The International Education Journal: Comparative Perspectives Vol. 16, No. 2, 2017, pp. 16-32
https://openjournals.lirary.sydney.edu.au/index.php/IEJ

Achievement flourishes in larger classes: Secondary school students in most countries achieved better literacy in larger classes

Abeer A. Alharbi
School of Education, University of Glasgow, UK:
abeer.alharbi@glasgow.ac.uk
College of Education, King Saud University, SA:
alhabeer@ksu.edu.sa
Gijsbert Stoet
School of Social Science, Leeds Beckett University, UK:
G.Stoet@leedsbeckett.ac.uk

There is no consensus among academics about whether children benefit from smaller classes. We analysed the data from the 2012 Programme for International Student Assessment (PISA) to test if smaller classes lead to higher performance. Advantages of using this data set are not only its size (478,120 15-year old students in 63 nations) and representativeness but also that the 2012 PISA data set, for the first time, includes the class size for each participating child. We found that, in most countries, children in smaller classes had a lower performance score in solving reading comprehension problems than those in larger classes. We further analysed the relationship between class size and factors that can explain this paradoxical phenomenon. Although grouping of students by ability and the socioeconomic status of parents played some role in some countries, these factors cannot fully explain the effect. We finish by discussing the overlooked potential advantages of larger classes.

Keywords: class size; Programme for International Student Assessment (PISA); Literacy; Larger Classes.

INTRODUCTION

Children's school performance is of critical importance for economies because school performance is directly related to a nation's opportunity to satisfy the need for skilled workers. It is because of this link between education and economic output that the Organisation for Economic Cooperation and Development (OECD) funds the Programme for International Student Assessment (PISA). One of the aims of PISA is to help countries to understand which factors contribute to successful educational systems, and, indeed, since the first PISA reports were published in 2000, numerous educational reforms have been inspired by these surveys (Ertl, 2006; Grek, 2009). It is argued that a 25 point increase in PISA scores for all OECD countries result in a 115 trillion US dollar GDP across nations (OECD, 2006, p. 27). Of all the factors that influence educational outcomes, class size is a much discussed factor. There is strong political and public support for the reduction of class size. In 1998, US President Clinton proposed a large initiative to reduce class sizes in primary schools and was quoted as saying: “When class sizes go down enough, learning goes up” (Broder, 1998). This initiative received US$1.2 billion to hire 100,000 teachers and to reduce class sizes in primary schools down to an average of 18 children (The White House, 2000). Nye, Hedges and Konstantopoulos (1999) noted that many states adopted policies to reduce classroom sizes. Chingos (2013) cited a 2007 representative public US survey which found that 77% of respondents (including teachers) supported the allocation of money to reduce class sizes. Also, 81% of public school employees wanted to improve work conditions that included reduction in class sizes rather than an increase in salaries.

There have been numerous academic studies addressing the question whether or not smaller classes are actually beneficial to educational performance, as well as studies of directly related questions, such as the associated costs of class size reductions; the latter is important because teacher salaries are a large part of a school’s expenditure (Clowes, 2004; OECD, 2013) and there is a teacher scarcity in a number of school subjects (UNESCO, 2013a, 2013b; Voke, 2003). In other words, even if a reduced class size would be beneficial in terms of performance improvements, it might not be value for money (Brewer, Krop, Gill, & Reichardt, 1999; Buckingham, 2003; Chingos, 2013), or it might, in fact, be impossible to provide the needed financial resources or teachers.

It is worth noting that the literature also addressed the interaction between class-size reduction and other factors. For example, Hattie (2005) and Harfitt (2015) argued that smaller classes are more beneficial in Western cultures where autonomy is valued, whereas larger classes are better for Eastern cultures that appreciate collectiveness. In East Asia, collectivist culture is shared among the Confucian heritage cultures (Oyserman, Coon, & Kemmelmeier, 2002; Phuong-Mai, Terlouw, & Pilot, 2005; Wursten & Jacobs, 2013; Yang, 1993). Various studies provided evidence that the positive impact of small classes is larger for children with difficulties and additional support needs (e.g., Bosworth, 2014; Ecalle, Magnan, & Gibert, 2006; Hanushek, 2002; Krassel & Heinesen, 2014; Molnar et. al., 1999; Mosteller, 1995; Mosteller, Light, & Sachs, 1996). However, such conclusions were contested by Cho, Glewwe, and Whitler (2012), who argue that smaller classes impact all children equally. Nonetheless, it might be the case that the benefit of class size depends on culture, which means that an international comparison of class size is important to study. Such an analysis can possibly tell for which countries it would make sense to invest in smaller classes.

Of course, the discussions about the value for money of class-size reduction only make sense if class-size reduction actually has a beneficial effect. Although there have been numerous studies about the effects of class size reduction, there is a lack of consensus. Both positive (Bascia, 2010; Breton, 2014; Cho et. al., 2012; Finn & Achilles, 1999: Finn, Gerber, & Boyd-Zaharias, 2005; Fredriksson, Öckert, & Oosterbeek, 2013; Jakubowski & Sakowski, 2006; Jepsen & Rivkin, 2009; Krueger, 2000; Nye et. al., 1999; Tienken & Achilles, 2006) and negative (Dobbelsteen, Levin, & Oosterbeek, 2002; Maasoumi, Millimet, & Rangaprasad, 2003; Morris, 1959 cited in OECD, 1974) associations between school performance and class sizes have been reported. Most of these studies are correlational with only a few experimental. Although experimental studies have theoretical advantages (e.g., they allow the determination of cause and effect), they have not lead to consensus. A good example of the lack of consensus is around one of the most famous experimental studies of class size reduction, namely the Student-Teacher Achievement Ratio (STAR) Project carried out in the 1980s in Tennessee, US. Based on the project's data, some researchers concluded there were benefits to smaller classes (Finn & Achilles, 1999; Finn et. al. 2005; Finn, Gerber, Achilles, & Boyd-Zaharias, 2001; Krueger, 1999; Nye, Hedges, & Konstantopoulos, 2000; Mosteller, 1995; Nye et. al., 1999) while others made a convincing argument that these data do not support such conclusions (Hanushek, 1997; 1999; 2002). A meta-analysis of studies about school resources (including class size) in both primary and secondary education concluded that smaller classes and schools are positively related to academic achievement in mathematics and reading (e.g., Greenwald, Hedges & Laine, 1996; Hedges & Stock, 1983). Some studies have concluded that there are also long-term and non-academic positive outcomes of smaller classes (e.g., Chetty et. al., 2011; Dee & West, 2011).

The lack of consensus about the benefits of smaller classes can partially be explained by confounding factors. It has been argued that the benefits disappear when other factors are controlled for (Cho et.al., 2012; Ehrenberg, Brewer, Gamoran, & Willms, 2001a; Hoxby, 2000; Wößmann, 2005; Wößmann, 2003b). Also, that the gains of class-size reduction could be achieved equally (if not better) by other factors, such as parental involvement and other family factors (Browning & Heinesen, 2007; Coleman et. al., 1966; Funkhouser, 2009; Nascimento, 2008; Wößmann, 2005) or institutional factors and school resources, including teachers' factors and teaching practices (Chingos, 2012; Ehrenberg, Brewer, Gamoran, & Willms, 2001b; Finn, Pannozzo, & Achilles, 2003; Fleming, Toutant, & Raptis, 2002; Funkhouser, 2009; Hall, 2012; Hanushek, 1986, 2003; Jackson, Johnson, & Persico, 2014; Harris & Plank, 2001; Jepsen, 2015**;** Jez & Wassmer, 2015; Mueller, 2013; Panizzon, 2015; Stern, 1987; Wößmann & West, 2006; Wößmann, 2003a). Further, methodological factors might be responsible for the differences (Akerhielm, 1995; Buckingham, 2003; Schanzenbach, 2014; Hanushek, 1999; Hoxby, 2000; Krueger, 1999, 2002, 2003; Lewit & Baker, 1997).

Given the lack of consensus about the relationship between school performance and class size, we analysed the 2012 PISA dataset. Since the year 2000, the PISA organisation has published academic performance in 15 year-old school children around the world. The 2012 survey is the latest dataset, which involved nearly half a million children in 65 countries, making it the largest international educational survey. One of the advantages of PISA is that children around the world are tested on the same set of problems. Questions are not only translated into local languages, but great effort has been put into the cross-cultural comparability of the questions asked (OECD, 2014). This design makes it possible to compare performance in different cultures. A specific advantage of the 2012 PISA data set is that it contains, for the first time for each participating student, the class size in the test-language classes (e.g., English class in English-speaking countries). Because this is the first time that the class size of each participating child in a large international educational survey is available, it allows for a more detailed correlational analysis between class size and performance than hitherto possible.

If it is true that smaller classes are beneficial for performance, we expected that children in smaller classes would score higher on the PISA survey of text comprehension tasks. We expected that if the effects of sorting by ability can explain a relationship between class size and performance, this effect should not be observed in children who are in schools which do not base their admission on ability and who do not sort children based on ability within the school.

METHODS

We analysed the raw data of the 2012 PISA data set (available via http://www.oecd.org/pisa) using the statistical software R (R Core Team, 2014). This data set contains the data from 485,490 school children in 68 countries and regions. The age of children participating in PISA ranged from 15 years and 3 months to 16 years and 2 months. In addition to the data of the US as a country, the US states Florida, Massachusetts, and Connecticut also participated separately, but we have excluded those data in order not to count the same country more than once. Similarly, we excluded the separate data from the Russian city Perm, because Russia as a whole participated. Finally, we excluded the data from Liechtenstein, which had too few participants (*n*=293) for a meaningful data analysis of class size and performance. Because the variable class size (PISA variable *ST72Q01*) was not available for all students (1.5%), the analyses involving the variable class size included 478,120 students.

The class size variable *ST72Q01* appeared to have some unrealistic outlier data (ranging from 0 to 200). In order to deal with these outliers, we calculated, for each country, the 5th and 95th percentile of class sizes and only included those data that fell in this range (see Table 1). This method is supported in the literature (e.g., Motulsky, 2014**;** Osborne & Overbay, 2004). We also analysed whether the effect of "streaming" can explain any correlation between class size and performance. Streamed schools are here defined as schools that either always use ability as an admission criterion or always use ability to assign children to classes (or both). Non-streamed schools are here defined as those schools that neither use ability for admission nor group children in classes by ability (note that the PISA data set allows cross-linking of school data with individual children's data, because the student data set has for each child a school identifier).

RESULTS

The range of class sizes varied considerably within and between countries. The low end of class size ranged from 5 in Kazakhstan to 31 in Vietnam, whereas the high end of class size ranged from 24 in Finland to 52 in Taiwan. While East Asia is known for its large classes, it should be pointed out that such large classes are found elsewhere as well (e.g., countries in Latin and South America, Turkey, and Jordan had classes of 40 or higher as well, Table 1). The variability in observed class sizes ranged between countries as well, from Greece ranging from 17 to 28 children per class to Mexico ranging from 15 to 51.

In Table 1, we report, for each country or economic region, the following information: 1) Range of class size (from 5th to 95th percentile of class size). 2) Correlation between class size and reading performance. 3) Correlation between class size and socio-economic status. 4) Correlation between class size and reading performance controlled for socio-economic status (partial correlations). 5) Percentage of children in schools that either always select based on ability or always group students by ability in classes. 6) Correlation between class size and reading performance controlled for socio-economic status only for children in non-streamed schools (calculated only if more than 1000 students in such schools in a country). For each of the 63 countries and economic regions, we calculated the Pearson correlation between class size and the reading comprehension scores. Correlations ranged from *r*=-.02 in the United Arab Emirates to *r*=.51 in France

(Table 1[1]). Except for the United Arab Emirates, Singapore, Jordan, Kazakhstan, and Tunisia, we found statistically significant correlations between school performance and class size in 58 (i.e., 92%) of countries. Thus, we did *not* find a positive relationship between smaller classes and performance in any of the countries.

Table 1: OECD countries, class size, streamed classes, correlations between class size and; performance, performance with controlled SES, and performance with controlled SES for non-streamed classes

Country	Class size range 5%	Class size range 95%	Class size and performance	Class size and SES	Class size and performance controlled for SES	Streamed (%)	Class size and performance
France	14	35	0.51	0.32	0.45	31	0.38
Netherlands	14	30	0.41	0.18	0.4	82	..
Bulgaria	10	29	0.4	0.28	0.33	80	..
Luxembourg	12	27	0.39	0.28	0.34	73	..
Belgium	10	26	0.37	0.25	0.32	28	0.26
Switzerland	10	25	0.36	0.21	0.32	67	0.18
Slovak Republic	11	30	0.33	0.24	0.27	56	..
Portugal	13	28	0.32	0.26	0.26	35	0.18
Hungary	13	36	0.31	0.21	0.26	84	..
Slovenia	12	31	0.3	0.17	0.27	27	-0.01
Lithuania	11	30	0.29	0.24	0.25	50	..
Hong Kong	20	41	0.27	-0.03	0.3	92	..
Israel	10	40	0.27	0.17	0.24	78	..
Latvia (LSS)	8	29	0.27	0.3	0.19	31	..
Montenegro	18	37	0.27	0.15	0.25	52	0.1
Estonia	9	31	0.26	0.27	0.22	47	..
Thailand	18	50	0.26	0.23	0.21	81	..
Shanghai	20	49	0.26	0.16	0.23	54	..
Ireland	12	30	0.25	0.17	0.22	58	..
Serbia	13	35	0.25	0.12	0.24	79	..
Czech Republic	12	30	0.24	0.1	0.23	50	..
Japan	24	43	0.23	0.07	0.23	94	..
Romania	18	33	0.23	0.14	0.2	51	..
Greece	17	28	0.22	0.21	0.18	6	0.17
Vietnam	31	49	0.22	0.11	0.22	88	..
Germany	15	30	0.21	0.15	0.18	52	..
Italy	13	29	0.21	0.13	0.19	54	0.18
New Zealand	12	30	0.21	0.13	0.18	57	..
Peru	14	40	0.21	0.19	0.15	31	0.15
Argentina	20	38	0.2	0.18	0.17	14	0.13

1 Please note that table 1 is mentioned in both sections (the methods and results), since it contains the data and its analysis.

Country	**Class size range 5%**	**Class size range 95%**	**Class size and performance**	**Class size and SES**	**Class size and performance controlled for SES**	**Streamed (%)**	**Class size and performance**
Australia	12	30	0.2	0.1	0.18	59	..
Canada	13	32	0.2	0.08	0.19	41	0.08
Austria	11	30	0.19	0.11	0.16	74	..
Korea	15	40	0.19	0.17	0.15	75	..
United Kingdom	12	30	0.19	0.01	0.21	75	..
Croatia	18	33	0.18	0.1	0.16	96	..
Russian Federation	7	29	0.18	0.23	0.11	19	..
Mexico	15	51	0.18	0.1	0.17	57	0.36
Indonesia	19	42	0.16	0.02	0.18	61	0.22
Finland	10	24	0.15	0.13	0.14	10	0.15
Iceland	7	30	0.14	0.14	0.12	22	..
Poland	15	30	0.14	0.14	0.11	19	0.1
Sweden	11	30	0.13	0.14	0.09	15	0.13
Spain	11	31	0.12	0.1	0.09	8	0.11
Macao	20	46	0.12	0.02	0.13	70	..
Costa Rica	15	36	0.08	0.12	0.05	49	..
Chile	20	45	0.07	-0.04	0.1	35	0.11
Malaysia	11	45	0.06	0.07	0.05	55	..
Norway	12	30	0.05	0.1	0.03	14	0.12
Turkey	7	40	0.05	0.04	0.04	43	0.05
Brazil	15	46	0.05	0.07	0.03	21	0.03
Colombia	19	45	0.04	0.1	0	44	0.18
Chinese Taipei	27	52	0.04	0.04	0.02	47	..
United States	12	35	0.04	0.01	0.04	40	..
Albania	9	39	0.04	..	..	57	..
Uruguay	12	36	0.04	0.08	0.01	26	0.03
Denmark	12	26	0.03	0.04	0.02	10	0.07
Qatar	20	38	0.03	-0.04	0.05	49	0.01
Kazakhstan	5	30	0.01	0.13	-0.05	56	..
Tunisia	15	34	0.01	0.08	-0.02	45	..
Jordan	14	47	-0.01	-0.04	0	34	0.13
Singapore	18	42	-0.01	-0.1	0.03	79	..
United Arab Emirates	12	35	-0.02	-0.13	0.02	81	..

Next, we investigated possible variables that can explain part of this pattern. We tested the effects of streaming and socio-economic status. Both these variables are highly relevant. Streaming means that children are grouped by ability (at class or school level), and it is possible that children with more learning difficulties were assigned to smaller classes. Further, it is well known that socio-economic status is a good predictor of school performance, including in PISA. For example, in the 2012 PISA data set, we found that the relationship between socioeconomic status and reading performance in PISA ranges from r=.12 in Macao to r=.49 in the Slovak Republic.

First, we report the role of socio-economic status. We correlated this variable with class size for each country and found that children from families with a higher socio-economic status sat in larger classes. Interestingly, the relationship between socioeconomic status and class size was similar to the relationship between performance and class size, *r* (60) =.770, *p*<.001 (Table 1). To deal with the possible confounding influence of socioeconomic status on performance, we calculated, for each country, the correlation between class size and performance controlled for socioeconomic status (using partial correlations). Using this calculation, the correlations ranged between *r*=-.046 in Kazakhstan to *r*=.452 in France (Table 1). Thus, the main difference is that with this control of socio-economic status, we found the expected negative relationship between class size and performance in only one country, namely Kazakhstan (albeit extremely weak), whereas, again, the positive effect was found in the large majority of countries (*n*=51 or 81%). It is possible that children were assigned to classes depending on their performance level, for example, because it is assumed that lower performing children need more attention and thus would benefit from a smaller class (Biddle & Berliner, 2002; Blatchford, Bassett, & Brown, 2008; Finn, et. al, 2001; Nye, Hedges, & Konstamtopoulos, 2002; Wilson, 2006) in which teachers have more time per child (Blatchford et. al, 2008). To test to what degree this can explain these data, we analysed the effect of streaming according to ability. The degree to which children were assigned to schools or classes by ability varied considerably between the participating countries. Greece, Spain, Denmark, and Finland have 10% or fewer participating students in streamed schools, whereas Hong Kong, Japan, and Croatia have over 90%. Of interest is that even countries with a generally comprehensive school system (like the UK), most children might be streamed within the school (in the UK, 75% of participating children are streamed by ability, Table 1).

One of the advantages of the large PISA data set is that, for many countries, we have sufficient data to just apply the analysis on children who are neither streamed through school admission or within the school. We tested to what degree class size and performance are related for children in schools that are not streamed at all (i.e., schools that do not admit or sort children based on ability). In some countries, few students were in such a school; to ensure we had sufficient numbers of children in a variety of class sizes, we only included the 26 countries that had at least a country total of at least 1000 participating students in the type of non-streamed schools. In these 26 countries, the correlation between performance and class size in non-streamed schools (while controlled for socioeconomic status using partial correlations) ranged from *r*=-.026 in Brazil to *r*=.436 in France (Table 1). Of these countries, 19 (or 73%) countries again showed a statistically significant positive correlation ranging between *r*=.074 in Qatar and *r*=.436 in France.

DISCUSSION

Our analysis of the 2012 PISA data shows that there is a positive relationship between class size and performance in reading comprehension in the majority of countries. Except for a very small effect in Kazakhstan, we found no countries where there is a clear positive benefit of sitting in a smaller class. Importantly, we found the same effect even when only taking into account children who attend schools that neither select nor group students by ability.

Implications of our findings

The main implication of our data analysis is that there is no strong evidence to believe that smaller classes are beneficial to student attainment (at least, for 15-year old students without special needs). Of course, it leaves open the question of whether performance could be raised by increasing class size. This is a key question, given that educational policy makers might conclude from our results that larger classes directly cause higher scores in, at least, language learning. That conclusion would *not* be warranted because our study is correlational, and correlation does *not* imply causation. In order to answer this question further, the reasons for the positive effects of larger classes need to be better understood. Some authors have suggested ways to improve performance in larger classes (Benbow, Mizrachi, Oliver, & Said-Moshiro, 2007; Blatchford, Goldstein, & Mortimore, 1998; Carter, Cushing, & Kennedy, 2008; Heiney, 2010; Henderson & Buising, 2000; Mintah, 2014). Unfortunately, though, there are not many studies that address how larger classes can be beneficial (Blatchford, Bassett & Brown, 2011; Hattie, 2005). One given reason for the effectiveness of larger classes is that many schools base admissions on ability or stream children by ability within schools (Maasoumi et. al., 2003; Mosteller et. al., 1996; Wößmann, 2003b, 2005). Yet, streaming cannot explain why we observed the effect in the majority of countries when only analysing data from children who are not being sorted by ability at all. Another reason for better performance in larger classes is that they offer more opportunities to learn from peers (Borland, Howsen, & Trawick, 2005 cited in Kornfeld, 2010; Dobbelsteen et. al., 2002). We speculate that there might also be an indirect effect of the approaches needed to deal with larger classes. For example, it might be case that larger classes require a different form of discipline, which might lead to less disruption in class which, in turn, leads to better performance. Such complex hypothetical causal pathways are difficult to prove and require more detailed studies.

Why our findings deviate from studies indicating benefits of smaller classes

In the introduction, we listed theoretical work relevant to the relationship between class size and educational performance and achievement. Most of this work pointed at the positive effects of smaller classes on academic achievement (e.g., Breton, 2014; Cho et. al., 2012; Finn & Achilles, 1999; Finn et. al, 2005; Fredriksson et. al, 2013; Jakubowski & Sakowski, 2006; Jepsen & Rivkin, 2009; Nye et. al., 1999; Tienken & Achilles, 2006). Relevant to our current work focusing on reading skills, it is interesting that positive effects of smaller classes have been reported to be larger in reading (the subject we focused on) than in mathematics (e.g., Camacho, 2006). Our findings raise the question: how is it possible that different studies come to quite different conclusions about the benefits of smaller or larger classes? Answering this question will help to develop a refined understanding of the relationship between class size and academic achievement.

In the following, we will focus on two factors we believe can explain part of the contrast between our study and other work showing benefits of smaller classes. In short, these factors are related to the studied children as well as to how benefits of smaller classes are measured.

The first factor is related to which children have been studied. Our study was exclusively carried out with 15 and 16 year olds. At this age group, children have already developed relatively high reading skills, and children are typically better able to study more independently than is the case at younger ages. Therefore, we believe that it would be unreasonable to extend our findings to primary school children in which benefits of smaller classes have been found (e.g., Finn & Achilles, 1999; Finn et. al. 2005; Finn et. al, 2001; Krueger, 1999; Mosteller, 1995; Nye et. al, 2000; Nye et. al., 1999). Apart from age groups, benefits of smaller classes have been shown for school children with special needs and from low income backgrounds (e.g., Bosworth, 2014; Ecalle, Magnan, & Gibert, 2006; Hanushek, 2002; Krassel & Heinesen, 2014; Molnar et.al., 1999; Mathis, 2016; Mosteller, 1995; Mosteller et.al, 1996; Zyngier, 2014). Again, we believe it would be unreasonable to extend our findings to schools with children with special needs, especially because these children will benefit from smaller classes.

The second factor that explains the difference between our conclusions and those of other studies are related to the outcomes measured. Our study focuses on a test measuring reading comprehension. Some studies analysing the benefits of small classes have focused on other outcomes, including long-term outcomes on college completion and earnings, as well as on non-cognitive skills (e.g., Chatty et. al., 2011; Dee and West, 2011; Harfitt & Tsui, 2015). Given the constraints of our data set, we could not include such variables.

In summary, theoretical advances in understanding the benefits of smaller classes needs to be put into the context on which age groups are studied, whether special needs students are included, and what incomes are being considered.

Limitations of the current study

The main limitation of this study is that our data only apply to 15-year olds. It would be of great interest to carry out the same analysis with children in primary schools. Although the Progress in International Reading Literacy Study (PIRLS) could address this question, it does not collect class size data per child. Further, the PISA class size variable was collected for classes in the host language only, which limits our analyses to the subject of reading comprehension. Although, we cannot generalize our conclusions to other subjects (e.g., mathematics or science literacy), it should be noted that scores in reading comprehension, mathematics, and science literacy are highly correlated (i.e., a child doing well in reading comprehension also does well in the other subjects (Stoet & Geary, 2015), and indeed the class size variable we used is highly correlated with the average class size of the schools participating in PISA. Therefore, we believe that the pattern observed here likely also generalizes to the subjects, mathematics and science.

Another limitation of this study is that our findings, like many large scale educational studies, are correlational in nature (which precludes conclusions about causal pathways). Given the importance of the relationship between class size and cognitive performance, we hope that educational policy makers would be willing to invest in an experimental or longitudinal study, which can answer the causal relationship between class size and cognitive performance.

CONCLUSION

In summary, we found a positive relationship between class size and educational performance in the majority of countries participating in PISA, even when controlling for streaming and socio-economic status. This finding seems incompatible with the idea that class-size reduction can increase attainment, at least for typically developing children around 15 years old. Given the importance of evidence-based educational policies, it is important to better understand the causal relationships between class-size and school achievement using experimental and longitudinal research approaches.

REFERENCES

Akerhielm, K. (1995). Does class size matter? *Economics of Education Review, 14* (3), 229-241. doi:10.1016/0272-7757(95)00004-4

Bascia, N. (2010). *Reducing class size: What do we know*, Ontario, CA: Canadian Education Association. Retrieved from: http://www.classsizematters.org/wp-content/uploads/2012/11/Reducing-Class-Size-What-do-we-Know.pdf

Benbow, J., Mizrachi, A., Oliver, D., & Said-Moshiro, L. (2007). *Large class sizes in the developing world: What do we know and what can we do?* US Agency for International Development Cooperative Agreement, No. GDG-A-00-03-00006-00. Retrieved from: http://www.equip123.net/docs/E1-LargeClassrooms.pdf

Biddle, B. J., & Berliner, D. C. (2002). Small class size and its effects. *Educational Leadership (Class Size, School Size), 59*(5), 12-23. Retrieved from: http://www.ascd.org/publications/educational-leadership/feb02/vol59/num05/toc.aspx

Blatchford, P., Goldstein, H., & Mortimore, P. (1998). Research on class size effects: A critique of methods and a way forward. *International Journal of Educational Research, 29*(8), 691-710. doi:10.1016/S0883-0355(98)00058-5

Blatchford, P., Bassett, P., & Brown, P. (2008). Do low attaining and younger students benefit most from small classes? Results from a systematic observation study of class size effects on pupil classroom engagement and teacher pupil interaction. Paper submitted to symposium: '*Class size effects: new insights into classroom, school and policy processes'*. American Educational Research Association Annual Meeting, New York.

Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction, 21* (6), 715-730. doi:10.1016/j.learninstruc.2011.04.001

Bosworth, R. (2014). Class size, class composition, and the distribution of student achievement. *Education Economics, 22*(2), 141-165. doi:10.1080/09645292.2011.568698

Breton, T. R. (2014). Evidence that class size matters in 4th grade mathematics: An analysis of TIMSS 2007 data for Colombia. *International Journal of Educational Development, 34*, 51–57. doi:10.1016/j.ijedudev.2013.04.003

Brewer, D. J., Krop, C., Gill, B. P., & Reichardt, R. (1999). Estimating the cost of national class size reductions under different policy alternatives. *Educational Evaluation and Policy Analysis,* Special Issue: Class Size: Issues And New Findings, *21*(2), 179-192. doi:10. 3102/01623737021002179

Broder, J. M. (1998). Clinton proposes legislation to achieve smaller class size. *The New York Times*, May 9 1998. Retrieved from: http://www.nytimes.com/1998/05/09/us/clinton- proposes-legislation-to-achieve-smaller-class-size.html

Browning, M., & Heinesen, E. (2007). Class size, teacher hours and educational attainment. *The Scandinavian Journal of Economics*, 109 (2) The Nobel Memorial Prize in Economics, 2006, 415-438. doi: 10.1111/j.1467-9442.2007.00492.

Buckingham, J. (2003). Class size and teacher quality. *Educational Research for Policy And Practice, (1)*, 71–86. doi: 10.1023/A:1024403823803

Camacho, C. M. (2006). *Class size reduction: Is it worth the cost? A meta-analysis of the research.* A non-published EdD dissertation submitted to the Department of Educational Studies at the University of Central Florida, Orlando, Florida

Carter, E. C., Cushing, L. S., & Kennedy, C. H. (2008). *Peer support strategies: improving students' social lives and learning*. Baltimore: Paul H. Brookes Publishing Co.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., & Yagan D. (2011). How does your kindergarten classroom affect your earnings? Evidence from project STAR. *Quarterly Journal of Economics, 126*(4), 1593-1660. doi: 10.1093/qje/qjr041.

Chingos, M. M. (2012). The impact of a universal class-size reduction policy: Evidence from Florida's State-wide mandate. *Economics of Education Review, 31*(5), 543–562. doi:10.10 16/ j.econedurev. 2012.03.002

Chingos, M. M. (2013). Class size and student outcomes: Research and policy implications. *Journal of Policy Analysis and Management, 32*(2), 411–438. doi: 10.1002/pam.21677

Cho, H., Glewwe, P., & Whitler, M. (2012). Do Reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools. *Economics of Education Review, 31*(3), 77–95. doi:10.1016/j.econedurev.2012.01.004

Clowes, G. A. (2004). *Just the facts: Teacher salaries and education spending*. Retrieved from The Hartland Institute: http://news.heartland.org/newspaper-article/2004/05/01/just-facts-teacher-salaries-and-education-spending

Coleman, J. S., Campbell, E.Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: US Government Printing Office.

Dee, T. S., & West, M. R. (2011). The non-cognitive returns to class size. *Educational Evaluation and Policy Analysis, 33*(1), 23–46. doi: 10.3102/0162373710392370

Dobbelsteen, S., Levin, J., & Oosterbeek, H. (2002). The causal effect of class size on scholastic achievement: Distinguishing the pure class size effect from the effect of changes in class composition. *Oxford Bulletin of Economics and Statistics, 64*(1), 17_38. doi: 10.1111/ 1468-0084.00003

Ecalle, J., Magnan, A., & Gibert, F. (2006). Class size effects on literacy skills and literacy interest in first grade: A large-scale investigation. *Journal of School Psychology, 44*(3), 191-209. doi:10.1016/ j.jsp.2006.03.002

Ehrenberg, R. G., Brewer, D. J., Gamoran, A., & Willms, J. D. (2001a). *The class size controversy* (CHERI Working Paper #14). Retrieved from Cornell University, ILR School Site: http://digitalcommons.ilr.cornell.edu/workingpapers/25/

Ehrenberg, R. G., Brewer, D. J., Gamoran, A., & Willms, J. D. (2001b). Class size and student achievement. *Psychological Science in the Public Interest, 2*(1), 1–30. Doi:10.1111/1529-1006.003

Ertl, H. (2006). Educational standards and the changing discourse on education: The reception and consequences of the PISA study in Germany. *Oxford Review of Education, 32*(5), 619–634. doi:10.1080/03054980600976320

Finn, J. D., & Achilles, C. M. (1999). Tennessee's class size study: Findings, implications misconceptions. *Educational Evaluation and Policy Analysis,* special Issue: Class Size: Issues and New Findings, *21*(2), 97–109. doi:10.3102/01623737021002097

Finn, J.D., Gerber, S.B., Achilles, C.M., & Boyd-Zaharias, J. (2001). The enduring effects of small classes. *Teachers College Record, 103*(2), 145–183.

Finn, J. D., Gerber, S. B., & Boyd-Zaharias, J. (2005). Small classes in the early grades, academic achievement, and graduating from high school. *Journal of Educational Psychology, 97* (2), 214–223. doi: 10.1037/0022-0663.97.2.214

Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The “why’s” of class size: Student behaviour in small classes. *Review of Educational Research, 73*(3), 321–368. doi: 10.3102/ 0034 6543073003321

Fleming, T., Toutant, T., & Raptis, H. (2002). *Class size and effects: A review.* Bloomington, Indiana: Phi Delta Kappa Educational Foundation

Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long term effects of class size. *The Quarterly Journal of Economics, 128* (1), 249–285. doi:10.1093/qje/qjs048

Funkhouser, E. (2009). The effect of kindergarten classroom size reduction on second grade student achievement: Evidence from California. *Economics of Education Review, 28*(3), 403–414. doi:10.1016/j.econedurev.2007.06.005

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361–396

Grek, S. (2009). Governing by numbers: The PISA “effect” In Europe. *Journal of Education Policy, 24*(1), 23–37. doi:10.1080/02680930802412669

Hall, L. (2012). The “come and go” syndrome of teachers in remote Indigenous schools: Listening to the perspective of Indigenous teachers about what helps teachers to stay and what makes them go. *The Australian Journal of Indigenous Education, 41*(2), 187–195. doi:10.1017/jie.2012.13.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24*(3), 1141–1177.

Hanushek, E.A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19*(2), 141–164.

Hanushek, E. A. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis,* Special Issue: Class Size: Issues and New Findings, *2* (2), 143–163. doi:10.3102/01623737021002143

Hanushek, E. A. (2002). Evidence, politics, and the class size debate. In L. Mishel, & R. Rothstein (Eds.), *The Class Size Debate* (pp. 37-65) Washington, DC: Economic Policy Institute.

Hanushek, E. A. (2003). The failure of input-based schooling policies. *The Economic Journal, 113*(485): F64–F98. doi: 10.1111/1468-0297.00099

Harfitt, G. J. (2015). *Class size reduction: Key insights from secondary school classrooms.* E-book. Singapore: Springer science, business and media. doi:10.1007/978-981-287-564-8

Harfitt, G. J., & Tsui, A. B. M., (2015). An examination of class size reduction on teaching and learning processes: A theoretical perspective. *British Educational Research Jo*urnal, *41*(5), 845–865. doi: 10.1002/berj.3165

Harris, D., & Plank, D.N. (2001). *Does class size reduction come at the expense of teacher quality?* Policy report no. 4. Michigan: The Education Policy Center at Michigan State University. Retrieved from http://education.msu.edu/epc/forms/Policy-and-research-Reports/REPORT4.PDF

Hattie, J. (2005). The paradox of reducing class size and improving learning outcomes. *International Journal of Educational Research, 43*(6), 387–425. doi:10.1016/ j.ijer. 2006.07.002

Hedges, L. V., & Stock, W. (1983). The effects of class size: An examination of rival hypotheses. *American Educational Research Journal, 20*(1), 63–85. doi:10.3102/ 00028312020001063

Heiney, P. A. (2010). Peer instruction in large classes. *Talk About Teaching and Learning, University of Pennsylvania Almanac*, *56*(22), 8. Retrieved from http://www.upenn.edu/almanac/volumes/v56/n22/tatl.html#sthash.y9CisM4O.dpuf

Henderson, L., & Buising, C. (2000). A peer-reviewed research assignment for large classes. *Journal of College Science Teaching*. Retrieved from http://www.nsta.org/publications/ news/story.aspx?id=40836

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics*, *115*(4), 1239–1285.

Jackson, C. K., Johnson, R., & Persico, C. (2014). The effect of school finance reforms on the distribution of spending, academic achievement, and adult outcomes. NBER Working Paper No. 20118. Retrieved from http://socrates.berkeley.edu/~ruckerj/Jackson_Johnson_Persico_SFR_LRImpacts.pdf

Jakubowski, M., & Sakowski, P. (2006). Quasi-experimental estimates of class size effect in primary schools in Poland. *International Journal of Educational Research, 45(3)*, 202–215. doi:10.1016/j.ijer.2006.11.003

Jepsen, C. (2015). Class size: does it matter for student achievement? *IZA World of Labor, 190*, 1–10. doi: 10.15185/izawol.190

Jepsen, C., & Rivkin, S. (2009). Class size reduction and student achievement: The potential trade off between teacher quality and class size. *Journal of Human Resources, 44*(1), 223–250. Doi: 10.1353/jhr.2009.0008

Jez, S., & Wassmer, R. (2015). The impact of learning time on academic achievement. *Education and Urban Society, 47*(3), 284–306. doi:10.1177/0013124513495275

Krassel, K. F., & Heinesen, E. (2014). Class-size effects in secondary school. *Education Economics, 22*(4), 412–426. doi: 10.1080/09645292.2014.902428

Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics, 114* (2), 497–532. doi: 10.1162/003355399556052

Krueger, A.B. (2000). *Economic considerations and class size* (Working Paper No. 447). Retrieved from Princeton University, Data Space website: http:// dataspace.princeton.edu/jspui/bitstream/88435/dsp019z902z86c/1/447.pdf

Krueger, A. (2002). A response to Eric Hanushek's evidence, politics, and the class size debate. In L. Mishel, & R. Rothstein (Eds.), *The class size debate* (pp. 67-87). Washington, DC: Economic Policy Institute.

Krueger, A. B. (2003). Economic considerations and class size. *The Economic Journal, 113* (485), F34-F63. doi: 10.1111/1468-0297.00098

Kornfeld, M. (2010). *The effects of class size on student academic achievement in rural state.* Unpublished EdD. Vermont, University of Vermont.

Lewit, E.M., & Baker, L. S. (1997). Class size. *The Future of Children Financing Schools, 7*(3), 112-121.

Maasoumi, E., Millimet, D. L., & Rangaprasad, V. (2003) *Robust inference concerning the impact of class size on student achievement.* The 8th Annual Texas Camp Econometrics and Georgetown University. Dallas. Retrieved from: http://www.researchgate.net/publication/228422044_Robust_Inference_Concerning_the_Impact_of_Class_Size_on_Student_Achievement

Mathis, W. J. (2016). *Research-based options for education policymaking: The effectiveness of class size reduction.* Boulder (Colorado): The National Education

Policy Center (NEPC). Retrieved from http://nepc.colorado.edu/publication/research-based-options

Mintah, E. K., (2014). Using group method of teaching to address the problem of large class size: An action research. *International Journal of Learning & Development, 4*(2), 82–97. doi:10.5296/ijld.v4i2.5707

Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., & Ehrle, K. (1999). Evaluating the SAGE program: A pilot program in targeted pupil-teacher reduction in Wisconsin. *Educational Evaluation and Policy Analysis,* Special Issue: Class Size: Issues and New Findings, *21*(2), 165–177. doi:10.3102/01623737021002165

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *Critical Issues for Children and Youths (The Future of Children), 5*(2), 113–127

Mosteller, F., Light, R. J., & Sachs, J. A. (1996). Sustained inquiry in education: Lessons from skill grouping and class size. *Harvard Educational Review, 66*(4), 797–842. doi: http://dx.doi.org/10.17763/ haer.66.4.36m 328762x21610x

Motulsky, H. (2014**).** *Intuitive biostatistics: A nonmathematical guide to statistical thinking.* 3rd edition, New York: Oxford University Press

Mueller, S. (2013). Teacher experience and the class size effect — experimental evidence. *Journal of Public Economics*, *98*, 44–52. doi:10.1016/j.jpubeco.2012.12.001

Nascimento, P.A.M.M. (2008). *School resources and student achievement: Worldwide findings and methodological issues*, M.Sc. in Economics of Education, Institute of Education, University of London.

Nye, B., Hedges, L.V., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis, 21*(2), 127-142. doi:10.3102/01623737021002127

Nye, B., Hedges, L. V., & Konstantopoulos, S. (2000). The effects of small classes on academic achievement: The results of the Tennessee class size experiment. *American Educational Research Journal, 37*(1), 123–151.

Nye, B., Hedges, L. V., & Konstamtopoulos, S. (2002). Do Low-Achieving Students Benefit More from Small Classes? Evidence from the Tennessee Class Size Experiment. *Educational Evaluation and Policy Analysis,* 24 (3), 201-217.

OECD (1974). *New patterns of teacher education and tasks: General analysis.* France: Organisation for Economic Co-operation and Development.

OECD (2006). *The high cost of low educational performance.* Paris: OECD Publishing.

OECD (2013). *Education indicators in focus*. Retrieved from National Centre for Education Statistics, Institute of Education Science: http://www.oecd.org/education/skills-beyond-school/EDIF%202013--N%C2%B012%20FINAL.pdf

OECD (2014). *PISA 2012 Technical Report*. Paris: OECD Publishing.

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation, 9*(6). Retrieved October 3, 2016 from http://PAREonline.net/getvn.asp?v=9&n=6.

Oyserman, D., Coon, H. M., Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. Psychological Bulletin, *128*(1), 3-72. http://dx.doi.org/10.1037/0033-2909.128.1.3

Panizzon, D. (2015). Impact of geographical location on student achievement: Unpacking the complexity of diversity. In A. Bishop, H. Tan, & T. N. Barkatsas (Eds.), *Diversity in Mathematics Education: Towards Inclusive Practices* (pp. 41-61). Switzerland: Springer International Publishing.

Phuong-Mai, N., Terlouw, C., & Pilot, A. (2005). Cooperative learning vs Confucian heritage culture's collectivism: confrontation to reveal some cultural conflicts and mismatch, *Asia Europe Journal, 3*(3), 403–419. doi:10.1007/s10308-005-0008-4

R Core Team (2014). *R: A language and environment for statistical computing.* R foundation for statistical computing, Vienna, Austria. URL http://www.R-project.org/

Schanzenbach, D.W. (2014). *Does class size matter?* Boulder, CO: National Education Policy Centre. Retrieved from http://nepc.colorado.edu/ publication/does-class-size-matter

Stern, D. (1987). *Teachers' salaries, class size, and student achievement in grades 5 and 6: Some new evidence.* Paper presented at the Annual Meeting of the American Educational Research Association. Washington, DC. Retrieved from ERIC database. (ED288241)

Stoet, G. & Geary, D.C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence, 48,* 137–151.

The White House (2000). *The Clinton-Gore administration a record of progress, the Clinton presidency: Eight years' of peace, progress and prosperity*. Retrieved from: http://clinton5. nara.gov/WH/Accomplishments/ eightyears-02.html

Tienken, C. H., & Achilles, C. M. (2006). Making class size work in the middle grades. *Journal of Scholarship and Practice, 3*(1), 26–34

UNESCO Institute for Statistic (2013a). *Global teacher shortage.* Retrieved from United Nations Scientific, Educational and Cultural Organisation: http://www.uis.unesco.org/ Education /Pages/world-teachers-day-2013.aspx

UNESCO Institute for Statistic (2013b) *A teacher for every child: Projecting global teacher needs from 2015 to 2030,* No. 27. Retrieved from United Nations Scientific, Educational and Cultural Organisation: http://www.uis.unesco.org/Education/Documents /fs27-2013-teachers-projections.pdf

Voke, H. (2003). Responding to the teacher shortage. In M. Scherer (Eds.), *Keeping good teachers* (pp. 3-13). Virginia: Association for Supervision and Curriculum Development (ASCD).

Wilson, V. (2006). *Does small really make a difference? An update: A review of the literature on the effects of class size on teaching practice and pupils' behaviour and attainment* (SCRE Research Report No 123) retrieved from the University of Glasgow, The Scottish Council for Research in Education (SCRE) Centre: http://eprints.gla.ac.uk/117736/1/117736.pdf

Wursten, H., & Jacobs, C., (2013). *The impact of culture on education.* The Hofstede Centre, Itim International retrieved from: https://geert-hofstede.com/tl_files/images/site/social/Culture%20and%20education.pdf

Wößmann, L. (2003a). Schooling resources, educational institutions and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics, 65*(2), 117–170. doi: 10.1111/14 68-0084.00045

Wößmann, L. (2003b). *Educational production in East Asia: The impact of family background and schooling policies on student performance*, IZA Discussion Paper No. 745, Institute for the Study of Labour, Bonn. Retrieved from: http://www.iza.org

Wößmann, L. (2005), Educational production in East Asia: The impact of family background and schooling policies on student performance. *German Economic Review, 6*(3), 331–353. doi: 10.1111/j.1468-0475.2005.00136.x

Wößmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review, 50* (3), 695–736. doi:10.1016/j.eu roecorev.2004.11.005

Yang, H.-J. (1993). *Communication patterns of individualistic and collective cultures: A value based comparison.* Paper presented at the Annual Meeting of the Speech Communication Association (79th, Miami Beach, FL, November 18-21, 1993). Retrieved from ERIC Number: ED366032

Zyngier, D. (2014). Class size and academic results, with a focus on children from culturally, linguistically and economically disenfranchised communities. *Evidence base, 1*, 1–23. Retrieved from: https://journal.anzsog.edu.au/publications/9/EvidenceBase2014Issue1.pdf