

education policy analysis archives

A peer-reviewed, independent,
open access, multilingual journal



Arizona State University

Volume 28 Number 67

April 20, 2020

ISSN 1068-2341

Is Shadow Education the Driver of East Asia's High Performance on Comparative Learning Assessments?¹

Jeremy Rappleye



Hikaru Komatsu

Kyoto University
Japan

Citation: Rappleye, J., & Komatsu, H. (2020). Is shadow education the driver of East Asia's high performance on comparative learning assessments? *Education Policy Analysis Archives*, 28(67).
<https://doi.org/10.14507/epaa.28.4990>

Abstract: East Asian students consistently top comparative assessments of academic achievement. Yet, rather than attempting to develop more sophisticated understandings of this difference, the most common reaction is to attribute East Asian performance to longer study hours and/or the attendance at schools focused on academic skill enhancement and test preparation (i.e., juku). Herein we seek to contribute to a richer debate both by presenting new data and findings in relation to Japan, and by highlighting new analytical strategies to understand the relationship between East Asian performance and shadow education. Specifically, we highlight that comparatively high levels of achievement among Japanese students were apparent even at the level of fourth graders, even though juku attendance was low prior to this stage. This suggests that juku attendance is not the primary factor for the high academic achievement of Japanese students. The wider significance of these findings lies in countering both common portrayals of East Asian success and factually inaccurate information disseminated by organizations such as the OECD. In so doing, researchers are in better position to

¹ Both authors contributed equally to this piece.

Journal website: <http://epaa.asu.edu/ojs/>

Facebook: /EPAAA

Twitter: @epaa_aape

Manuscript received: 8/18/2019

Revisions received: 11/15/2019

Accepted: 11/15/2019

elaborate new, more sophisticated theories that explain East Asia's consistently world-leading academic achievement.

Keywords: alternative theories; international large-scale assessments; scale mismatch; stereotypes; TIMSS

¿Es la educación *en la sombra* el impulsor del alto rendimiento de este de Asia en las evaluaciones de aprendizaje comparativo?

Resumen: Los estudiantes de este de Asia constantemente superan las evaluaciones comparativas de rendimiento académico. La comprensión común de esta diferencia es atribuir el rendimiento del este asiático a horas de estudio más largas y / o la asistencia a escuelas enfocadas en la mejora de habilidades académicas y la preparación de exámenes (es decir, juku). Este estudio presenta nuevos datos y hallazgos de Japón y destaca estrategias de noticias para comprender la relación entre el desempeño de este de Asia y la educación *en la sombra*. Los niveles de logro comparativamente altos entre los estudiantes japoneses fueron evidentes a nivel de cuarto grado, a pesar de que la asistencia al juku era baja antes de esta etapa. Esto sugiere que la asistencia al juku no es el factor principal para el alto rendimiento académico de los estudiantes japoneses. Los hallazgos contrarrestan tanto las representaciones comunes del éxito de este de Asia como la información inexacta difundida por organizaciones como la OCDE. La investigación futura necesita desarrollar teorías que expliquen mejor el logro académico líder mundial de este de Asia.

Palabras clave: teorías alternativas; evaluaciones internacionales a gran escala; desajuste de escala; estereotipos TIMSS

¿A educação *na sombra* do impulsor do alto desempenho da Ásia leste nas avaliações de aprendizagem comparativa?

Resumen: Os estudiosos da Ásia leste podem superar as avaliações comparativas de desempenho acadêmico. A composição comum dessa diferença é atribuir o retorno a este número de horas de estudo asiáticas, maior do que o número de sessões de assistência enfocadas na grande quantidade de habilidades acadêmicas e na preparação de exames (es decir, juku). Este estudo apresenta novos dados e resultados do Japão e exhibe estratégias de notícias para incluir a relação entre o desempenho da Ásia e a educação *na sombra*. Os níveis de logotipos comparativamente altos entre os estudantes japoneses japoneses evidenciam um nível de desenho animado, um peso que ajuda na era juku antes de esta etapa. É sugerido que a ajuda no juku não é o principal fator para o alto desempenho acadêmico dos estudantes japoneses. Os controles contrariam tanto as representações comuns do movimento da Ásia leste como as informações inexatas difundidas por organizações como o OCDE. A investigação futura necessária descreve as teorias que explicam melhor o logotipo acadêmico líder mundial da Ásia leste.

Palavras-chave: teorías alternativas; avaliações internacionais a gran escala; desajuste de escala; estereotipos TIMSS

Is Shadow Education Driving East Asian Achievement?

As global interest in international large-scale assessments (ILAs) continues apace, one enduring puzzle has been the drivers of high achievement across East Asia. Not only have East Asian students – as a whole – consistently outscored other countries by one full standard deviation over the past decade in the Programme for International Student Assessment (PISA) assessments, but the gap between the East Asian countries and other countries on TIMSS has grown wider in recent decades (TIMSS International Study Center, 2016). While a few scholars have argued that achievement score differences must be understood as rooted, at least to some extent, in deeper divergences in views on the nature of knowledge, the meaning of education, and concepts of selfhood (e.g., Komatsu & Rapple, 2017b; Leung, 2001; Tan, 2018), the majority of analysts have maintained that score variations can be explained by institutional and/or systemic arrangements. Since the inception of PISA in the late 1990s, OECD analyses have consistently championed the latter view. Evidence for this can be found in the OECD's repeated assertion that East Asia's world leading test scores are due, in part, to the prevalence of shadow education. For example, a 2011 OECD report on Japan argued that “the growing investment in *juku* suggests that they positively influence school performance” and “contribute to Japan's results on the PISA assessments” (OECD, 2011, p. 18). The OECD's official 2015 PISA results analysis highlighted concerns over a “performance gap” and explicitly mentions Japan, Korea, Macau, and Chinese Taipei, hinting strongly that “differences are the result of private tutoring and a pervasive shadow education system” (OECD, 2015, p. 6).

A similar narrative can be found in leading Western media outlets with recent articles. For example, an article in *The Economist* focused on Japanese *jukus* (2011) and South Korea *bagwons* (2015). In 2005, *Bloomberg News* carried the story “Japan: Crazy for Cramming”, following a 2002 *Time International* cover story entitled “School Daze: Cramming. Bullies. Rote Learning. East Asia's Schools Are Failing Their Students. Big Changes are Planned, But Will They Come Soon Enough?” (Beech, 2002). The tone and tenor of this coverage is largely indistinguishable from a 1992 front-page *New York Times* article seeking to explain Japan's world-leading success in the 1980s: “How Do Japanese Students Do It? They Cram” (Weisman, 1992). This narrative is a long standing and frequently revamped one. Yet, it is one, we argue, that obstructs a deeper discussion from unfolding.

There is also no dearth of academic scholarship suggesting the link between East Asian achievement and the prevalence of shadow education. For example, a recent 2018 piece in *Science*, written by leading researchers of educational assessment at Harvard University, discussed “East Asian Successes,” but suggested these could largely be explained by shadow education:

...in East Asia, as elsewhere, private tutoring is widespread. Korea is the most prominent example, where approximately half the 2012 PISA participants reported receiving private tutoring, often focused on test preparation....One plausible interpretation of Korea's and other East Asia jurisdictions' ILSA results is that they are not a result of the public education system but rather this substantial private investment (Singer & Braun, 2018, p. 39).

While that *Science* article offers no deeper statistical backing for the conjecture, there are other studies which do. For example, working in-country to collect data and using multiple regression analysis, Entrich (2014) “found shadow education to be a decisive determinant for higher achievement scores in international comparison....the higher usage ratio of out-of-school education provides a reasonable explanation for the Japanese success in PISA.” (p. 95)

However, researchers, many from East Asia, who have looked deeply into the issue have been unable to verify the purportedly strong links between shadow education and higher achievement, either within countries or between countries. An extended study from 2015 out of the University of Tokyo utilizing PISA 2006 data to compare the effects of supplementary tutoring on mathematics achievement in Japan and the United States suggested “no overall effect of out-of-school tutoring in Japan” (Mori, 2015, p. 70). In Korea, drawing on new large-scale longitudinal educational surveys, a 2014 comprehensive study found that for middle school achievement “cram schools made a small difference in achievement gains in math, whereas other forms of shadow education (e.g., individual tutoring, correspondence courses, on-line tutoring services, and EBS) made little difference” (Byun, 2014, p. 15, see also Park, 2013). Similarly, a 2011 analysis of Taiwanese middle school students’ mathematics achievement utilizing propensity score matching and nationally representative panel data detected a “positive but fairly small” effect (Kuan, 2011, p. 363).

At present, as one Korean scholar confides, the links between shadow education and achievement in East Asia remain unclear for methodological reasons: “although the effects of shadow education on academic achievement have been widely investigated in recent years, emerging empirical evidence has been inconsistent, contradictory, and even confusing,” underscoring that this uncertainty “precludes definitive conclusions about the effects of shadow education” (Byun, 2014, p. 40). Our intended contribution in this piece is situated within this wider constellation and uncertainty, and is best conceptualized on two distinct levels.

First, within the on-going scholarly investigations of shadow education and achievement in East Asia, we seek to put forth several new analytical strategies that may help nuance and advance existing research.² In doing so, we primarily focus on Japan, where there is both an abundance of unexamined domestic survey data and relative dearth of research in the English language (as compared to Korea and Taiwan). Our analytical innovation centers on finding new ways to take up the problem that do not rely on individual-scale analyses to describe country-scale phenomenon. As we detail below, this problem of scale mismatch is found in Entrich (2014), among many others. Certainly, such individual-scale analyses have their own advantages and purposes, and will remain important. But we seek to underscore that it is problematic if educational research lacks analytical strategies for understanding whether or not shadow education plays a decisive role in determining country-scale achievement and between-country differences, particularly at this juncture when OECD analyses and media outlets casually observing ILSA results confidently suggest this is the case.

Against this backdrop, we propose a novel strategy founded on the fact that shadow education differs among different grades (i.e., school year of students) and subjects. If shadow education is really the primary factor explaining achievement differentials among countries, we would observe more pronounced between-country differences in achievement in those grades and subjects wherein shadow education is more pervasive. Concretely, this means analyzing (i) the link between shadow education and achievement link for particular subjects, and (ii) not at age 15 (PISA) but at the elementary school level (i.e., fourth grade). To underline the importance of the scale mismatch problem, we also show that juku attendance did not explain international differences, nor inter-regional differences in achievement (when using Japanese domestic surveys).

² It is important to note that issues of shadow education extend far beyond links to achievement, encompassing crucial issues of inequality (e.g., Dawson, 1999), policy planning (e.g., Bray, 1999), and local variations and distortions of the ideal of public education (e.g., Brehm & Silova, 2014). Space does not allow us to connect our discussion with these issues herein.

The second dimension of our intended contribution is to challenge the authority of the OECD-style portrayal of East Asia's world leading achievement. While many researchers seem content to simply labor away on the analytical question, we find ourselves increasingly dissatisfied with analyses that do not engage with the wider policy and epistemic context in which we work. That is, the findings (and continued uncertainty) we present herein are not simply of academic interest, but also raise the question of the basis of authority the OECD and Western media outlets stand upon when drawing conclusions about East Asia's academic achievement. As we show, there is little evidence to support the idea that East Asia's achievement is primarily driven by shadow education, but yet the OECD and Western media repeatedly present that view in an authoritative voice. Importantly, this problem is not limited to issues surrounding shadow education but extend much more broadly to portrayals of entire systems across East Asia. Take for example, the OECD's 2012 report entitled *Lessons from PISA for Japan* that states:

At the primary school level, juku participation increased from 16% in 1985 to 26% in 2007, and at the lower secondary level, from 44% to 53%. *At the upper secondary level, participation in private tutoring is even greater* (Figure 2.2.4; OECD, 2012, p. 72, italics added)

Here the word private tutoring means juku. But as shown in Fig. 1 (below), this latter assertion (italics) is factually inaccurate, even blatantly so. Moreover, the OECD concludes that this rising juku attendance is “driven by the severe competition to enter the country's top universities.” (OECD, 2012; see also OECD, 2011, p. 14). But a series of recent papers, we have shown how learning time, pressure surrounding study, and the stresses surrounding college entrance exams – the so-called East Asian ‘Exam Hell’ – is no longer existent in Japan (Komatsu & Rappleye, 2018, 2020; Rappleye & Komatsu, 2018). More broadly, in other recent studies, we have shown how the OECD portrayals of the ‘East Asian Miracle’ are also unsupported by data (e.g., Komatsu & Rappleye, 2019), and how the purported links between PISA Scores and economic growth rates worldwide suggested by the OECD are also based on flawed statistics (Komatsu & Rappleye, 2017a; Rappleye & Komatsu, 2020). All of this points to how stereotypes – *later presented as rigorous, evidence-based research on East Asian education* – continue to be disseminated by the OECD and Western media outlets.

Rather than improving our understanding, these stereotypes actually frustrate and mislead. The consequence, we argue, is that the deeper, alternative perspectives on, say, the nature of knowledge, the meaning assigned to education, and concepts of selfhood existent in East Asia (e.g., Markus & Kitayama, 1991) are replaced not simply by structural analyses (i.e., *a priori* premised on the idea of no deeper diversity of worldviews), but upon disappointingly shallow ones. Thus, whilst we continue to investigate the analytical question herein, we also seek to clear up some of the blatant misinformation promoted by the OECD about East Asia. In so doing, we seek to invite readers' attention to the wider political and epistemic questions in play here as well, not simply the analytical details. We return to address these wider themes in the final section of our conclusion.

Materials, Data, and Methods

Situation of Shadow Education Attendance in Japan

To understand the patterns of shadow education attendance across Japan, we used data collected by the Ministry of Education, Culture, Sports, Science and Technology (MEXT, 2008). As supplement, we also used data collected by the NHK Broadcasting Culture Research Institute (2013). NHK is the national broadcasting organ and research arm of the Japanese government and is

charged with conducting non-partisan public surveys (analogous to Pew Survey research in the USA).

MEXT recorded data for shadow education attendance rates for first – sixth graders (i.e., primary students) and seventh – ninth graders (lower secondary students). This survey was conducted in 1976, 1985, 1993, 2002, and 2007. In each survey cycle, MEXT randomly selected municipal governments nationwide for which shadow education attendance would be examined. The final sample size of students was approximately about 1% of the total student population in Japan (total population: 53,458 students in the 2007 academic year). MEXT then examined the nature and quality of shadow education attendance for the sample students using questionnaire. This questionnaire included questions about (1) the specific form of shadow education (schools focused on academic skill enhancement and test preparation, home tutoring, remote tutoring, or other types such as sports, music, and calligraphy) and (2) which subject the student learns in shadow education. Hereafter, the term *juku* is used to exclusively refer to schools focused on academic skill enhancement and test preparation.

NHK Broadcasting Culture Research Institute data included the mean *juku* attendance rate for 10th – 12th graders (i.e., upper secondary students) as well as the mean *juku* attendance rate for seventh – ninth graders. The survey was conducted in 1982, 1987, 1992, 2002, and 2012. In each survey cycle, NHK researchers first randomly selected municipal governments, and then visited the same municipalities repeatedly, selecting students at random using the municipal population registry. NHK researchers then visited their homes to conduct interviews. In the interviews, NHK researchers examined whether or not the student attended a *juku*. The sample size for 10th – 12th graders ranged between 969 – 1350 students in all survey cycles.

Considering the random sampling strategy used by MEXT and the NHK Broadcasting Culture Research Institute, we anticipated that the data would contain only very small systematic biases. Indeed, the mean *juku* attendance rate for seventh – ninth graders derived from the former dataset was almost the same as that derived from the latter dataset, suggesting little systematic biases in the data. It would be thus reasonable to conflate data for first – ninth graders from the former dataset and those for 10th – 12th graders in our analysis. Using these data, we analyzed in which grade and in which subject *juku* attendance was common. Based on this analysis, we generated a hypothesis about when the achievement gap between Japan and other countries should become apparent if shadow education attendance was indeed the primary factor for explaining Japan's achievement.

Achievement Gap Between Japan and Other Countries

We used data derived from the Third International Mathematics and Science Study (TIMSS, TIMSS & PIRLS International Study Center, 2018a,b). Our selection of TIMSS was because TIMSS had achievement data for two different grades, i.e., fourth and eighth grades, whereas PISA includes achievement data only for one fixed age (15 year olds). TIMSS thus allows us to examine in which grade the achievement gap between East Asia (in this case only Japan) and other countries is most pronounced.

To examine the achievement gap between Japan and other countries, we used achievement data of Math and Science for fourth and eighth graders obtained by TIMSS (1995, 2003, 2007, 2011, and 2015). We did not use data obtained by TIMSS 1999, because TIMSS 1999 did not survey achievement for fourth graders. To estimate achievement gap between Japan and other countries (or regions), we first selected countries which regularly participated in TIMSS (i.e., England, Hong Kong, Hungary, Iran, Norway, Singapore, Slovenia, and United States). We then calculated the difference in Japan's achievement score and the mean score for these regular participant countries.

We also calculated the difference divided by the standard deviation (SD) of the scores for the regular participant countries on the assumption that the regular participant countries were the control group (see Ellis, 2010). The focus of our analysis was whether or not the achievement gap observed here meshed with the variation in the juku attendance rates between different grades and between math and science.

Regional Variations in Juku Attendance and Achievement

To examine the relationship between the regional variation in juku attendance and that in academic achievement in Japan, we used data collected by MEXT's National Survey of Academic Achievement and Learning Conditions conducted in 2007 (MEXT, 2008). This survey aimed to assess academic achievement and learning conditions for all sixth graders and ninth graders in public schools in Japan. Virtually all public primary schools (19,251 among the 19,361 schools) participated in this survey. This survey included data of the students' achievement of two math subjects and two language subjects for sixth and ninth graders for different prefectures. This survey also included data for the juku attendance rates for different prefectures. We utilized data for the mean of the correct answer rates for the two math assessments and the juku attendance rates of sixth graders. These data were derived from the National Institute for Education Policy Research (2018) and the Prefecture Rankings (2018).

Using these data, we examined the relationship between juku attendance rates and math achievement among different prefectures. We would obtain a positive correlation if juku attendance was the primary factor determining academic achievement. To examine the correlation, we used the Pearson's correlation coefficient (r) and its 95% confidence interval (CI). CI was calculated using the bootstrapping method (Efron, 1979; Fox, 2008). Specifically, we obtained random samples from the original data with replacement and calculated r values 10,000 times and then identified the range in which 95% of the r values fell (see Komatsu & Rappleye, 2017b).

We were cognizant of the possibility that data for academic achievement of sixth graders might be biased due to the exclusion of academic achievement data for students attending national (*kokuritsu*) and private schools (i.e., usually considered superior academically), although this bias would likely be small due to much lower numbers of national and private primary schools (70 and 217 schools, respectively) as compared with numbers of public (*koritsu*) primary schools (19,648 schools) across Japan. Yet to consider the effect of this potential bias on our results, we also classified prefectures according to the rate of students attending national and private schools. We then examined the relationship between juku attendance and math achievement for prefectures having a similar rate of students attending national and private schools to confirm.

One final note: throughout this study we did not conduct significance testing. The primary reason is that statistical testing is quite often misleading. With a sufficiently large sample size, one can always find a statistically significant difference between any two variables and a statistically significant correlation between the two variables. What is most important is not whether there is a difference or a correlation, but the magnitude of the difference or the strength of the correlation. The confusion created by statistical testing was noted by numerous statisticians for many years (Bakan, 1966; Berkson, 1938; Komatsu, Shinohara, & Otsuki, 2015; Nuzzo, 2014; Thompson, 1996, 2002). What is important from the perspective of education research is that these statisticians include Thompson (1996), who penned the guidelines for statistical reporting for the American Education Research Association. Thompson (1996, 2002), in resonance with other statisticians, recommended reporting effect sizes and confidence intervals instead of reporting statistical significance (i.e., p values).

Results

Situation of Shadow Education Attendance in Japan

In Japan, the attendance of shadow education for academic subjects (i.e., jukus, home tutoring, and remote tutoring) was common at the lower secondary level (grades 7–9), but it was not very common at the primary level (grades 1–6; Fig. 1a). At the primary level, non-academic activities were the focus, particularly in the early grades. It is true that juku attendance increased with grade even at the primary level, but the incremental change becomes considerable after the fourth grade (Fig. 1b). Home tutoring and remote tutoring changed little with grade and were supplementary in Japan.

We observed a clear contrast in juku attendance between math and science. A majority of students attending juku took math lessons, while very few students took science lessons (Fig. 2). For example, the mean percentage of first – fourth graders taking math lessons was 15.0%. The same percentage for those taking science lessons was merely 0.9%.

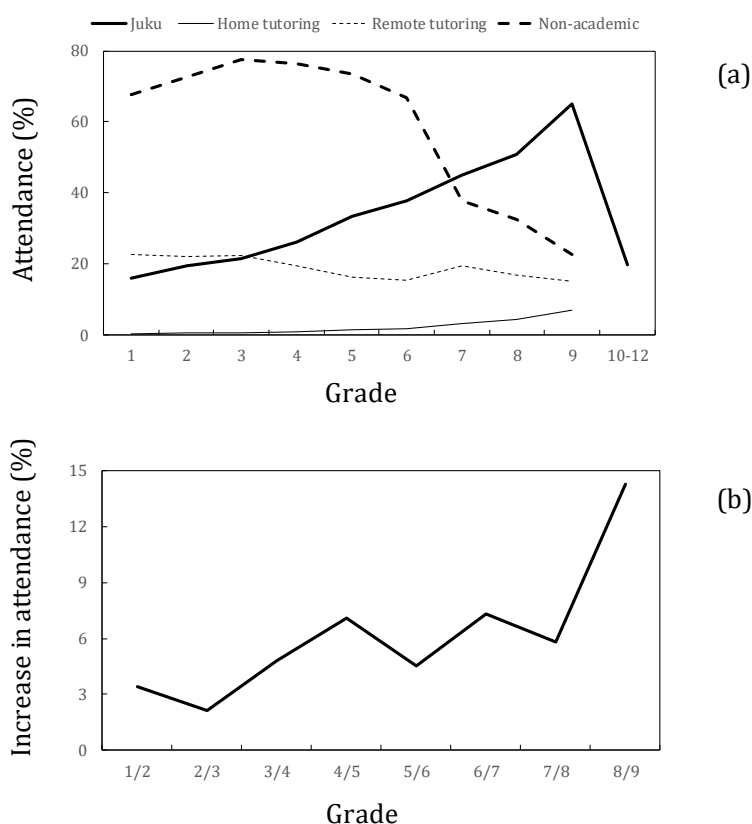


Figure 1: (a) Variations in juku attendance rates among different grades and (b) increases in juku attendance rates between grades.

Note: The data year was 2007 for grades 1–9 and 2012 for grades 10–12. Original data were derived from MEXT (2008) and the NHK Broadcasting Culture Research Institute (2013).

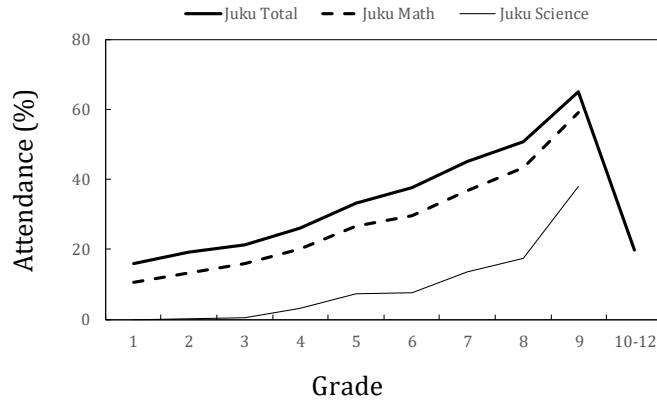


Figure 2: Variations in juku attendance rates for different subjects (i.e., math and science).

In relation to claims reviewed above that juku attendance was the primary cause for Japan's academic achievement, we can generate the hypothesis that the achievement gap between Japan and other countries would be pronounced only in math at higher grades (i.e., seventh – ninth grades). This hypothesis will be tested using TIMSS data below.

It is important to note here that the period when TIMSS data were available was exactly the period when juku attendance was the most prevalent in Japan's history (Fig. 3). Juku attendance had increased between 1976 and 1993 and it has remained fairly consistent since then. Juku attendance before 1976 would be less prevalent. Indeed, the survey of juku attendance in 1976 was triggered by the perception that juku attendance had become unprecedentedly widespread (Yamada, 2014). Data after 2007 were lacking in the MEXT dataset, but juku attendance is unlikely to have changed greatly since 2007. A different dataset covering the period 2009–2017 provided by a major private education institute in Japan (Benesse, 2017) observed no pronounced changes in juku attendance in the period, although the data were not fully comparable with MEXT data due to some differences in the definition of juku attendance.

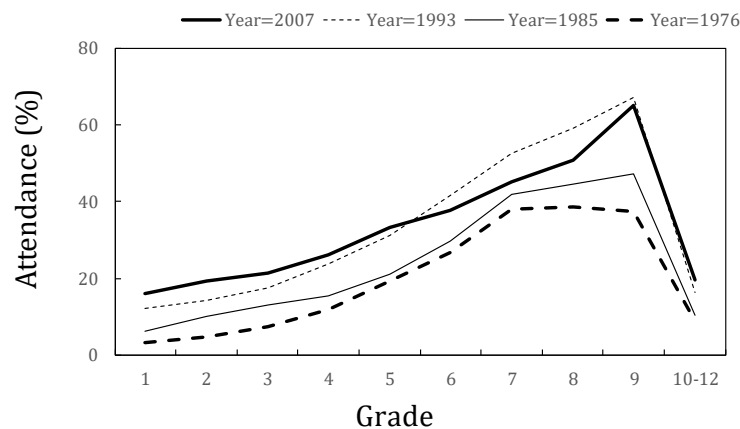


Figure 3: Variations in juku attendance rates for different survey years.

Note: The survey years were for the surveys for first – ninth graders (MEXT 2008). Data for grades 10–12 were derived from the survey conducted by the NHK Broadcasting Culture Research Institute (2013) in the closest year.

Achievement Gap Between Japan and Other Countries

Table 1 summarizes achievement of math and science for Japan and other countries which regularly participated in TIMSS. The achievement gap between Japan and other countries was apparent not only at the eighth-grade level but importantly already present at the fourth grade level. For example, Japan's math scores for fourth and eighth graders were 593 and 586 points, respectively, in TIMSS 2015. These scores were much higher than the mean values for other countries (536 and 526 points, respectively). These results held even when comparing Japan with each regular participant (Table A1 of Appendix A). In other words, the achievement gap between Japan and each participant was apparent already at the fourth grade level. We obtained qualitatively the same results for science. Japan's science scores for fourth and eighth graders were 569 and 571 points, respectively. These scores were again much higher than the mean values for other countries (529 points for both cases). Japan's scores were generally higher than the other regular participating countries by one full standard deviation regardless of the grade and subject in TIMSS 2015 (Table 1).

Table 1

Achievement of Japan and other regular participants in TIMSS (England, Hong Kong, Hungary, Iran, Norway, Singapore, Slovenia, and the United States)

	Japan (points)	Mean of regular participants (points)	SD of regular participants (points)	Japan-Mean (points)	(Japan-Mean) /SD
2015					
Math, 4 th grade	593	536	57	57	1.01
Math, 8 th grade	586	526	54	61	0.90
Science, 4 th grade	569	529	48	41	1.18
Science, 8 th grade	571	529	39	42	0.94
2011					
Math, 4 th grade	585	531	53	54	0.98
Math, 8 th grade	570	514	57	56	1.02
Science, 4 th grade	559	524	36	35	1.01
Science, 8 th grade	558	527	32	31	1.04
2007					
Math, 4 th grade	568	520	62	48	1.30
Math, 8 th grade	570	510	55	61	0.91
Science, 4 th grade	548	524	44	24	1.81
Science, 8 th grade	554	523	32	31	1.03
2003					
Math, 4 th grade	565	508	62	57	1.10
Math, 8 th grade	570	511	59	59	1.00
Science, 4 th grade	543	510	47	33	1.44
Science, 8 th grade	552	527	36	25	1.45
1995					
Math, 4 th grade	567	499	58	68	0.86
Math, 8 th grade	581	513	53	68	0.79
Science, 4 th grade	553	495	48	58	0.83
Science, 8 th grade	554	521	31	34	0.92

Note. Original data for the regular participant countries are shown in Table A1 in the Appendix. Data for 1995 were inconsistent between older and more recent TIMSS reports. We used data derived from recent TIMSS reports (TIMSS & PIRLS International Study Center, 2018a, pp. 27–30; TIMSS & PIRLS International Study Center, 2018b, pp. 34–37). We confirmed that our conclusions hold even when using data derived from old TIMSS reports (e.g., TIMSS International Study Center, 1997a, p. 24; TIMSS International Study Center, 1997b, p. 22).

In light of the fact that juku attendance was uncommon at the primary level (particularly prior to Grade 4) and that primary school juku attendance in science subjects was virtually non-existent, TIMSS data did not support the hypothesis that juku attendance was the primary factor for high academic achievement of Japan.

Regional Variations in Juku Attendance and Achievement

Table 2 shows regional variations in juku attendance and math achievement for sixth graders in Japan. Juku attendance rates varied considerably among prefectures. The range was between 22.1% (Akita) and 57.9% (Tokyo). In general, juku attendance was relatively high for urban prefectures (e.g., Tokyo, Kanagawa, Hyogo, and Nara). However, academic achievement for these prefectures was not always high. The achievement scores for Hyogo and Nara (62 points for both prefectures) were less than the mean score for all prefectures (63 points), although the scores for Tokyo and Kanagawa (65 and 64 points, respectively) were slightly higher than the mean. We rather found an interesting paradox, as viewed by those who would advocate the juku-achievement hypothesis: rural prefectures where juku attendance rates were low often had the highest levels of academic achievement. Indeed, the achievement score for Akita was 67 points, which was the second highest among all prefectures. The score for Akita was higher than the mean by 2.3 *SD*, whereas the juku attendance rate for Akita was lower than the mean by 2.4 *SD*. Furthermore, the highest achievement score was recorded by Ishikawa of which juku attendance rate (28.2%) was considerably lower than the mean for all prefectures (41.8%). The score for Ishikawa was higher than the mean by 3.4 *SD*, whereas the juku attendance rate for Ishikawa was lower than the mean by 0.6 *SD*. We further found the overall correlation between juku attendance rates and academic achievement was not positive ($r = -.16$ with CI between $-.44$ and $.16$, Fig. 4a).

Table 2
Regional variations in juku attendance rates and academic achievement for six graders among prefectures in Japan

Prefecture	Population (thousand persons)	Population density (persons per km ²)	Juku attendance (%)	Achievement (%)	Attendance for private or national schools (%)
Aichi	7,483	1,446.7	49.3	61.0	0.7
Akita	1,023	87.9	22.1	67.0	1.3
Aomori	1,308	135.6	28.1	64.0	0.9
Chiba	6,223	1,206.5	51.0	61.5	1.2
Ehime	1,385	244.1	46.2	66.0	0.8
Fukui	787	187.7	38.0	66.5	1.0
Fukuoka	5,102	1,023.1	39.6	62.5	1.4
Fukushima	1,914	138.9	33.2	63.0	1.4
Gifu	2,032	191.3	46.9	61.0	1.1
Gunma	1,973	310.1	44.0	61.0	1.4
Hiroshima	2,844	335.4	45.6	64.0	2.0
Hokkaido	5,382	68.6	36.7	60.5	0.8
Hyogo	5,535	658.8	54.4	62.0	1.5
Ibaraki	2,917	478.4	41.2	64.0	1.2
Ishikawa	1,154	275.7	36.9	69.0	1.2
Iwate	1,280	83.8	25.8	63.0	1.1
Kagawa	976	520.2	44.0	63.5	2.0
Kagoshima	1,648	179.4	32.3	62.0	1.4

Table 2 cont.

Regional variations in juku attendance rates and academic achievement for six graders among prefectures in Japan

Prefecture	Population (thousand persons)	Population density (persons per km ²)	Juku attendance (%)	Achievement (%)	Attendance for private or national schools (%)
Kanagawa	9,126	3,777.7	56.9	61.5	2.6
Kochi	728	102.5	46.2	64.0	2.8
Kumamoto	1,786	241.1	35.3	63.0	0.7
Kyoto	2,610	566.0	50.2	64.0	4.3
Mie	1,816	314.5	49.7	61.0	1.3
Miyagi	2,334	320.5	37.8	60.5	1.3
Miyazaki	1,104	142.7	33.2	61.5	1.2
Nagano	2,099	154.8	39.5	62.0	1.3
Nagasaki	1,377	333.3	37.2	60.5	1.6
Nara	1,364	369.6	53.8	61.5	5.0
Niigata	2,304	183.1	35.5	63.0	1.1
Oita	1,166	183.9	39.4	63.5	1.5
Okayama	1,922	270.1	46.1	62.5	1.5
Okinawa	1,434	628.4	37.8	63.5	1.8
Osaka	8,839	4,639.8	50.7	61.5	1.9
Saga	833	341.2	36.5	62.0	1.3
Saitama	7,267	1,913.4	45.4	60.5	0.8
Shiga	1,413	351.7	48.5	59.5	0.8
Shimane	694	103.5	32.1	60.0	1.0
Shizuoka	3,700	475.8	47.4	62.0	1.2
Tochigi	1,974	308.1	41.5	62.0	1.0
Tokushima	756	182.3	48.8	62.5	3.0
Tokyo	13,515	6,168.7	57.9	65.0	4.8
Tottori	573	163.5	35.7	61.5	1.3
Toyama	1,066	251.0	37.1	64.5	0.8
Wakayama	964	203.9	52.9	62.5	2.4
Yamagata	1,124	120.5	28.2	60.5	1.1
Yamaguchi	1,405	229.8	42.4	62.0	1.1
Yamanashi	835	187.0	41.7	60.0	3.9
Mean	2,704	655.3	41.7	62.5	1.6
<i>SD</i>	2,700	1181.6	8.2	1.9	1.0

These results did not change qualitatively when using data for prefectures having similar rates of students attending national and private schools. The rates of students attending national and private schools for most prefectures were less than 3.0%, but there were several prefectures having higher rates (Table 2). We thus examined the relationships between juku attendance rates and achievement rates for three categories having sufficient samples to examine correlation, i.e., those having the rates of students attending national and private schools being less than 1.0%, from 1.0% to 2.0%, and from 2.0% to 3.0% (Figs. 4b, 4c, and 4d, respectively). In all cases, we failed to observe positive correlations. *r* values were -.27 (with the between -.93 and .32), -.38 (with CI between -.58 and .08), and -.94 (with CI between -1.00 and .98), respectively. Note that the very wide ranges of the CIs for the first and third cases suggest that calculating CIs based on such limited samples was not

very meaningful. Juku attendance thus did not explain intra-national regional differences in academic achievement, as well as between-country differences, underlining the importance of the scale mismatch problem.

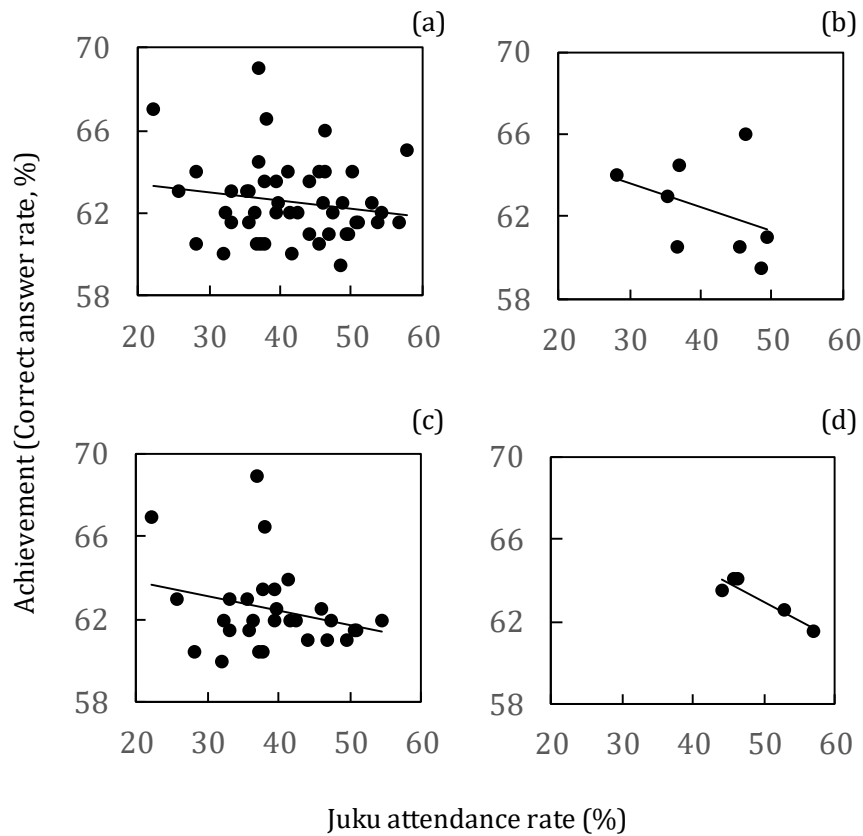


Figure 4: Relationships between juku attendance and achievement for (a) all prefectures, and prefectures having a percentage of students attending private or national schools being (b) less than 1.0%, (c) from 1.0% to 2.0%, and (d) from 2.0% to 3.0%.

Discussion

We are, of course, aware that the correlational analyses above do not indicate that juku attendance had no effect on improving academic achievement. However, this data does strongly suggest that juku attendance is not the primary determinant of the superior performance of Japan, when viewed at the country-scale. To clarify and extend the implications of our findings, we first link to other research and then return to the question of how future research might build on the new analytical strategies we have introduced herein.

Links to Other English-Language Research

The results of our analyses mesh with other previous data and studies. First, Japan's achievement was already high in FIMS (1964) and SIMS (1980-82) which examined math achievements of 13-year-olds for various countries. For example, the mean FIMS score of Japan was 31.16 points (with *SD* being 16.90 points; Postlethwaite, 1967, p. 94). This was much higher than that of the United States whose mean score was 17.85 points (with *SD* being 13.21 points). However, the juku attendance rates of primary students in the 1960s and 1970s was lower than in

the 1990s and later (see Results). It is thus difficult to attribute Japan's achievement in FIMS and SIMS only to juku attendance of primary students.

Second, our results complement findings in *The Learning Gap* by Stevenson and Stigler (1992). Stevenson and Stigler studied math achievement for first and fifth graders in Sendai (Japan) and Minneapolis (the United States) in 1980. They conducted a similar study in Sendai and Chicago in 1987. Their main point was that the achievement gap between Japan and the United States expanded between the first and fifth grades. More importantly, we note that the achievement gap was already present at the first grade level in both studies (conducted in 1980 and 1987). This achievement gap cannot be explained by juku attendance given the low rate of juku attendance by Japanese students at the first grade level. We further underscore that Miyagi prefecture which includes Sendai that was the focus of the Stevenson and Stigler's empirical study is not a region where juku attendance is prevalent (Table 2) and, moreover, juku attendance in the 1980s was less common than in the 1990s and later (Fig. 3).

Towards Analytical and Methodological Innovation

Our results showed that the common hypothesis that shadow education attendance is the leading cause for high academic achievement of East Asian countries cannot be supported by country-scale analyses. This stands in direct opposition to results from several individual scale studies, such as Entrich (2014), that we highlighted at the outset. One major problem we find with Entrich (2014) is scale mismatch between analysis and conclusions. Entrich (2014) found that shadow education is a major factor corresponding to *between-individual* variations in achievement in Japan. However, he concludes that shadow education plays a decisive factor explaining *between-country* variations in achievement. Is it actually possible that a factor explaining micro-scale variations does not explain macro-scale variations? To explore this question, let us take a simple example of two hypothetical countries having different achievement scores in international assessments. We shall assume that all students in Country A study 50 hours a week, while all students in Country B study 2 hours a week. The difference in study time would be one plausible factor explaining the between-country achievement variation. However, study time cannot be a factor explaining between-individual variations in achievement in Country A because of no between-individual variation in learning time. Here we can imagine that individual-scale analysis cannot provide direct evidence for a country-scale phenomenon. The key conceptual-turned-methodological point here is to nuance and advance the discussion with attention to the issue of scale, i.e. selecting an appropriate scale to understand the target phenomenon. We suggest that education researchers should conduct country-scale analyses of East Asian countries to examine the relationship between East Asian performance in international assessments and shadow education, as we have done. Due to the limited number of studies examining the relationship at a country scale, results of individual scales are often being misused to interpret academic achievement at a country scale.

Here it is worth pausing to think about the potential reasons for the shortage of country-scale studies. One apparent reason, of course, would be that historically individual data were more easily available and education scholars persist in using methodologies and strategies developed in the period when international data was unavailable. We, however, surmise another reason, one more philosophical. Education scholars seem to often implicitly assume that the realities are independent of perspective. This assumption might lead scholars to smaller scale analyses. Indeed, we have frequently received negative comments by reviewers of our other manuscripts that country-scale (or other aggregated) data cannot capture the fine-grained realities of education. Our position is different from this. We believe that the "realities" are only defined in relation to perspective and frame. There is little doubt that today national perspectives/frames are being replaced by

international comparative ones, due in large part to the fervor surrounding international large-scale assessments such as TIMSS and PISA. Meanwhile, most education researchers remain wedded to the idea that classrooms are the primary reality of education. But it could be argued that we are in the midst of a slow shift wherein the “primary reality” is the global scale. Without recognizing this, politically expedient but untested hypotheses can easily take hold. Put somewhat less philosophically, reality is defined in relation to the problem, and there has been a large shift in the dominant problem over the past two decades due to the instruments like PISA and the research it makes possible (see Komatsu & Rappleye, 2017b). Our perspectivism position is quite common in other research fields, not only in philosophy but across the natural sciences (e.g., Jarvis & McNaughton, 1986; Komatsu et al., 2012; Waring & Running 1998). We hope this philosophical issue and its practical implications – *How should educational research change in light of new global instruments for policy and research?* – can be taken up more by a new generation of education scholars.

In this context, it should be unmistakable that we disagree with the way the OECD uses its data. The OECD quite often ignores rich domestic data which are potentially useful to check the validity of its arguments (e.g., the data used in Figures 1-3). In this piece, we brought into conversation ILSA data and rich domestic data for shadow education, suggesting that another analytical innovation should be a synthesizing of international and domestic data. This mutual exchange is essential for identifying relevant hypotheses explaining the target phenomenon. When we rely exclusively on international data to conduct analyses, we are often left with many “equally plausible” hypotheses. Yet, this diversity creates the possibility for someone to simply pick the hypothesis which is politically expedient (e.g., Morris, 2012), while rejecting other hypotheses without any deeper ensuing inquiry. Multiple hypotheses are not harmful so long as there is substantive dialogue, but without it the situation is ripe for selective reporting for one’s own benefit. We feel this is frequently the case with OECD analyses: much more attention is given to identifying problems so as to pave the way for selling OECD-style solutions or ‘policy authority’ rather than in furthering understanding of different approaches to education worldwide (Komatsu & Rappleye, in review). We hope that our study, modest as it is, will lead more education scholars to engage in the exchange between international and domestic data, eventually leading to less potential “selling” of merely ideological positions.

Conclusion: From Japan to East Asia? From Analytical Questions to Onto-Epistemic Alternatives?

It remains unclear to what extent our findings from Japan can help in understanding the situation elsewhere in East Asia, i.e. Korea and Taiwan, systems with both similarities and highly significant divergences (see Aizawa, Kagawa, & Rappleye, 2018). Yet, given our preliminary analyses of those contexts woven in at various points above, we can readily imagine similar findings. Future research along those lines is necessary, led by researchers more familiar than ourselves with those contexts. As stated at the outset, our intended contribution herein was both analytic advance and reopening of potential explanations cut short by uninformed, OECD-style explanations.

We might well end the analysis there, but that would still leave open the most crucial question for us: If not shadow education, what drives East Asia’s world leading achievement? If we were to listen to OECD experts and the reports they write, the question itself would never arise: it would already be answered by reference to structural features. But to the degree to which with those explanations are refuted, it become possible to engage with alternative explanations.

A leading scholar of differences in mathematics achievement worldwide based in Hong Kong wrote two decades ago:

...the essential difference between the features of East Asia and the West rest on the different views about **who or what the centre in the teaching and learning process should be**. This is in turn based on fundamental differences in cultural values such as the nature of human beings, the nature of mathematics, etc. (Leung, 2001, p. 47, bold in original).

Although many scholars today would blush at the apparent reification inherent in the notion of “essential” differences and quickly problematize what “culture” might mean, could we accept the thrust of the argument here if we called this complex simply a “worldview”? That is, what if we simply understood “essential differences” as divergences in the deeper perspectives on the nature of knowledge, the meaning assigned to education, and concepts of selfhood? Could we accept it if “culture” was understood as less a fixed and reified entity, more a “mind set”?

Closely examining differences in PISA 2006 Science data, our recent study put forth the idea that structural and institutional factors do not explain observed differences (Komatsu & Rappleye, 2017b). The anomalies between theories espoused by OECD analysts do not square with observed data, suggesting the need for new theories. And it has been well established why and how the OECD needs to avoid “culture” if it is to maintain its legitimacy (Auld & Morris, 2016). In that vein, our study put forth the view that in Japan, and perhaps elsewhere in East Asia, an alternative theory of learning, what we called Type II Learning, was partially responsible for high achievement. By working to remove the usual tropes about East Asia, i.e., students are not creative, under tremendous pressure, education is geared towards exams and economic growth, and here that performance is secretly driven by shadow education, that piece tried to keep open space to think about the alternatives East Asia presents, joining others who have come before (e.g., Li, 2012; Stigler & Hiebert, 1999; Stigler & Stevenson, 1992; Tobin, Hsueh, & Karasawa, 2009; Tobin, Wu, & Davidson, 1989; Takayama, 2015, 2018).

But why aren’t those alternatives more visible already, after decades of research and global comparisons? The reasons are multiple and complex, of course, but the same Hong Kong-based scholar of mathematics cited earlier captures some of the main dynamics very well:

...mathematics education, as a discipline, unlike the disciplines of mathematics and education is a relatively young field of study. If a narrow definition of mathematics education is taken, its root can be traced to the emergence of learning psychology at the beginning of the 20th century. This was exactly the time when East Asian countries were either colonies of or subjected to heavy influence from Western countries. In the area of education, instead of developing a theory of mathematics education of its own, educators in East Asian countries either adopted a Western model of mathematics education or failed to develop any theory of mathematics education at all. Yet even without a theory of its own, teachers in these countries at the classroom level seem to have developed rather distinctive ways of teaching mathematics... (Leung, 2001, p. 37)

The basic pattern described herein, where theories are imported and overlaid upon existing practice at the practical level, finds many similarities around the region. Recent studies too have stated, building on sociological theory developed in Japan, how these attempted “revolutions from above” driven by the repeated import of Western theories have failed to deeply change practice in ways that make it convergent on the West (Rappleye, 2018). Moreover, this pattern leads to considerable confusion about achievement across East Asia: the pedagogy looks ‘backward’ when viewed through Western lenses but it may be precisely that pedagogy, more specifically the underpinning worldviews

and “mind sets” from which it emerges, that is the main reason for world-leading achievement (Komatsu & Rappleye, 2017b). It is crucial here to note that when “culture” is understood as something embedded in context (which it surely is) there is little hope to “learn from” rather than simply “learn about” (Takayama, 2015). But when it is understood as a “mind set”, worldview, and/or as a set of onto-epistemic possibilities available to anyone, these East Asian approaches to achievement become a resource for learning for anyone, which – if sustained over long periods of time – would likely lead to higher levels of achievement. In fact, we already see affirmation of this general approach in, say, the importation of Japanese Lesson Study worldwide.

In light of this, we hope that other scholars will join in these lines of research, as a necessary prerequisite for the more arduous task of developing original theory (i.e., working up what Tobin et al. (2009, p. 242) usefully describe as ‘unmarked beliefs’ and ‘an implicit cultural logic’ into explicit symbolic systems). We feel that keeping open the possibilities to think differently is now an even more urgent and vital task, not least because the statistics the OECD uses and descriptions of the region they present are often factually flawed yet increasingly influential (see also Rappleye et al., 2020; recall Fig. 1). In the face of renewed OECD and Western media attempting to inscribe their own cultural views about education onto the minds of not just the East Asia region, but the wider world, we renew the call for the “push back” not just at the analytical level but also at the deeper level: from *different* onto-epistemic depths. Highlighting these differences should not be mistaken as divisive and reactionary, but instead as desideratum for continued learning and achievement worldwide.

References

- Aizawa, S., Kagawa, M., & Rappleye, J. (Eds) (2018). *High school of all in East Asia: Comparing experiences*. Routledge. <https://doi.org/10.4324/9781315157719>
- Auld, E., & Morris, P. (2016). PISA, policy and persuasion: Translating complex conditions into educational ‘best practice’. *Comparative Education*, 52(2), 202-229. <https://doi.org/10.1080/03050068.2016.1143278>
- Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin*, 66, 423–437. <https://doi.org/10.1037/h0020412>
- Beech, H. (2002). School daze: cramming. bullies. rote learning. East Asia’s schools are failing their students. Big changes are planned, but will they come soon enough? *Time International*, 159(14), 44-57.
- Benesse (2017). *Gakko gaino kyoikuni kansuru chosa 2017 [Survey of Extra-curricular Activities, 2017]*. Benesse.
- Berkson, J. (1938). Some difficulties of interpretation of interpolation encountered in the application of the chi-square test. *Journal of the American Statistical Association*, 33, 526–536.
- Bray, M. (1999). *The shadow education system: Private tutoring and its implication for planners*. UNESCO IIEP. https://bsc-sslm.edu.hku.hk/f/acadstaff/376/Shadow_Education_System.pdf. Accessed 25 January 2019.
- Bloomberg News. (2005). Japan: Crazy for cramming. Bloomberg Business Week. <https://www.bloomberg.com/news/articles/2005-04-17/japan-crazy-for-cramming>. Accessed 25 January 2019.
- Brehm, W., & Silova, I. (2014). Hidden privatization of public education in Cambodia: Equity implications of private tutoring. *Journal of Educational Research Online* 6(1), 94-116.

- Byun, S. Y. (2014). Shadow education and academic success in Republic of Korea. In H. Park & K. Kim (Eds.) *Korean education in changing economic and demographic contexts* (pp. 39-58). Springer. https://doi.org/10.1007/978-981-4451-27-7_3
- Dawson, W. (2010). Private tutoring and mass schooling in East Asia: reflections on inequality in Japan, South Korea, and Cambodia. *Asia Pacific Education Review*, 11, 14-24. <https://doi.org/10.1007/s12564-009-9058-4>
- The Economist. (2011). *Japan's cramming schools: Testing times*. <https://www.economist.com/asia/2011/12/31/testing-times>. Accessed 25 January 2019.
- The Economist. (2015). *Education in South Korea: The crème de la cram*. Available at: <https://www.economist.com/asia/2015/09/19/the-creme-de-la-cram>. Accessed 25 January 2019.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1), 1–26.
- Ellis, P. D. (2010). *The essential guide to effect sizes*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511761676>
- Entrich, S. (2014). Effects of investments in out-of-school education in Germany and Japan. *Contemporary Japan*, 26(1), 71-102. <https://doi.org/10.1515/cj-2014-0004>
- Fox, J. (2008). *Applied regression analysis and generalized linear models*. Sage Publications.
- Jarvis, P. G., & McNaughton, K. G. (1986). Stomatal control of transpiration: Scaling up from leaf to region. *Advances in Ecological Research*, 15, 1-49. [https://doi.org/10.1016/S0065-2504\(08\)60119-1](https://doi.org/10.1016/S0065-2504(08)60119-1)
- Komatsu, H., & Rappelleye, J. (2017a). A new global policy regime founded on invalid statistics? Hanushek, Woessman, PISA, and economic growth. *Comparative Education*, 53, 166-193. <https://doi.org/10.1080/03050068.2017.1300008>
- Komatsu, H., & Rappelleye, J. (2017b). A PISA paradox? An alternative theory of learning as a possible solution for variations in PISA scores. *Comparative Education Review*, 61(2), 269-297. <https://doi.org/10.1086/690809>
- Komatsu, H., & Rappelleye, J. (2018). Is Exam Hell the cause of high academic achievement in East Asia? The case of Japan and the case for transcending stereotypes. *British Education Research Journal*, 44(5), 802-826. <https://doi.org/10.1002/berj.3468>
- Komatsu, H., & Rappelleye, J. (2019). Refuting the OECD-World Bank development narrative: Was East Asia's 'economic miracle' primarily driven by high educational achievement and cognitive skills? *Globalisation, Societies, and Education*, 17(2), 101-116.
- Komatsu, H., & Rappelleye, J. (2020) Reimagining modern education: Contributions from modern Japanese philosophy and practice? *ECNU Review of Education*. <https://doi.org/10.1177/2096531120905197>
- Komatsu, H. & Rappelleye, J. (in review). Rearticulating PISA: From extending the dominant worldview to locating different worldviews. *Globalisation, Societies, and Education*.
- Komatsu, H., Cho, J., Matsumoto, K., & Otsuki, K. (2012). Simple modeling of the global variation in annual forest evapotranspiration. *Journal of Hydrology*, 420-421, 380-390. <https://doi.org/10.1016/j.jhydrol.2011.12.030>
- Komatsu, H., Shinohara, Y., & Otsuki, K. (2015). Models to predict changes in annual runoff with thinning and clearcutting of Japanese cedar and cypress plantations in Japan. *Hydrological Processes*, 29, 5120-5134. <https://doi.org/10.1002/hyp.10520>
- Kuan, P. (2011). Effects of cram schooling on mathematics performance: Evidence from junior high students in Taiwan. *Comparative Education Review*, 55(3), 342-368. <https://doi.org/10.1086/659142>

- Leung, F. (2001). In search of an East Asian identity in mathematics education. *Educational Studies in Mathematics*, 47(1), 35-51. <https://doi.org/10.1023/A:1017936429620>
- Li, J. (2012). *Cultural foundations of learning: East and West*. Cambridge University Press.
- Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253. <https://doi.org/10.1037/0033-295X.98.2.224>
- MEXT. (2008). *Kodomono gakkougaideno gakushukatsudoni kansuru jittaichosa hokoku [A Report on Extra-Curricular Learning]*. http://www.mext.go.jp/b_menu/houdou/20/08/08080710.htm. Accessed 25 January 2019.
- Mori, I. (2015). The effects of supplementary tutoring on students' mathematics achievement in Japan and the United States. ISS Discussion Paper Series. <https://www.iss-u-tokyo.ac.jp/publications/dpf/pdf/f-178.pdf>. Accessed 15 December 2018.
- Morris, P. (2012). Pick 'n' mix, select and project; policy borrowing and the quest for 'world class' schooling: an analysis of the 2010 schools White Paper. *Journal of Education Policy*, 27(1), 89-107. <https://doi.org/10.1080/02680939.2011.596226>
- National Institute for Educational Policy Research. (2018). *Heisei 29nendo zenkoku gakuryoku gakushujokyo chosa kekkaishiryō (todofukenbetsu) [Reports on the results of 2017 Survey of Academic Achievement and Learning Environments]*. <http://www.nier.go.jp/17chousakekkahoukoku/factsheet/17prefecture-City/>. Accessed 25 January 2019.
- NHK Broadcasting Culture Research Institute. (2013). *Chugakusei kokoseino seikatsuto ishiki chosa 2012 [2012 Survey on life and attitudes of secondary students]*. NHK Shuppan.
- Nuzzo, R. (2014). Statistical errors. *Nature*, 506, 150-152. <https://doi.org/10.1038/506150a>
- OECD. (2011). *Education reform in Japan*. OECD Economics Department Working Papers. https://www.oecd-ilibrary.org/economics/education-reform-in-japan_5kg58z7g95np-en. Accessed 25 January 2019.
- OECD. (2012). *Strong performers and successful reformers in education: Lessons from PISA*. Author.
- OECD. (2015). *PISA 2015 results, volume 2*. Author. <http://dx.doi.org/10.1787/9789264267510-en>
- Park, H. (2013). *Re-evaluating education in Japan and Korea: Demystifying stereotypes*. Oxford: Routledge. <https://doi.org/10.1086/675632>
- Postlethwaite, N. (1967). *School organization and student achievement: A study based on achievement in mathematics in twelve countries*. Almqvist and Wiksell. http://ips.gu.se/english/research/research_databases/compeat/Before_1995/FIMS/FIMS_Postlethwaite. Accessed 25 January 2019.
- Prefecture Rankings. (2018). *Juku attendance of primary students*. <https://todo-ran.com/t/kiji/14737>. Accessed 25 January 2019.
- Rapplee, J. (2018). Borrowings, modernity, and de-axialization: Rethinking the educational research agenda for a global age. In A. Yonezawa, Y. Kitamura, B. Yamamoto, & T. Tokunaga (Eds.), *Japanese education in a global age: sociological reflections and future directions* (pp. 53-74). Springer. https://doi.org/10.1007/978-981-13-1528-2_4
- Rapplee, J., & Komatsu, H. (2018). Stereotypes as Anglo-American exam ritual? Comparisons of students' exam anxiety in East Asia, America, Australia, and the United Kingdom. *Oxford Review of Education*, 44(6), 730-754. <https://doi.org/10.1080/03054985.2018.1444598>
- Rapplee, J., & Komatsu, H. (2020). Is knowledge capital theory degenerate? PIAAC, PISA, and economic growth. *Compare*. <https://doi.org/10.1080/03057925.2019.1612233>
- Rapplee, J., Komatsu, H., Uchida, Y., Krys, K., & Markus, H. (2020). 'Better policies for better lives'? Constructive critique of the OECD's (mis)measure of student well-being. *Journal of Education Policy*, 35(2), 258-282. <https://doi.org/10.1080/02680939.2019.1576923>

- Singer, J., & Braun, H. (2018). Testing international education assessments. *Science*, 360(6384), 38-40. <https://doi.org/10.1126/science.aar4952>
- Stevenson, H., & Stigler, J. (1992). *The learning gap*. Simon & Schuster.
- Stigler, J. W., & Hiebert, J. (1999). *The teaching gap*. Free Press.
- Takayama, K. (2015). Provincializing the world culture theory debate: Critical insight from a margin. *Globalisation, Societies, and Education*, 13(1), 34-57. <https://doi.org/10.1080/14767724.2014.967485>
- Takayama, K. (2018). How to mess with PISA: Learning from Japanese kokugo experts. *Curriculum Inquiry*, 42(2), 220-237. <https://doi.org/10.1080/03626784.2018.1435975>
- Tan, C. (2018). *Comparing high performing education systems: Understanding Singapore, Shanghai, and Hong Kong*. Routledge.
- Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Research*, 25, 26–30. <https://doi.org/10.3102/0013189X025002026>
- Thompson, B. (2002). What future quantitative social science research could look like: Confidence intervals for effect sizes. *Educational Researcher*, 31, 25–32. <https://doi.org/10.3102/0013189X031003025>
- TIMSS International Study Center. (1997a). *Mathematical achievement in the primary years: IEA's Third International Mathematics and Science Study (TIMSS)*. Center for the Study of Testing, Evaluation, and Educational Policy, Boston College. <https://timssandpirls.bc.edu/timss1995i/TIMSSPDF/amtimss.pdf>. Accessed 8 November 2019.
- TIMSS International Study Center. (1997b). *Science achievement in the primary years: IEA's Third International Mathematics and Science Study (TIMSS)*. Center for the Study of Testing, Evaluation, and Educational Policy, Boston College. Available at: <https://timssandpirls.bc.edu/timss1995i/TIMSSPDF/astimss.pdf>. Accessed 8 November 2019.
- TIMSS & PIRLS International Study Center. (2016). TIMSS 2015. Available at: <http://timssandpirls.bc.edu/timss2015/>. Accessed 25 January 2019.
- TIMSS & PIRLS International Study Center. (2018a). TIMSS 2015 International Results in Mathematics. Available at <http://timss2015.org/wp-content/uploads/filebase/full%20pdfs/T15-International-Results-in-Mathematics-Grade-4.pdf>. Accessed 25 January 2019.
- TIMSS & PIRLS International Study Center. (2018b). TIMSS 2015 international results in science. Available at: <http://timss2015.org/wp-content/uploads/filebase/full%20pdfs/T15-International-Results-in-Science.pdf>. Accessed 25 January 2019.
- Tobin, J., Wu, D., & Davidson, D. (1989). *Preschool in three cultures*. Yale University Press.
- Tobin, J., Hsueh, Y., & Karasawa, M. (2009). *Preschool in three cultures revisited*. University of Chicago Press.
- Waring, R. H., & Running, S. W. (1998). *Forest ecosystems*. Academic Press.
- Weisman, S. R. (1992). How do Japan's students do it? They cram. *The New York Times*. <https://www.nytimes.com/1992/04/27/world/how-do-japan-s-students-do-it-they-cram.html>. Accessed 25 January 2019.
- Yamada, K. (2014). *Nihonno kyoiku bunnkashiwō manabu [Cultural History of Japanese Education]*. Minerva Shobo.

Appendix

Table A-1
Achievement of regular participants (Japan, England, Hong Kong, Iran, Norway, Singapore, Slovenia, and the United States) in TIMSS (points).

	JPN	ENG	HKG	HUN	IRN	NOR	SGP	SVN	USA
2015									
Math, 4 th	593	546	615	529	431	493	618	520	539
Math, 8 th	586	518	594	514	436	487	621	516	518
Science, 4 th	569	536	557	542	421	493	590	543	546
Science, 8 th	571	537	546	527	456	489	597	551	530
2011									
Math, 4 th	585	542	602	515	431	495	606	513	541
Math, 8 th	570	507	586	505	415	475	611	505	509
Science, 4 th	559	529	535	534	453	494	583	520	544
Science, 8 th	558	533	535	522	474	494	590	543	525
2007									
Math, 4 th	568	541	607	510	402	473	599	502	529
Math, 8 th	570	513	572	517	403	469	593	501	508
Science, 4 th	548	542	554	536	436	477	587	518	539
Science, 8 th	554	542	530	539	459	487	567	538	520
2003									
Math, 4 th	565	531	575	529	389	451	594	479	518
Math, 8 th	570	498	586	529	411	461	605	493	504
Science, 4 th	543	540	542	530	414	466	565	490	536
Science, 8 th	552	544	556	543	453	494	578	520	527
1995									
Math, 4 th	567	484	557	521	387	476	590	462	518
Math, 8 th	581	498	569	527	418	498	609	464	492
Science, 4 th	553	528	508	508	380	504	523	494	542
Science, 8 th	554	533	510	537	463	514	580	514	513

About the Authors

Jeremy Rappleye

Kyoto University

rappleye.jeremy.6n@kyoto-u.ac.jp

ORCID: <http://orcid.org/0000-0002-0717-8253>

Jeremy Rappleye is Associate Professor at Kyoto University, Graduate School of Education. His recent work has focused on developing a mutually moving dialogue on education free of first assumptions common in Anglo-American research circles, as discussed recently in *Reimagining Modern Education* (Komatsu & Rappleye, 2020) and *Comparative Education as Cultural Critique* (Rappleye, 2020).

Hikaru Komatsu

Kyoto University

kmthkr@gmail.com

ORCID: <http://orcid.org/0000-0003-1195-1648>

Hikaru Komatsu is an interdisciplinary environmental studies scholar. After spending more than 15 years in the field of environmental science, he realized the importance of culture for achieving sustainability. He is currently examining how culture is related to sustainability and how culture can be rearticulated for sustainability through education.

education policy analysis archives

Volume 28 Number 67

April 20, 2020

ISSN 1068-2341



Readers are free to copy, display, distribute, and adapt this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives**, the changes are identified, and the same license applies to the derivative work. More details of this Creative Commons license are available at <https://creativecommons.org/licenses/by-sa/2.0/>. **EPAA** is published by the Mary Lou Fulton Institute and Graduate School of Education at Arizona State University. Articles are indexed in CIRC (Clasificación Integrada de Revistas Científicas, Spain), DIALNET (Spain), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), QUALIS A1 (Brazil), SCImago Journal Rank, SCOPUS, Socolar (China).

Please send errata notes to Audrey Amrein-Beardsley at audrey.beardsley@asu.edu

Join **EPAA's Facebook community** at <https://www.facebook.com/EPAAAPE> and **Twitter feed** @epaa_aape.

education policy analysis archives
editorial board

Lead Editor: **Audrey Amrein-Beardsley** (Arizona State University)

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Associate Editors: **Melanie Bertrand, David Carlson, Lauren Harris, Eugene Judson, Mirka Koro-Ljungberg, Daniel Liou, Scott Marley, Molly Ott, Iveta Silova** (Arizona State University)

Cristina Alfaro
San Diego State University

Gary Anderson
New York University

Michael W. Apple
University of Wisconsin, Madison

Jeff Bale
University of Toronto, Canada
Aaron Bevenot SUNY Albany

David C. Berliner
Arizona State University
Henry Braun Boston College

Casey Cobb
University of Connecticut

Arnold Danzig
San Jose State University
Linda Darling-Hammond
Stanford University

Elizabeth H. DeBray
University of Georgia

David E. DeMatthews
University of Texas at Austin

Chad d'Entremont Rennie Center
for Education Research & Policy

John Diamond
University of Wisconsin, Madison

Matthew Di Carlo
Albert Shanker Institute

Sherman Dorn
Arizona State University

Michael J. Dumas
University of California, Berkeley

Kathy Escamilla
University of Colorado, Boulder

Yariv Feniger Ben-Gurion
University of the Negev

Melissa Lynn Freeman
Adams State College

Rachael Gabriel
University of Connecticut

Amy Garrett Dikkers University
of North Carolina, Wilmington

Gene V Glass
Arizona State University

Ronald Glass University of
California, Santa Cruz

Jacob P. K. Gross
University of Louisville
Eric M. Haas WestEd

Julian Vasquez Heilig California
State University, Sacramento

Kimberly Kappler Hewitt
University of North Carolina
Greensboro

Aimee Howley Ohio University

Steve Klees University of Maryland
Jaekyung Lee SUNY Buffalo

Jessica Nina Lester
Indiana University

Amanda E. Lewis University of
Illinois, Chicago

Chad R. Lochmiller Indiana
University

Christopher Lubienski Indiana
University

Sarah Lubienski Indiana University

William J. Mathis
University of Colorado, Boulder

Michele S. Moses
University of Colorado, Boulder

Julianne Moss
Deakin University, Australia

Sharon Nichols
University of Texas, San Antonio

Eric Parsons
University of Missouri-Columbia

Amanda U. Potterton
University of Kentucky

Susan L. Robertson
Bristol University

Gloria M. Rodriguez
University of California, Davis

R. Anthony Rolle
University of Houston

A. G. Rud
Washington State University

Patricia Sánchez University of
University of Texas, San Antonio

Janelle Scott University of
California, Berkeley

Jack Schneider University of
Massachusetts Lowell

Noah Sobe Loyola University

Nelly P. Stromquist
University of Maryland

Benjamin Superfine
University of Illinois, Chicago

Adai Tefera
Virginia Commonwealth University

A. Chris Torres
Michigan State University

Tina Trujillo
University of California, Berkeley

Federico R. Waitoller
University of Illinois, Chicago

Larisa Warhol
University of Connecticut

John Weathers University of
Colorado, Colorado Springs

Kevin Welner
University of Colorado, Boulder

Terrence G. Wiley
Center for Applied Linguistics

John Willinsky
Stanford University

Jennifer R. Wolgemuth
University of South Florida

Kyo Yamashiro
Claremont Graduate University

Miri Yemini
Tel Aviv University, Israel

archivos analíticos de políticas educativas
consejo editorial

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Editores Asociados: **Felicitas Acosta** (Universidad Nacional de General Sarmiento), **Armando Alcántara Santuario** (Universidad Nacional Autónoma de México), **Ignacio Barrenechea**, **Jason Beech** (Universidad de San Andrés), **Angelica Buendía**, (Metropolitan Autonomous University), **Alejandra Falabella** (Universidad Alberto Hurtado, Chile), **Carmuca Gómez-Bueno** (Universidad de Granada), **Veronica Gottau** (Universidad Torcuato Di Tella), **Carolina Guzmán-Valenzuela** (Universidade de Chile), **Antonia Lozano-Díaz** (University of Almería), **Antonio Luzon**, (Universidad de Granada), **María Teresa Martín Palomo** (University of Almería), **María Fernández Mellizo-Soto** (Universidad Complutense de Madrid), **Tiburcio Moreno** (Autonomous Metropolitan University-Cuajimalpa Unit), **José Luis Ramírez**, (Universidad de Sonora), **Axel Rivas** (Universidad de San Andrés), **César Lorenzo Rodríguez Uribe** (Universidad Marista de Guadalajara), **María Verónica Santelices** (Pontificia Universidad Católica de Chile)

Claudio Almonacid
Universidad Metropolitana de
Ciencias de la Educación, Chile

Miguel Ángel Arias Ortega
Universidad Autónoma de la
Ciudad de México

Xavier Besalú Costa
Universitat de Girona, España

Xavier Bonal Sarro Universidad
Autónoma de Barcelona, España

Antonio Bolívar Boitia
Universidad de Granada, España

José Joaquín Brunner Universidad
Diego Portales, Chile

Damián Canales Sánchez
Instituto Nacional para la
Evaluación de la Educación,
México

Gabriela de la Cruz Flores
Universidad Nacional Autónoma de
México

Marco Antonio Delgado Fuentes
Universidad Iberoamericana,
México

Inés Dussel, DIE-CINVESTAV,
México

Pedro Flores Crespo Universidad
Iberoamericana, México

Ana María García de Fanelli
Centro de Estudios de Estado y
Sociedad (CEDES) CONICET,
Argentina

Juan Carlos González Faraco
Universidad de Huelva, España

María Clemente Linuesa
Universidad de Salamanca, España

Jaume Martínez Bonafé
Universitat de València, España

Alejandro Márquez Jiménez
Instituto de Investigaciones sobre la
Universidad y la Educación,
UNAM, México

María Guadalupe Olivier Tellez,
Universidad Pedagógica Nacional,
México

Miguel Pereyra Universidad de
Granada, España

Mónica Pini Universidad Nacional
de San Martín, Argentina

Omar Orlando Pulido Chaves
Instituto para la Investigación
Educativa y el Desarrollo
Pedagógico (IDEP)

José Ignacio Rivas Flores
Universidad de Málaga, España

Miriam Rodríguez Vargas
Universidad Autónoma de
Tamaulipas, México

José Gregorio Rodríguez
Universidad Nacional de Colombia,
Colombia

Mario Rueda Beltrán Instituto de
Investigaciones sobre la Universidad
y la Educación, UNAM, México

José Luis San Fabián Maroto
Universidad de Oviedo,
España

Jurjo Torres Santomé, Universidad
de la Coruña, España

Yengny Marisol Silva Laya
Universidad Iberoamericana,
México

Ernesto Treviño Ronzón
Universidad Veracruzana, México

Ernesto Treviño Villarreal
Universidad Diego Portales
Santiago, Chile

Antoni Verger Planells
Universidad Autónoma de
Barcelona, España

Catalina Wainerman
Universidad de San Andrés,
Argentina

Juan Carlos Yáñez Velazco
Universidad de Colima, México

arquivos analíticos de políticas educativas
conselho editorial

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Editoras Associadas: **Andréa Barbosa Gouveia** (Universidade Federal do Paraná), **Kaizo Iwakami Beltrao**, (Brazilian School of Public and Private Management - EBAPE/FGV), **Sheizi Calheira de Freitas** (Federal University of Bahia), **Maria Margarida Machado**, (Federal University of Goiás / Universidade Federal de Goiás), **Gilberto José Miranda**, (Universidade Federal de Uberlândia, Brazil), **Marcia Pletsch** (Universidade Federal Rural do Rio de Janeiro), **Maria Lúcia Rodrigues Muller** (Universidade Federal de Mato Grosso e Science), **Sandra Regina Sales** (Universidade Federal Rural do Rio de Janeiro)

Almerindo Afonso

Universidade do Minho
Portugal

Alexandre Fernandez Vaz

Universidade Federal de Santa
Catarina, Brasil

José Augusto Pacheco

Universidade do Minho, Portugal

Rosanna Maria Barros Sá

Universidade do Algarve
Portugal

Regina Célia Linhares Hostins

Universidade do Vale do Itajaí,
Brasil

Jane Paiva

Universidade do Estado do Rio de
Janeiro, Brasil

Maria Helena Bonilla

Universidade Federal da Bahia
Brasil

Alfredo Macedo Gomes

Universidade Federal de Pernambuco
Brasil

Paulo Alberto Santos Vieira

Universidade do Estado de Mato
Grosso, Brasil

Rosa Maria Bueno Fischer

Universidade Federal do Rio Grande
do Sul, Brasil

Jefferson Mainardes

Universidade Estadual de Ponta
Grossa, Brasil

Fabiany de Cássia Tavares Silva

Universidade Federal do Mato
Grosso do Sul, Brasil

Alice Casimiro Lopes

Universidade do Estado do Rio de
Janeiro, Brasil

Jader Janer Moreira Lopes

Universidade Federal Fluminense e
Universidade Federal de Juiz de Fora,
Brasil

António Teodoro

Universidade Lusófona
Portugal

Suzana Feldens Schwertner

Centro Universitário Univates
Brasil

Debora Nunes

Universidade Federal do Rio Grande
do Norte, Brasil

Lílian do Valle

Universidade do Estado do Rio de
Janeiro, Brasil

Geovana Mendonça Lunardi

Mendes Universidade do Estado de
Santa Catarina

Alda Junqueira Marin

Pontifícia Universidade Católica de
São Paulo, Brasil

Alfredo Veiga-Neto

Universidade Federal do Rio Grande
do Sul, Brasil

Flávia Miller Naethe Motta

Universidade Federal Rural do Rio de
Janeiro, Brasil

Dalila Andrade Oliveira

Universidade Federal de Minas
Gerais, Brasil