How to Assign Students into Sections to Raise Learning

Ming Ming Chiu
Purdue University
5156 Beering Hall
West Lafayette, IN, 47907
+1-765-496-0119
chiu23@purdue.edu

Bonnie Wing-Yin Chow
City University of Hong Kong
Department of Applied Social Sciences
Kowloon, Hong Kong
+852-3442-8954
wychow@cityu.edu.hk

Sung Wook Joh
Seoul National University
Business School
Seoul, Republic of Korea 151-716
+82-2-880-9384
swjoh@snu.ac.kr

ABSTRACT

Grouping students with similar past achievement together (tracking) might affect their reading achievement. Multilevel analyses of 208,057 fourth grade students in 40 countries showed that clustering students in schools by past achievement was linked to higher reading achievement, consistent with the benefits of customized, targeted instruction. Meanwhile, students had higher reading achievement with greater differences (variances) among classmates' past achievement, reading attitudes, or family SES; these results are consistent with the view that greater student differences yield more help opportunities (higher achievers help lower achievers, so that both learn), and foster learning from their different resources, attitudes and behaviors. Also, a student had higher reading achievement when classmates had more resources (SES, home educational resources, reading attitude, past achievement), suggesting that classmates shared their resources and helped one another. Modeling of non-linear relations and achievement subsamples of students supported the above interpretations. Principals can use these results and a simpler version of this methodology to re-allocate students and resources into different course sections at little cost to improve students' reading achievement.

CCS Concepts

• **Applied Computing → Education**

Keywords

Ability grouping; classmates; inequality; socioeconomic status; international assessment

1. INTRODUCTION

As classmates' interactions with each student differ, their effects differ across students, so some arrangements of students into classrooms yield superior learning overall compared to others. For example, grouping students by ability (*tracking*) is a common but controversial education policy. Past studies of tracking has yielded mixed results (*positive effects*: e.g., [1]; *negative effects*: e.g., [2]; and *non-significant effects*: e.g., [3]). To explain these contradictory results, we propose that the impact of tracking on student achievement differs across levels (across schools vs. within school), differs across types of classmate resources (past achievement, reading attitude, family socio-economic status [SES]) and depends on a student's own academic ability. Armed with this knowledge and methodology, school principals can analyze their data and use evidence-based arrangements of students into course sections for superior learning instead of idiosyncratic systems (e.g., gender balance, parent requests, etc.) [4].

Data or analytic limitations might also account for some conflicting results. While some schools track openly, other schools do so quietly without a public tracking policy, even though students can often recognize different achievement patterns across groups [5]. Hence, we examine the distribution of students by their *actual* past achievement, instead of deferring to incomplete school declarations of tracking [6]. Analytic limitations of past studies include omitted variable bias, multi-collinearity, and failure to model the nested data structure of students within classes within countries [7, 8].

To disentangle the effects of different types of tracking on reading achievement, this study considers different levels of factors and extends past research on tracking in four ways through analyses of 208,057 fourth grade primary school students in 40 countries. (Many primary schools track their students, some as early as first grade [9].) First, we examine how tracking students both across schools (*school-level*) and across classes within schools (*class-level*) affects reading achievement. Second, we examine how both the *amount* and *variation* in different types of classmate resources (*past achievement*, *reading attitude*, *family socio-economic status* [*SES*]) are related to a student's reading achievement. Third, we analyze whether these relations differ across students with *different levels of past reading achievement*. Lastly, we overcome the statistical limitations of past studies through representative sampling, inclusion of central variables, structured sets of variables, and multilevel analyses. By understanding how tracking operates, this study can help explain different tracking effects and inform school policies on allocation of students into different course sections to improve student learning.

2. TRACKING MECHANISMS

Under a *tracking* policy, students with similar past academic achievement are grouped together for instruction. Whether tracking raises or reduces student achievement depends on (a) the impact of student similarities vs. differences and (b) whether classmates compete or share resources.

2.1 Student Similarities vs. Differences

Education administrators can track students by placing those with similar academic competences together, separating them from other

LAK '17, March 13-17, 2017, Vancouver, BC, Canada

DOI: http://dx.doi.org/10.1145/3027385.3027439

students with higher or lower levels of past academic achievement. Or, they may mix together students with different past achievements.

2.1.1 Clusters of similar students

Students with similar academic competences can be assigned together to the same school (*academic or vocational streaming*) or to the same classes within a school (*course-by-course tracking*) [1]. Grouping similar students together can improve their learning by enabling teachers to customize instruction to similar students, capitalizing on student preferences to interact with and help similar peers and enacting self-fulfilling prophecies of their labels and norms [10].

When facing students with similar academic competences, educators can customize the curriculum, lessons, teaching materials, and teaching pace to the needs of each group of similar students, which can improve their learning (*customized instruction*) [11]. In contrast, when the competences of students differ widely, teachers may focus on the learning needs of a subset of students to the detriment of other students with much higher or much lower competences [12]. Such customized instruction is easier for a streamed school with similar students rather than for a tracked class with similar students (but different students across classes) [10]. In a streamed school, teachers design lessons for all students within a small range of competences [4]. In contrast, tracking within a school requires much more teacher time and effort to create lessons for classes of students at different competences [11]. Hence, we expect instruction customization to occur more often and improve student achievement more in streamed schools than in other schools (with or without internal tracking of classes).

As students often prefer to interact with others who are similar to themselves, those with similar academic competences might be more likely than others to interact, befriend, and help one another to learn, compared to dissimilar others (*homophily bias*) [13]. As a result, academically similar students in streamed schools might be more likely to help one another and contribute to a school-wide community culture of mutual support, compared to students in non-streamed schools [14]. (While academically similar classmates in tracked classes within a non-streamed schools might help one another more, they might be less helpful to academically different schoolmates in other classes.)

Although students placed in schools or classes labeled as high-achieving might benefit from self-fulfilling prophecies by enacting high expectations and norms (*assimilation*) [15], students labeled as low-achieving might correspondingly suffer. Teachers and parents of students in schools labeled as high-achieving typically have high expectations of them, which students often internalize [16]. As a result, these students, their teachers and their parents tend to invest more time, effort and other resources to improve their learning, compared to those in non-streamed schools [17].

Assimilation effects might be stronger when tracking within a school rather than streaming across schools. When students are tracked within a school, teachers can devote more attention, effort and other resources to students in higher tracks than to students in lower tracks [18]. This unequal distribution of teacher resources not only increases the gap between high- and low-achieving students but its unfairness can demoralize low-achieving students; as a result, the drop in the academic performance of low-achieving students might exceed the gain in that of high-achieving students, thereby yielding lower overall academic performance [14].

In contrast, the negative effects of assimilation and demoralization might be weaker for students within a streamed school. As academic comparisons with other schools are more distant from students and teachers' immediate experiences, such labels might have less impact on their behaviors, especially as they acclimate to their streamed school environment [14]. (However, such labels can still influence parents and attract teachers to reputable schools in education systems with open markets for hiring teachers [unlike school systems like South Korea that randomly rotate teachers to different schools every five years [14].) Furthermore, students within a school share the same label, so teachers and staff are less likely to treat students differently [18]. As a result, these students are more likely to view their teachers as fair, to be less demoralized, and show higher overall academic performance compared to students in schools with tracked classes.

Hence, streamed schools might have more benefits from customized instruction and homophily, and less harm from assimilation and demoralization, compared to schools with tracked classes. As many schools systems do not explicitly stream their schools, we operationalized the degree of streaming across schools with a *school clustering* by past achievement measure (ratio of student past achievement variance across schools over total variance of past achievement in a country) [18]. Hence, we propose the following hypothesis:

H-1. In education systems with greater school clustering by past achievement, students have higher academic achievement than otherwise.

2.1.2 Mixing different students

While clustering similar ability students together aids instruction customization and capitalizes on homophily, mixing different ability students together can facilitate helping behaviors and learning from schoolmates/classmates with different experiences. Mixed classes offer more possible pairs of a higher-achieving student helping a lower-achieving student (*help opportunities*). Helping benefits both the recipient who receives additional information and explanation, and the giver who often learns more by re-organizing and elaborating his or her knowledge to give a suitable explanation to the recipient [19]. Unlike clustering which operates across schools, help opportunities occur primarily in direct interactions among classmates.

H-2a. When classmates have greater *variance* in their past achievements, a student has higher academic achievement.

Compared to less diverse groups, more diverse groups often have weaker interpersonal relations, more disagreements and less early learning, but their greater range of experiences and resolution of disagreements can increase their later learning [20]. Due to *homophily bias* within a group, members often categorize themselves into subgroups based on similarity (similar to ingroup members and different from outgroup members, *social categorization*) [21]. People trust ingroup members more, cooperate with them more often, and have better relationships with them, compared to outgroup members. Moreover, diverse groupmates often have unfamiliar ideas, attitudes, and experiences [22] that can conflict, thereby igniting disagreements that hinder interpersonal relations initially [21]). Thus, less diverse groups often initially function better than more diverse groups do [20].

However, diverse groups' different ideas and disagreements can legitimize different opinions, thereby stimulating group members to pay attention, share more ideas, and reduce premature consensus [23]. Group members' diverse views also help them recognize flaws and correct them to yield better ideas [24]. If diverse groupmates can reconcile their different views, resolve their disagreements, and integrate their information, they can create and learn new ideas [25]. Lastly, divergent views can stimulate a group to reflect and improve on its own functioning [26]. Thus, over longer time periods, diverse groups produce ideas that are more diverse and learn more compared to homogeneous groups [27]. As these students have shared a classroom for several months, we focus on the long-term effects of diversity. Also, many countries do not have much racial diversity across classmates, so we examine diversity of classmates' family SES and propose the following hypothesis.

H-2c. When classmates have greater *variance* in their family SES, a student has higher academic achievement.

Greater differences among classmates increase the extremes of diametric opposites, which can draw attention to them and aid learning. Consider two classes whose students' reading attitudes have the same mean but greater variance in the second class than the first. As reading attitude extremes are likely greater in the second class than the first [8], the second class likely has both the student with the best reading attitude (let us call her Heidi) and the one with the worst (Lola). Heidi's concrete behaviors embody her reading attitude (e.g., reads many books; tells their stories by acting them) and yield beneficial consequences (high reading quiz scores; smiles at her graded quizzes; teacher praises her, reading awards, etc.). In contrast, Lola rarely reads ("I never read books –they're boring"), has low reading quiz scores, frowns at her graded quizzes and so on.

As extremes, Heidi and Lola mutually highlight their differences (*contrasting cases*), thereby attracting and focusing more attention on the two of them than on either one alone or on classmates with smaller reading attitude differences (*focusing function*, Schwartz & Bransford, 1998). This focusing function helps classmates include important differences and omit unnecessary information from their working memories (*cognitive load theory*) [28], which are smaller in young children than in adults [29]. As contrasting cases of people rather than abstract ideas, Heidi and Lola serve as detailed reference points for inferences about students in the continuum between them (*cognitive reference point reasoning*) [30]. Rather than observing other students or either Heidi or Lola alone, a student who focuses on both of them can create a network of their contrasting reading attitudes and related information in his or her short-term working memory (*hippocampus*) and then encode it into his or her long-term memory (*cerebral cortex*) [29], thereby learning effectively and efficiently about reading attitude, remembering it reliably, and acting on it accordingly to learn more than otherwise [31]. Hence, we propose this hypothesis.

H-2b. When classmates have greater *variance* in their reading attitudes, a student has higher academic achievement.

2.2 Classmates Compete vs. Share Resources

The impact of tracking on student achievement also depends on the extent to which classmates compete or share resources. When competing with classmates with greater cognitive, social and material resources in a zero-sum game, a student could have lower academic achievement. Classmates can serve as a collective ruler against which to judge a student's relative competence (*comparative reference-group view*). When surrounded by lower-achieving classmates, a student often has greater confidence in his or her competence (self-concept), expectation of future success, motivation, and academic achievement (*social comparison*) [32]. Conversely, higher-achieving classmates can demoralize a student, reduce his or her self-concept, lower future expectations, and yield lower academic achievement.

H-3. When classmates have a higher *mean* past academic achievement, a student has lower academic achievement.

In this social comparison view, tracking benefits lower-achieving students at the expense of higher achieving students.

On the other hand, classmates, especially high-achieving ones, can help a student learn directly or indirectly [33]. Classmates can directly help a student by sharing information. For example, a higher-achieving classmate can help a student correctly spell an unfamiliar word.

A classmate can also help students learn indirectly through motivation or norms. For example, a classmate can dramatically enact a scene from a storybook, which can entice and motivate other students to discuss it and learn about it [33]. Over time, students' greater motivation helps them exert more effort and persevere when facing difficulties [34].

Classmates, especially higher-achieving ones with higher status, can help create and maintain norms of positive attitudes toward reading, regular learning behaviors and high reading achievement [35]. Classmates can articulate and model positive academic attitudes, such as sharing their enjoyment of specific stories. Furthermore, they can discuss their readings daily to promote peer pressure toward regular reading of new books. Together, classmates can cultivate a culture of positive reading attitudes and behaviors in which to immerse a student, which typically yields higher reading achievement [36]. Note that hypothesis H-4a competes with H-3 above.

H-4a. When classmates have a higher *mean* past academic achievement, a student has higher academic achievement.

H-4b. When classmates have a higher *mean* attitude toward reading, a student has higher academic achievement.

A student can benefit not only from family resources but also from classmates' family resources. Families can use their financial, human, cultural and social capital to give their children learning opportunities. Specifically, families with more money (*financial capital*) can buy more educational resources (books, calculator, etc.) to create a richer learning environment [18]. Furthermore, high SES students often spend more time with their parents (due to less parent time on housework and multi-tasking parents), so they can benefit more from their parents' human, social and cultural capital. Families with more education, knowledge or skills (*human capital*) often create better learning environments for their children, foster better attitudes toward reading and teach them more skills compared to other families [18]. Likewise, high SES families often have cultural possessions or experiences (*cultural capital*) that can help their children learn society's cultural knowledge, skills and values to adapt to their school culture [34]. High SES families also often have large social networks of relatives, friends and acquaintances with skills or resources (*social capital*) that can help their children learn [35]. Using their greater financial, human, social and cultural capital, higher SES students can better understand others' expectations, behave properly at school, have closer relationships with teachers and classmates, and learn more in school than lower SES students do.

Similarly, a student can benefit from a classmate's family resources directly or indirectly [35]. In the most direct case, a student visits a classmate's home and uses the latter's educational resources. A student may work with a classmate on the classmate's computer (classmate family financial capital), discuss a book with the classmate's mom (classmate family human capital), discuss a painting in the living room (classmate family cultural capital) or chat with a family friend over dinner (classmate family social capital). Less directly, a student can learn from a classmate' learning experiences at home [35], for example, when the classmate talks about it ("my mom was telling me about the presidential election").

H-4c. When classmates have a higher *mean* family SES, a student has higher academic achievement.

Tracking provides higher-achieving students with higher-achieving classmates who often have more intellectual resources. However, tracking provides low-achieving students with low-achieving classmates who often have more intellectual resources. Hence, tracking might benefit high achieving students more from low-achieving students.

2.3 The Present Study

This study examines the effects of tracking on students' reading achievement through analyses of 208,057 fourth grade primary school students in 40 countries. We focus on three research questions. First, does tracking at school- or class-levels linked to better reading achievement? Second, do these links operate through the mechanisms of *customized instruction, homophily*, *giving help*, *receiving help*, *diversity*, *competition/social comparison*, *sharing resources* or *superior benefits for high-achieving students*? Lastly, do these relations differ across subsamples of the lowest-achieving 10%, 20% and 50% of students and the highest-achieving 50%, 20% and 10% of students?

As past studies have shown that reading achievement is related to the following variables, we included them in our regression model to reduce *omitted variable bias* [8]: country economic growth (gross domestic product per capita) [18], family income inequality in a country (Gini index) [18], family SES [35], home educational resources [18], student gender [36], reading self-concept [38], attitude toward reading [35], parent attitude toward reading [35], school climate [39], and teacher gender [40].

3. METHODS

3.1 Data

In 40 countries, the International Association for the Evaluation of Education Achievement Progress in International Reading Literacy Study (IEA-PIRLS) assessed 208,057 fourth-grade students, and their parents, teachers and school principals completed questionnaires [41]. IEA chose at least 150 representative schools in each country, based on neighborhood SES and student intake. From each school, IEA selected one or two 4^{th} grade classes, yielding a sample size of at least 4,000 students per country (stratified sampling) [41]. Participating students completed an 80-minute assessment booklet and then a 15–30-minute questionnaire. The World Bank collected economic data for each country (annual income and income inequality) [42].

3.2 Variables

Table 1. Summary Statistics (N = 208,044 Students)

Variable	Mean	SD	Min	Max
Reading achievement	496	113	5	813
Past reading achievement before school (Recognize alphabet; Read words; Read sentences; Write alphabet; Write words.) Reliability = 0.95.	0.00	1.00	-2.49	1.82
Log (income per person) Log Gross Domestic Product per capita	9.73	0.65	7.97	10.5
Inequality of Family income (Gini index)	35.44	7.90	25	58
School inequality based on students' past reading achievement; Ratio of variance across schools / country variance	0.12	0.09	0.05	0.54
Family socio-economic status (parents' educations, jobs, incomes); Reliability = 0.94.	0.00	1.00	-2.95	2.84
Parents' attitude towards reading; Reliability = 0.82.	0.00	1.00	-2.94	1.69
Home education resources (books; children books; computer; child desk/table); Reliability = 0.76.	0.01	1.00	-2.54	1.98
Class mean SES	0.00	0.61	-2.72	2.18
Class mean home education resources	0.01	0.68	-2.54	1.98
Class mean parents' attitude towards reading	0.00	0.37	-2.67	1.41
School violence (bullying by me and of me; injury by me and of me) Reliability = 0.93	0.00	1.00	-1.37	2.09
Female teacher (vs. male)	0.84			
HW time mismatch (in teacher and student reported times)	0.89	0.61	0.00	8.19
Girl (vs. boy)	0.49			
Student's reading attitude Reliability = 0.79	0.00	1.00	-3.00	1.59
Student's reading self-concept Reliability = 0.80	0.00	1.00	-2.77	1.54
Class mean past reading achievement	0.00	0.48	-2.49	1.66
Class mean attitude towards reading	0.00	0.41	-1.66	1.59
Past reading achievement -class variance	0.80	0.35	0.00	3.89
SES -class variance	0.66	0.31	0.00	3.35
Students' attitude towards reading -class variance	0.87	0.35	0.00	3.54

Note: Indices were standardized (mean = 0; SD = 1). All reliabilities refer to the composite score reliability coefficient.

3.3 Analysis

The outcome variable $\mathbf{Reading}_{ijk}$ of student i in school j in country k has a grand mean intercept β_{000}, with unexplained components (*residuals*) at the student-, school-, and country-levels (e_{ijk}, f_{jk}, g_k). First, we enter past student achievement (*Past_Literacy_Skills*).

$$\mathbf{Reading}_{ijk} = \beta_{000} + e_{ijk} + f_{0jk} + g_{00k} + \beta_{rjk}\mathit{Past\_Literacy\_Skills}_{ijk} + \beta_{00s}\mathbf{Country}_{00k} + \beta_{tjk}\mathbf{Family}_{ijk} + \beta_{ujk}\mathbf{Classmates\_Families}_{ijk} + \beta_{vjk}\mathbf{School\&Teacher}_{ijk} + \beta_{wjk}\mathbf{Student}_{ijk} + \beta_{xjk}\mathbf{Classmates}_{ijk} + \beta_{zjk}\mathbf{Classmates\_Variance}_{ijk} \quad (1)$$

Then, we entered a vector of s variables at the country level: GDP per capita, Gini index, school clustering of students by past reading achievement before school (**Country**, see table 1) to test whether more school clustering is linked to higher reading achievement (hypothesis H-1). To test for non-linear effects, squared terms (e.g., $Gini^2$) were added. We tested whether sets of predictors were significant with a nested hypothesis test (χ^2 log likelihood) [8]. Non-significant variables were removed. Next, we added family variables: family SES, parents' attitude towards reading and home education resources (**Family**). Then, we added class means of family SES, home education resources, and parents' attitudes toward reading (**Classmates_Families**) to test whether higher reading achievement is linked to higher classmate SES (H-4c). Next, we added school and teacher variables: school violence, female teacher, and homework mismatch (**School&Teacher**), followed by student variables: gender, student's attitude towards reading, student's reading self-concept (**Student**). Then, we added the class mean of past reading achievement and class mean of attitude towards reading (**Classmates**). This tests the competing hypotheses of whether higher classmates' past achievement is linked to a student's *lower* reading achievement (H-3) or *higher* reading achievement (H-4a). It also tests whether classmates with superior reading attitudes are linked to a student's higher reading achievement (H-4b).

Lastly, we added the class variances of past student achievement, SES, and students' attitude towards reading (**Classmates_Variance**) to test whether higher academic achievement is linked to higher variance in classmate past achievement (H-2a), to higher variance in classmate reading attitude (H-2b) or to higher variance in classmate SES (H-2c).

To determine whether high-achieving vs. low-achieving students have greater access to or benefits from available resources, we tested if the above links differed across student sub-samples. We created six sub-samples of students by reading achievement (bottom 10%, bottom 20%, bottom 50%, top 50%, top 20%, and top 10% from each country).

4. RESULTS

4.1 Summary Statistics

This sample's countries ranged from poor, very unequal nations (e.g., Iran) to rich, relatively equal ones (e.g., Luxembourg). See Table 1 for overall summary statistics.

4.2 Explanatory Model

Past reading achievement, country, family, school, teacher, student, classmates and class variance variables accounted for differences in students' reading achievement (see Table 2). The differences in reading achievement at the country-, school-, and student-levels were 58%, 13%, and 29%. All results discussed below describe first entry into the regression, controlling for all previously included variables.

4.2.1 Past Reading Achievement

Students whose past reading achievement exceeded the mean by 10% averaged 5 more points in reading achievement (Table 2, model 1). Past reading achievement accounted for 2% of the differences in students' reading achievement.

4.2.2 Country

Meanwhile, students in richer or more equal countries scored higher (see Table 2). If a richer country's GDP per capita exceeded the country mean by 10%, its students averaged 4 more points in reading achievement. (Regressions with linear GDP per capita did not fit the data as well, explaining less of the variance in students' reading achievement.) Meanwhile, when a country's GINI exceeded the mean by 10%, its students averaged 20 points lower in reading.

If a country's clustering of students by past reading achievement exceeded the mean by 10%, students averaged 45 more points in reading achievement, supporting hypothesis H-1 (instruction customization). Furthermore, school clustering showed a non-linear effect (see Figure 1). The effect size increases up to a maximum when school clustering is 0.30 (~95th percentile of a normal curve or 2.0 SDs above the mean; 2.0 = [0.30 – 0.12] / 0.09 = ([X – mean] / SD). Extreme school clustering beyond two standard deviations (> 0.30) tends to reduce student reading achievement. Together, country variables accounted for about 29% the total differences in reading achievement and for 51% of its variance across countries.

Table 2. Multilevel Regression of Students' Reading Achievement with Unstandardized Coefficients (and Standard Errors in Parentheses)

	Reading achievement		
Explanatory variable	Beta	SE	p
Past reading achievement before school	9.05	(0.12)	***
Log GDP per capita	25.03	(3.34)	***
GINI	-7.03	(0.13)	***
Clustering students by past reading achievement²	-1913.59	(118.78)	***
Clustering students by past reading achievement	1100.19	(62.99)	***
Family SES	8.07	(0.15)	***
Home education resources	8.87	(0.15)	***
Parents' attitude towards reading	5.29	(0.12)	***
Class mean SES²	-5.47	(0.61)	***
Class mean SES	20.95	(0.73)	***
Class mean home education resources²	1.58	(0.57)	**
Class mean home education resources	22.37	(0.83)	***
Class mean past reading achievement2	14.97	(0.67)	***
Class mean past reading achievement	10.36	(0.73)	***
School violence	-3.43	(0.11)	***
Female teacher	4.02	(0.61)	***
HW mismatch	-5.29	(0.37)	***
Girl	9.60	(0.23)	***
Students' attitude towards reading	8.25	(0.12)	***
Students' reading self-concept	21.57	(0.11)	***
Class mean students' reading attitude	9.88	(0.69)	***
SES -class variance	2.50	(0.77)	**
Past reading achievement -class variance²	-5.64	(0.94)	***
Past reading achievement -class variance²	13.46	(2.09)	***
Students' reading attitude -class variance²	-8.00	(1.13)	***
Students' reading attitude -class variance	22.47	(2.37)	***
Variance at each level	Variance explained		
Country (58%)	0.64		
School (13%)	0.43		
Student (29%)	0.29		
Total variance explained	0.51		

Note. A constant term was added. *$p < .05$, **$p < .01$, ***$p < .001$.

4.2.3 Family

Family variables were linked to students' reading achievement (see Table 2). If parents have 10% better reading attitude, their

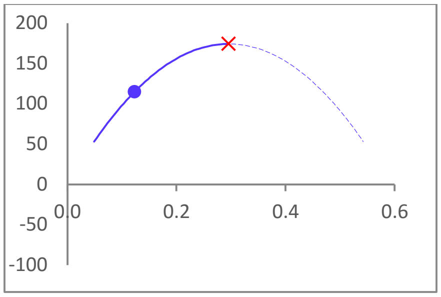

Figure 1. School Clustering by Past Achievement. All figures' graphs have the same reading achievement y-axis scale of -100 to 200. Each blue dot indicates an explanatory variable's mean value, and the solid blue curve indicates values within two standard deviations of this mean. The dashed curves indicate extreme values outside two standard deviations of this mean. Some curves have a turning point (minimum or maximum), indicated by a red X.

children averaged 2 points higher in reading. Meanwhile, students with 10% more home education resources averaged 4 points higher in reading. Family variables accounted for an extra 7% of the variance in students' reading achievement.

4.2.4 *Classmates' Families*

Classmate family variables were also linked to students' reading achievement (see Table 2). If classmate family SES exceeded the mean by 10%, students averaged 4 points higher in reading achievement, supporting the classmate sharing hypothesis H-4c, and this non-linear effect decreases for higher classmate family SES (*diminishing marginal returns*; see Figure 2) [18].

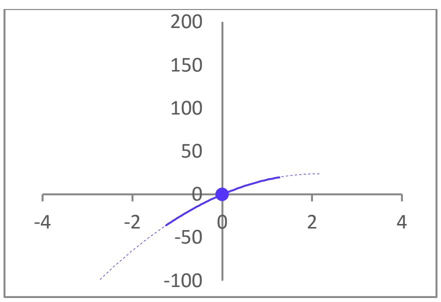

Figure 2. Class Mean SES

Students with classmates who had 10% more education resources at home averaged 5 points higher reading achievement. (This technically non-linear effect was nearly linear, see Figure 3.)

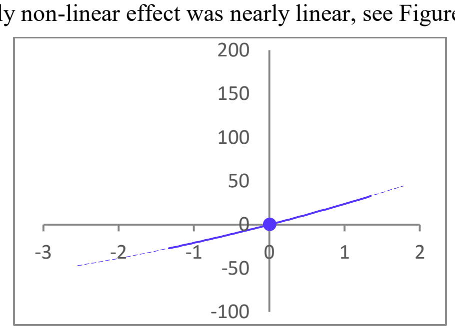

Figure 3. Class Mean Home Education Resources

4.2.5 *School and teacher*

School and teacher variables were linked to students' reading achievement (see Table 2). Students in schools with 10% less school violence averaged 2 points higher in reading, and students with a female teacher averaged 5 points higher than those with a male teacher. Also, if teachers' misjudgment of students' homework time exceeded the mean by 10%, their students averaged 1 point lower in reading. School and teacher variables accounted for an extra 1% of the variance in students' reading achievement.

4.2.6 *Student*

Girls outscored boys by 13 points on average (Table 2, model 6). Moreover, students with 10% better reading attitude or 10% higher reading self-concept averaged 3 or 6 points higher in reading, respectively (see Table 2). Student variables accounted for an extra 4% of the variance in students' reading achievement.

4.2.7 *Classmates*

When classmates' past reading achievement exceeded the mean by 10%, students averaged 2 points higher reading achievement, supporting hypothesis H-4a (classmates share) and rejecting hypothesis H-3 (classmates compete). See Table 2. This nonlinear effect has a minimum when classmate past achievement is -0.30 (~10th percentile at -0.63 SD (below the mean); -0.63 = [-0.30 – 0] / 0.48); at lower values beyond this turning point, student reading achievement scores tend to be higher (see Figure 4). Together, classmates' family variables accounted for an extra 6% of the variance in students' reading achievement.

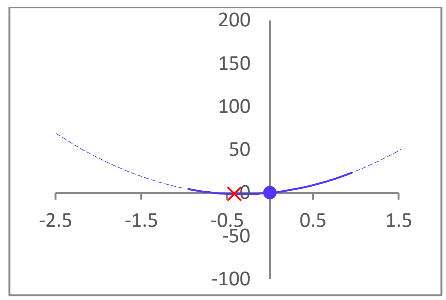

Figure 4. Class Mean Past Reading Achievement

When classmates' reading attitudes exceeded the mean by 10%, students averaged 1 point higher in reading (see Table 2), supporting hypothesis H-4b; the link is slightly stronger for higher classmate reading attitudes (see Figure 5). This classmate variable accounted for an extra 0.1% of the variance in students' reading achievement.

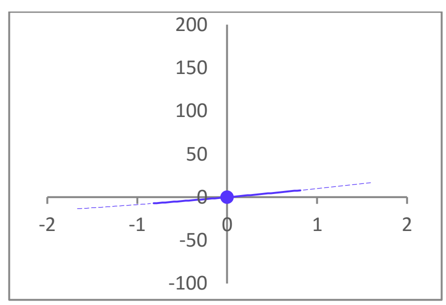

Figure 5. Class Mean Reading Attitude

4.2.8 *Classmate variance*

Classmate variances were also linked to student reading achievement (see Table 2). When variance of classmates' past acheivement exceeded the mean by 10%, students averaged 1 point higher in reading achievement, supporting hypothesis H-2a (giving and receiving help). The strength of this link increases up to a maximum at a variance of 1.19 (~86th percentile for a normal curve or 1.11 standard deviations above the mean; 1.11 = [1.19 – 0.80] / .35 = ([X – mean] / SD). See Figure 6. Extreme values at greater

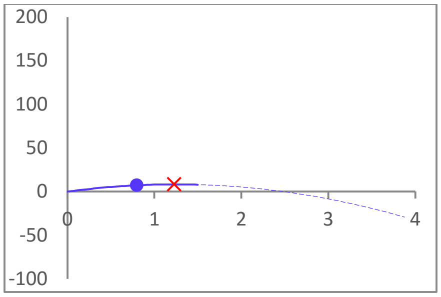

Figure 6. Past Achievement – Class Variance

variances tend to reduce students' reading achievement. Also, students in classes with 10% more variance in classmates' reading attitudes averaged 2 points higher in reading, supporting hypothesis H-2b (attitude diversity). The strength of this link increases up to a maximum at a variance of 1.40 (~87th percentile for a normal curve or 1.51 standard deviations above the mean; 1.51 = [1.40 – 0.87] / .35 = ([X – mean] / SD)). See Figure 7. Lastly, students in classes with 10% more SES variance averaged 0.2 points higher in reading achievement, supporting H-2c (SES diversity). Class variances accounted for an extra 1% of the variance in students' reading achievement. Other variables were not significant.

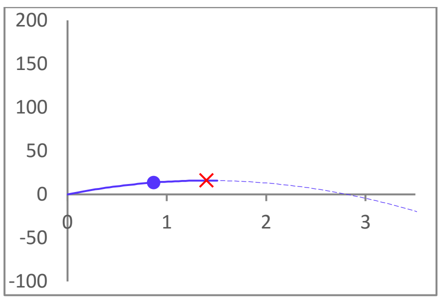

Figure 6. Reading Attitude – Class Variance

4.3 Achievement Sub-samples

Many variables showed similar, significant results in all achievement sub-samples (see Table 3): past reading achievement, log GDP per capita, GINI, clustering of students by past reading achievement, parents' attitude towards reading, home education resources, classmates' mean SES, classmates' home education resources, school violence, students' gender, students' reading self-concept, and class variance in students' attitude towards reading. (The consistency of non-linear results can be seen by graphing the quadratic functions. Due to space considerations, these graphs are not included but are available upon request.)

However, the achievement subsample results varied for the following attributes: teacher gender, homework mismatch, and many classmate-related variables (SES variance; mean and variance of their parents' reading attitudes; mean and variance of past achievement; and reading attitude). Teacher gender and homework mismatch were not significant for the top 10% of students in each country, suggesting that these factors do not affect the highest-achieving students (perhaps because higher ability students have successful, resilient ways of learning even in less hospitable environments). Furthermore, variance in classmates' SES was significant for only the top 50% and top 20% of students by achievement, suggesting that these students primarily benefit from other students' diverse resources and experiences; the results are not significant for the other subsamples.

Also, the classmate reading attitudes results for the lowest-achieving 10%, 20% and 50% of students were similar to the overall result. However, the results were only significant for the lowest-achieving 10%, 20% and 50% of students. This result suggests that only the lower achieving students benefit from more classmates with better attitudes toward reading.

For many subsamples of students, the links between mean and variance of classmate past achievement and student reading achievement were similar to the overall result. However, lower mean classmate past achievement was linked to higher student reading achievement for the lowest-achieving 10% and highest-achieving 20% and 10% of students, which is consistent with the classmate help hypothesis. Meanwhile, classmates' past achievement variance is positively related to reading achievement for the lowest-achieving 10%, 20% and 50% of students but negatively related to reading achievement for the highest-achieving 20% and 10% of students and not significant for the highest-achieving 50% of students. We discuss these results below.

None of the above results show substantial homophily or labeling effects. Analyses using standardized scores within each country yielded similar results. Analyses of residuals showed no influential outliers.

5. DISCUSSION

We examined the effects of tracking on students' reading achievement through analyses of 208,057 fourth grade primary school students in 40 countries. There are five major findings. First, tracking across schools (streaming) was linked to higher reading achievement. Specifically, school clustering of students by past reading achievement yielded higher reading achievement, supporting the instruction customization hypothesis. Second, a student generally benefited from classmates' resources (SES, home educational resources, reading attitude, past achievement), suggesting that classmates shared their resources and this sharing of resources yielded higher academic achievement. Third, tracking at the class level was linked to lower reading achievement. Greater variance of classmate family SES, reading attitude or past achievement were all linked to higher reading achievement, supporting the helping and diversity hypotheses. Fourth, several relations were non-linear with turning points (minimums or maximums): school clustering by past reading achievement, mean and variance of classmate past achievement, and variance of classmate reading attitude. Fifth, some relations differed across achievement subsamples, notably classmate reading attitude, classmate past achievement, variance of classmate SES, and variance of classmate past achievement.

5.1 School Clustering

The school clustering by past reading achievement results likely reflect instruction customization rather than homophily. As homophily was not evident in the classroom-level results, it likely does not account for the broader, school clustering result. Instead, the school clustering and classmate results are consistent with the view that instruction customization is likely more efficient and effective at the school level (streaming) than at the class level. They suggest that greater school clustering of students by past reading achievement –up to about two standard deviations above the mean –might help educators target curricula and instruction to specific schools of students with similar academic competence to help them learn more. The lower reading achievement of students in education systems with extremely high clustering (more than two standard deviations above the mean) suggest that classmates with extremely similar past reading achievement can be harmful; instead, moderate classmate variation is useful, consistent with the classmate-level results.

Table 3. Multilevel Regressions of Reading Achievement with Six Subsamples by Achievement (Bottom 10%, 20%, and 50%; and Top 10%, 20%, and 50%).

	Regressions of Reading Achievement with Six Achievement Subsamples					
	Bottom			Top		
Explanatory variable	10%	20%	50%	50%	20%	10%
Past reading achievement	2.70 ***	3.18 ***	4.72 ***	3.83 ***	2.22 ***	1.93 ***
Log GDP per capita	19.89 ***	17.43 ***	12.17 ***	41.22 ***	40.11 ***	34.20 ***
GINI	-2.53 ***	-3.52 ***	-3.61 ***	-6.29 ***	-4.94 ***	-3.15 ***
Clustering of students by past reading achievement²	-1443.74 ***	-1493.31 ***	-1419.10 ***	-1878.35 ***	-1620.42 ***	-1425.54 ***
Clustering of students by past reading achievement	480.53 ***	524.28 ***	574.52 ***	1122.71 ***	900.99 ***	786.00 ***
Family SES	2.20 ***	2.46 ***	3.64 ***	4.26 ***	2.54 ***	1.77 ***
Home education resources	2.11 ***	3.21 ***	7.08 ***	4.09 ***	3.17 ***	3.08 ***
Parents' attitude towards reading	0.61 **	1.00 ***	1.73 ***	1.67 ***	0.43 **	0.45 *
Class mean SES²	0.37	0.21	-1.75 ***	0.87 *	3.13 ***	2.45 ***
Class mean SES	7.75 ***	7.69 ***	6.98 ***	10.83 ***	4.71 ***	2.18 **
Class mean home education resources²	-0.06	0.56	1.21 *	-1.74 ***	-3.71 ***	-3.59 ***
Class mean home education resources	8.57 ***	11.15 ***	14.21 ***	10.07 ***	8.35 ***	7.50 ***
Class mean past reading achievement²	2.25 **	3.17 ***	4.59 ***	5.20 ***	-0.87	-1.74 **
Class mean past reading achievement	-4.21 ***	-1.79 *	2.78 ***	1.15 *	-2.02 ***	-0.89
School violence	-0.55 *	-0.69 ***	-1.90 ***	-1.93 ***	-1.15 ***	-1.29 ***
Female teacher	1.97 **	1.90 ***	3.58 ***	2.13 ***	1.71 ***	-0.71
HW mismatch	-3.85 ***	-4.59 ***	-4.77 ***	-1.53 ***	-0.82 *	-0.24
Girl	1.15 **	2.63 ***	4.67 ***	1.90 ***	1.25 ***	1.21 ***
Students' attitude towards Reading	1.00 ***	1.50 ***	2.44 ***	3.92 ***	1.83 ***	0.93 ***
Students' reading self-concept reading	2.79 ***	5.35 ***	9.24 ***	9.22 ***	5.83 ***	4.15 ***
Class mean students' attitude towards reading²	0.29	0.03	1.49	0.05	0.57	-1.49
Class mean students' attitude towards reading	6.70 ***	7.25 ***	8.63 ***	0.22	-0.03	-0.12
SES -class variance	0.08	0.64	0.62	1.44 *	2.62 ***	1.19
Past reading achievement -class variance²	-7.57 ***	-7.31 ***	-9.95 ***	1.51	1.33	1.37
Past reading achievement -class variance	16.56 ***	14.61 ***	20.04 ***	-2.65	-3.90 *	-5.61 **
Variance of classmates' attitude towards reading²	-9.91 ***	-8.76 ***	-7.52 ***	-2.64 **	-2.17 *	-1.45
Variance of classmates' attitude towards reading	27.03 ***	21.89 ***	20.78 ***	8.06 ***	6.37 **	5.60 *
Variance at each level			Variance Explained			
Country	0.62	0.63	0.63	0.63	0.64	0.64
School	0.37	0.39	0.39	0.42	0.42	0.43
Student	0.28	0.28	0.29	0.37	0.37	0.37
Total variance explained	0.49	0.50	0.50	0.53	0.53	0.53

Note. Each regression included a constant term. **p* < .05, ***p* < .01, ****p* < .001.

5.2 Classmate Resources

The classmate mean results show that a student generally benefited from classmates' sharing of resources (social capital [35]). A student whose classmates had higher family SES, home educational resources, reading attitudes, or past achievements often had higher reading achievement than other students. Classmate family SES and home educational resources both had positive relations with student reading achievement, consistent with the view that classmates shared educational materials, ideas and experiences with a student to aid their reading achievement [35]. Classmate family SES's diminishing marginal returns show that it benefits lower SES students more than higher SES students; hence, mixing high SES students with low SES students maximizes the classmate SES's benefits [14]. If SES is highly correlated with student past achievement in a country however, we cannot both track students into schools by achievement and ensure that all students have some classmates with high SES.

Meanwhile, classmates with better reading attitudes primarily benefited low-achieving students, showing no significant effect on high-achieving students. This result suggests that classmate reading attitude has a ceiling effect, helpful for low-achieving students who might have poorer reading attitudes but not for high-achieving students who may already have sufficiently positive attitudes toward reading (reading attitude has a weak positive correlation with reading achievement in these data, $r = .12$). Hence, assigning some students with positive reading attitudes into the same classes as low-achieving students might improve the latter's reading achievement without harming the former.

5.3 Tracking across Classes

Very high or very low classmate past achievement was linked to higher student reading scores, supporting the hypothesis that students learn more by receiving help from higher-achieving classmates or by giving help to lower-achieving classmates [19]. The achievement subsamples further showed that lower mean classmate past achievement was linked to higher student reading achievement for the highest-achieving 10% and 20% of students and for the lowest-achieving 10% of students. As a high-achieving student is less likely than other students to ask for help from their classmates, she can learn more with more low-achieving classmates, to whom she can give help (more help opportunities). On the other hand, a very low-achieving student benefits from classmates with lower reading achievement, perhaps because they are more likely to understand his or her learning difficulties and help appropriately (Vygotsky's zone of proximal development [43]). Or, a low-achieving student might have more help opportunities with still lower-achieving classmates. Together, these results suggest that a variety of low- and high-achieving students in a classroom creates more helping opportunities that can aid the learning of both givers and receivers of help.

The results involving classmate variances (reading attitude, family SES, or past achievement) do not support tracking across classes (within school). Greater variance in classmate reading attitude was linked to greater reading achievement, consistent with the view that students learn from observing the consequences of classmates with different reading attitudes and appreciating the importance of desirable ones. Variance in classmate reading attitude also showed diminishing marginal returns, indicating that some variation in classmate reading attitude is sufficient to realize much of the benefits and that maximizing variation is not necessary.

While higher classmate family SES generally benefits all students, greater variance in classmate family SES is linked to greater reading achievement only for the highest-achieving 50% and 20% of students, with no significant effects on other students. Hence, mixing students with different family SES together might benefit these high-achieving students without harming other students.

Lastly, the link between variance in classmate past achievement and reading achievement differs across achievement subsamples. For the lowest-achieving 10%, 20% and 50% of students, greater variance in classmate past achievement is linked to higher reading achievement. This result supports the view that a greater variety of classmates across achievement levels offers more opportunities for a low-achieving student to give and receive help to learn more.

Meanwhile, the highest-achieving 20% and 10% of students have higher reading scores when their classmates have similar, low past achievement levels. Perhaps, a high achieving student can elaborate his or her knowledge with a helpful explanation to a lower achieving student, but has difficulty creating several explanations for students at different, lower-ability levels –which might be frustrating, time-consuming or harmful to his or her learning. Meanwhile, variance in classmate past achievement is not significantly related to reading achievement for the highest-achieving 50% of students, possibly because they generally face sufficient variation in classmate ability to both give and receive help. Together, these results suggest placing high-achieving students with a limited range of low-achieving classmates.

This study not only provides general advice for assigning fourth grade students to classrooms but also showcases a methodology for principals to use on their own school data to customize assignment of their own students. While this analysis uses data from many countries and schools, a principal can analyze his or her own school data across time for assigning his or her students.

6. CONCLUSION

This study investigated the links between grouping students by ability and their reading achievement, suggesting possible mechanisms and showing how it might influence the learning of students at different levels of reading ability. Greater clustering of students by past reading achievement in schools is linked to greater reading achievement, suggesting that streaming across schools can be effective. However, tracking at the class level was not linked to greater reading achievement. Instead, greater variances in classmates' family SES, in classmates' attitudes toward reading and in classmates' past reading achievement were linked to greater reading achievement. Moreover, when classmates had greater family SES, home education resources, attitudes toward reading or past reading achievement, a student had higher reading achievement. This result suggests that classmates shared educational resources, attitudes, ideas and experiences to help a student learn. The nonlinear links between several factors (especially school clustering and classmate past achievement) and reading achievement, and the results of the achievement subsample analyses further support these interpretations. These results suggest low-cost education policies, namely re-allocation of students and resources, for improving students' reading achievement.

7. ACKNOWLEDGMENTS

We appreciate the research assistance of Yik Ting Choi.

8. REFERENCES

[1] Chmielewski, A.K., 2014. An international comparison of achievement inequality in within-and between-school tracking systems. *American J. of Education*, 1203, 293-324.

[2] Hanushek, E. A., and Wöbmann, L. 2006. Does educational tracking affect performance and inequality? *The Economic J.*, 116, 510, C63-C76.

[3] Opdenakker, M.C. and Damme, J., 2001. Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research J.*, 27, 4, pp.407-432.

[4] Chiu, M. M. and Walker, A., 2007. Leadership for social justice in Hong Kong schools Addressing mechanisms of inequality. *J. of Educational Administration*, *45,* 6, 724-739.

[5] Boaler, J., 2013. Ability and mathematics. *Forum*, 55, 1, 143-152.

[6] Hanushek, E.A., Kain, J.F., Markman, J.M. and Rivkin, S.G., 2003. Does peer ability affect student achievement? *J. of applied econometrics*, *18*5, 527-544.

[7] Goldstein, H., 2011. *Multilevel statistical models*. John Wiley & Sons, Sydney.

[8] Kennedy, P., 2003. *A guide to econometrics*. Blackwell, Cambridge, MA.

[9] Lleras, C. and Rangel, C., 2009. Ability grouping practices in elementary school and African American/Hispanic achievement. *American J. of Education*, *115,* 2, 279-304.

[10] Watanabe, M. 2008. Tracking in the era of high stakes state accountability reform. *Teach. Coll. Rec.*, *110,* 3, 489-534.

[11] Smith, B. J. 2013. Dropouts and differentiation. In *Differentiated instruction*, E. F. Sparapani, Ed. University Press of America, Lanham, MD, 221-230.

[12] Westwood, P. 2013. *Inclusive and adaptive teaching*. Routledge, New York, NY.

[13] Brechwald, W. A. and Prinstein, M. J. 2011. Beyond homophily. *J. of Research on Adolescence*, 21, 1, 166-179.

[14] Chiu, M. M. 2008a. Inequality mechanisms that hurt both privileged and disadvantaged students' learning. In I. H. Wadell's Ed. *Income Distribution*. Nova Science Publishers, Hauppauge, NY*,* 79-98.

[15] Jansen, M., Schroeders, U., Lüdtke, O. and Marsh, H. W. 2015. Contrast and assimilation effects of dimensional comparisons in five subjects: An extension of the I/E model. *J. of Educational Psychology,* 107, 4, 1086.

[16] Pintrich, P. R. 2003. A motivational science perspective on the role of student motivation in learning and teaching contexts. *J. of Educational Psychology,* 95, 667-686.

[17] Kaplan, S. N., Guzman, I. and Tomlinson, C. A. 2009. *Using the parallel curriculum model in urban settings, grades K-8*. Corwin, Thousand Oaks, CA.

[18] Chiu, M. M. and Khoo, L. 2005. Effects of resources, inequality, and privilege bias on achievement. *American Educational Research J.,* 42, 575-603.

[19] Blatchford, P., Pellegrini, A. D. and Baines, E. 2016. *The Child at School*. Routledge, New York, NY.

[20] Watson, W. E., Johnson, L. and Zgourides, G. D. 2002. The influence of ethnic diversity on leadership, group process, and performance. *Int. J. of Intercult. Rel.*, 26*,* 1, 1-16.

[21] Van Knippenberg, D. and Schippers, M. C. 2007. Work group diversity. *Annual Review of Psychology*, 58, 515-541.

[22] Sharan, Y. 2010. Cooperative learning for academic and social gains. *European J. of Education*, 45, 2, 300-313.

[23] De Dreu, C. K. W. & West, M. A. 2001. Minority dissent and team innovation. *J. of Appl. Psychology,* 86, 1191-1201.

[24] Chiu, M. M. 2008. Flowing toward correct contributions during groups' mathematics problem solving. *J. of the Learning Sciences,* 17, 3, 415-463.

[25] Paulus, P. B. and Brown, V. 2003. Ideational creativity in groups. In *Group creativity*, P. B. Paulus and B. A. Nijstad Eds. Oxford University Press, New York, NY, 110-136.

[26] Schippers, M. C., Den Hartog, D. N. and Koopman, P. L. 2007. Reflexivity in teams A measure and correlates. *Applied psychology*, 56, 2, 189-211.

[27] Van Offenbeek, M. 2001. Processes and outcomes of team learning. *European J. of Work and Organizational Psychology*, 10, 303-317.

[28] Sweller, J. 2010. Cognitive load theory. In *Cognitive load theory*, J. L. Plass, R. Moreno and R. Brunken, Eds. Cambridge University Press, New York, NY, 29–47

[29] Baddeley, A. 2003. Working memory. *Nature reviews neuroscience*, 4, 10, 829-839.

[30] Tribushinina, E. 2008. *Cognitive reference points*. Netherlands Graduate School of Linguistics, Utrecht, Netherlands.

[31] Glogger, I., Holzapfel, L., Kappich, J., Schwonke, R., Nuckles, M. and Renkl, A. 2013. Development and evaluation of a computer-based learning environment for teachers. *Education Research International*.

[32] Liu, W. C., Wang, C. K. J. and Parkins, E. J. 2005. A longitudinal study of students' academic self-concept in a streamed setting: The Singapore context. *British J. of Educational Psychology,* 75, 4, 567-586.

[33] Skibbe, L. E., Phillips, B. M., Day, S. L., Brophy-Herb, H. E. and Connor, C. M. 2012. Children's early literacy growth in relation to classmates' self-regulation. *J. of Educational Psychology,* 104, 541-553.

[34] Chiu, M. M. and Chow, B. W.-Y. 2010. Culture, motivation, and reading achievement: High school students in 41 countries. *Learning and Individual Differences*, *20, 57*9-592.

[35] Chiu, M. M. and Chow, B. W. Y. 2015. Classmate characteristics and student achievement in 33 countries. *J. of Educational Psychology,* 107, 1, 152-169.

[36] Chiu, M. M. and McBride-Chang, C. 2006. Gender, context, and reading. *Scientific Studies of Reading,* 10, 331–362.

[37] Baker, D. P., Goesling, B. and Letendre, G. K. 2002. Socioeconomic status, school quality, and national economic development. *Comparative Education Review,* 46, 291-312.

[38] Chiu, M. M. and Klassen, R. M. 2009. Calibration of reading self-concept and reading achievement among 15-year-olds. *Learning and Individual Differences,* 19, 372-386.

[39] Thapa, A., Cohen, J., Guffey, S. and Higgins-D'Alessandro, A. 2013. A review of school climate research. *Review of Educational Research*, 83, 3, 357-385.

[40] Winters, M. A., Haight, R. C., Swaim, T. T. and Pickering, K. A. 2013. The effect of same-gender teacher assignment on student achievement in the elementary and secondary grades. *Economics of Education Review*, 34, 69-75.

[41] Martin, M. O., Mullis, I. V. S. and Kennedy, A. M. 2003. *PIRLS 2001 technical report.* TIMSS & PIRLS International Center, Boston College, Chestnut Hill, MA.

[42] World Bank. 2007. *The world development report 2006*. Oxford University Press, New York, NY.

[43] Vygotsky, L. S. 2014. Genesis of the higher mental functions. In P. Light, S. Sheldon, M. Woodhead's (Eds.) *Learning to think* (pp. 35-45). Routledge, New York, NY.