Unemployed	1=yes, 0=no	.06	
Not in labor force	1=yes, 0=no	.36	
Community			
Urban	1=yes, 0=no	.27	
Suburb / town	1=yes, 0=no	.40	
Rural	1=yes, 0=no	.32	
Country-level variables			
GDP per capita	Gross Domestic Product (GDP) per capita, logged	10.23	.79
Asian countries	1=yes [China, Chinese Taipei, Japan, and Korea], 0=no	.18	

Variable	Definition and description	Mean	SD
Math score	Aggregate score in PISA 2009 mathematics	506.26	41.17
Math score compared to OECD average			
Statistically above average	1=yes, 0=no	.45	
No difference	1=yes, 0=no	.23	
Statistically below average	1=yes, 0=no	.32	
Math ranking	Country ranking in PISA 2009 mathematics	21.85	13.35
Math simplified ranking			
Rank: Top 10	1=yes, 0=no	.23	
Rank: 11-20	1=yes, 0=no	.22	
Rank: 21 and below	1=yes, 0=no	.55	
Science score	Aggregate score in PISA 2009 science	508.71	31.28
Science score compared to OECD average			
Statistically above average	1=yes, 0=no	.49	
No difference	1=yes, 0=no	.20	
Statistically below average	1=yes, 0=no	.31	
Science ranking	Country ranking in PISA 2009 science	21.85	12.97
Science simplified ranking			
Rank: Top 10	1=yes, 0=no	.22	
Rank: 11-20	1=yes, 0=no	.21	
Rank: 21 and below	1=yes, 0=no	.57	
Reading score	Aggregate score in PISA 2009 reading	504.47	32.94
Reading score compared to OECD average			
Statistically above average	1=yes, 0=no	.40	
No difference	1=yes, 0=no	.28	
Statistically below average	1=yes, 0=no	.32	
Reading ranking	Country ranking in PISA 2009 reading	20.99	13.39
Reading simplified ranking			
Rank: Top 10	1=yes, 0=no	.25	
Rank: 11-20	1=yes, 0=no	.24	
Rank: 21 and below	1=yes, 0=no	.51	

Table 2: Confidence in Educational System and Performance in PISA

Country	N	Confidence in education		Score			Ranking		
		Mean	SD	Math	Science	Reading	Math	Science	Reading
Australia	1,624	3.09	.81	514.34	527.27	514.90	15	10	9
Belgium	2,563	3.56	.75	515.27	506.58	505.95	14	21	11
Bulgaria	894	2.98	1.10	428.07	439.29	429.08	46	46	46
Chile	1,287	2.59	.93	421.06	447.47	449.37	49	44	44
China	4,045	3.60	.91	600.08	574.62	555.83	1	1	1
Chinese Taipei	1,024	2.52	1.00	543.18	520.42	495.24	5	12	23
Croatia	1,087	3.16	.89	459.94	486.36	475.75	40	37	36
Czech Republic	1,592	3.23	.96	492.81	500.50	478.19	27	24	34
Denmark	1,288	3.65	.68	503.28	499.34	494.92	19	26	24
Finland	1,238	3.93	.65	540.50	554.08	535.88	6	2	3
France	2,712	2.99	.70	496.78	498.23	495.62	22	27	22
Germany	1,567	3.34	.84	512.78	520.41	497.31	16	13	20
Israel	1,059	2.97	1.04	487.14	507.98	495.64	32	20	21
Italy	1,006	2.90	.83	482.91	488.83	486.05	35	35	29
Japan	1,108	2.65	.72	528.99	539.43	519.86	9	5	8
Korea	1,306	2.88	.92	546.23	537.99	539.27	4	6	2
Lithuania	1,035	2.91	.83	476.60	491.41	468.44	37	33	40
Netherlands	1,283	3.21	.74	525.84	522.22	508.40	11	11	10
Norway	1,595	3.71	.68	497.96	499.88	503.23	21	25	12
Poland	1,014	2.89	.82	494.80	508.07	500.48	25	19	15
Portugal	906	2.91	.85	486.89	492.95	489.33	33	32	27
Russian Federation	1,281	2.97	1.12	467.81	478.30	459.40	38	39	43
Slovak Republic	1,072	3.54	.95	496.68	490.27	477.44	23	34	35
Slovenia	957	3.44	.83	501.47	511.76	483.08	20	17	31
Spain	2,169	3.14	.96	483.49	488.25	481.04	34	36	33
Sweden	920	3.30	.80	494.24	495.11	497.45	26	29	19
Switzerland	1,136	3.59	.74	533.96	516.57	500.50	8	15	14
Turkey	1,308	3.26	1.21	445.45	453.91	464.19	43	43	41
United Kingdom	815	3.06	.77	492.41	513.71	494.18	28	16	25
United States	1,440	3.05	.82	487.40	502.00	499.83	31	23	17
Total	42,331	3.21	.93	-	-	-	-	-	-

Table 3: Hierarchical Linear Models of Confidence in Educational System, by Country Performance in PISA in Mathematics (n=42,331)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Individual-level variables						
Female	.028**	.028**	.028**	.028**	.028**	.028**
	(.009)	(.009)	(.009)	(.009)	(.009)	(.009)
Age	-.009***	-.009***	-.009***	-.009***	-.009***	-.009***
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Age squared	.000***	.000***	.000***	.000***	.000***	.000***
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Parent (yes)	.081***	.081***	.081***	.081***	.081***	.081***
	(.010)	(.010)	(.010)	(.010)	(.010)	(.010)
Education						
Secondary education	-.095***	-.095***	-.095***	-.095***	-.095***	-.095***
	(.011)	(.011)	(.011)	(.011)	(.011)	(.011)
University degree or higher	-.052***	-.052***	-.052***	-.052***	-.052***	-.052***
	(.012)	(.012)	(.012)	(.012)	(.012)	(.012)
Income	.004	.004	.004	.004	.004	.004
	(.004)	(.004)	(.004)	(.004)	(.004)	(.004)
Employment status						
Unemployed	-.080***	-.080***	-.080***	-.079***	-.079***	-.080***
	(.019)	(.019)	(.019)	(.019)	(.019)	(.019)
Out of labor force	-.037***	-.037***	-.037**	-.037***	-.037***	-.037***
	(.011)	(.011)	(.011)	(.011)	(.011)	(.011)
Community						
Suburb/town	.038***	.038***	.038***	.038***	.038***	.038***
	(.011)	(.011)	(.011)	(.011)	(.011)	(.011)
Rural	.090***	.090***	.090***	.090***	.089***	.090***
	(.012)	(.012)	(.012)	(.012)	(.012)	(.012)
Country-level variables						
GDP per capita		.100	-.074	-.073	-.159*	-.030
		(.082)	(.070)	(.081)	(.078)	(.078)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Asian countries		-.252	-.880***	-.616***	-.927***	-.870***
		(.171)	(.178)	(.172)	(.180)	(.244)
Math score			.009***			
			(.002)			
Math score compared to OECD average						
No difference				-.263+		
				(.137)		
Below				-.552***		
				(.143)		
Math ranking					-.026***	
					(.005)	
Simplified ranking						
Rank 11-20						-.389+
						(.216)
Rank 21 and below						-.699***
						(.211)
Random components						
Level-1 variance	.866	.863	.863	.863	.863	.863
Level-2 variance	.338	.335	.230	.227	.261	.255
Intraclass correlation coefficient (ICC)	.131	.116	.066	.064	.083	.080
Percentage of change in ICC compared with the null model	.8%	12.1%	50.0%	51.5%	37.1%	39.4%

Note: Standard errors in parentheses. Constants are not shown. Reference groups for: Education – less than secondary school, employment status – employed, community – urban.
+ p<0.10, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Hierarchical Linear Models of Confidence in Educational System, by Country Performance in PISA in Science and Reading (n=42,331)

	PISA Performance in Science				PISA Performance in Reading			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Score	.008***				.010***			
	(.002)				(.003)			
Score compared to OECD average								
No difference		.029				-.306*		
		(.165)				(.136)		
Below		-.239				-.303+		
		(.151)				(.170)		
Ranking			-.015**				-.017**	
			(.006)				(.006)	
Simplified ranking								
Rank 11-20				-.370+				-.094
				(.195)				(.186)
Rank 21 and Below				-.356+				-.343*
				(.199)				(.166)
Random components								
Level-1 variance	.863	.863	.863	.863	.863	.863	.863	.863
Level-2 variance	.264	.299	.283	.295	.258	.287	.272	.287
Intraclass correlation coefficient (ICC)	.086	.107	.097	.104	.082	.099	.090	.099
Percentage of change in ICC compared with the null model	34.8%	18.9%	26.5%	21.5%	37.9%	25.0%	31.8%	25.0%

Note: Standard errors in parentheses. Constants are not shown. All models include controls for: Gender, age, age squared, parental status, education, income, employment status, community, GDP per capita, and Asian countries. Reference groups for: Education – less than secondary school, employment status – employed, community – urban.
+ p<0.10, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$