IJIET, e-ISSN 2548-8430, p-ISSN 2548-8422, Vol. 3, No. 2, July 2019

International Journal of Indonesian Education and Teaching
http://e-journal.usd.ac.id/index.php/IJIET
Sanata Dharma University, Yogyakarta, Indonesia

THE INFLUENCE OF EARLY LITERACY CHARACTERISTICS AND SES ON THE LITERACY ACHIEVEMENT OF STUDENTS WHO SPEAK NON-DOMINANT LANGUAGES IN INDONESIA

Kevin M. Wong
New York University
kevinwong@nyu.edu
https://doi.org/10.24071/ijiet.v3i2.1965
received 24 June 2019; revised 26 June 2019; accepted 30 June 2019

Abstract
Across the world, children from non-dominant language speaking families are not performing as well as their peers who speak a dominant language when they enter school. The current study examines the case of Indonesia, investigating the influence of language background status, early literacy characteristics, and socioeconomic status on literacy achievement in Indonesia. Drawing from the PIRLS 2011 dataset (N = 2,725), findings reveal that there is a significant association between each variable and literacy achievement, and that socioeconomic status explains literacy achievement most strongly among 4thgrade students in Indonesia. Implications are discussed.

Keywords: education, international testing, literacy, multilingualism

Introduction

The influence of early literacy characteristics and SES on the literacy achievement of students who speak non-dominant languages in Indonesia

In 2013, the number of international migrants reached 232 million worldwide (UNDESA, 2013). With the increase of global migration patterns in recent decades, and the settlement of ethnic minority populations for generations in host countries, nations have become increasingly multicultural and multilingual. As such, more and more countries are faced with the challenge of creating linguistic and cultural accommodations to address the holistic educational needs of language minority students.

Indonesia is one such country as the fourth most populous country in the world with a population of over 250 million (CIA World Factbook, 2014), over 300 ethnic groups and 700 languages (Lewis, Simons & Fennis, 2015). In fact, Greenberg's Linguistic Diversity Index places Indonesia among the top 30 most diverse countries in the world with an index of 0.816 on a scale of 0 to 1 where 1 represents the probability of picking two people who speak different mother tongues (Greenberg, 1956; Lewis, Simons & Fennig, 2015). Without question, millions of people in Indonesia speak non-dominant languages (NDL) – languagesthat do not hold official status or command high prestige in comparison to dominant languages (DL) of society (Benson & Kosonen, 2013).

Examining these groups of speakers, comparative studies reveal that there is a discrepancy in the literacy performance between NDL and DL students in various countries (Ogle, Begnum & Scott, 2008; Ogle, Miller & Malley, 2006), comparing countries with a similar socioeconomic status (CPE, 2015; Ogle, Miller & Malley, 2006) and also by geographic location in the Nordic region (Ogle, Begnum & Scott, 2008). In these studies, DL students consistently outperformed NDL students in literacy achievement. With the exception of a few studies (Oey-Gardiner, 1991; Widjaja, 1989), there is not much documentation in the literature regarding literacy and NDL speaking students in Indonesia. One aim of the current study is therefore to examine whether these trends in literacy achievement apply to the Indonesian context, a context rich with linguistic diversity.

In addition, the current study recognizes the importance of family background and home resources in literacy outcomes. Studies document that a lack of home resources is associated with poor literacy performance among NDL students (Organisation for Economic Cooperation and Development; OECD, 2006). Likewise, the U.S. National Literacy Panel on Language Minority Children concludes that families play a significant role in the literacy development of their children (Goldberg, Rueda, & August, 2006). The current study thus investigates the influence of home characteristics on literacy achievement, specifically looking at early literacy characteristics and home economic resources as key factors that might affect the literacy achievement of NDL students in Indonesia.

Literature review

Across the world, children from NDL speaking families are not performing as well as their DL speaking peers in schools. Ogle, Begnum and Scott (2008) compared home and school characteristics of NDL and DL speaking fourth grade students in Denmark, Norway and Sweden to understand how these factors affected literacy achievement on the PIRLS 2006 assessment – an international assessment that measures progress in reading and literacy. Homes with NDL speakers were identified as families with neither parent speaking the language of the test. Ogle, Begnum and Scott (2008) compared the percentage of NDL and DLresponses to home and school characteristics across countries; compared the average PIRLS scaled scores and standard errors for NDL and DL groups; and ran a multiple regression analysis to explore the correlation between NDL and DL literacy achievement with home and school characteristics. Findings revealed that home characteristics explained more variation in literacy achievement among NDL and DL groups than school characteristics, concluding that home characteristics variables were better able to predict literacy achievement than language background.

In addition to language background, a number of studies document the relationship between academic achievement and early literacy characteristics, such as preschool or Kindergarten attendance and home literacy activities. Caughy, DiPietro and Strobine (1994) explored the benefits of preschool on short-term reading scores. Using the large-scale National Longitudinal Survey of Youth for 5- and 6-year-old children in the United States, researchers found that children

in poverty impacted contexts benefitted from preschool. Following the progress of these children longitudinally, however, the effects of preschool were unaccounted for by the time children reached second grade. Researchers also found a significant interaction between reading achievement and home environment rather than income, suggesting that children with higher scores on home environment were likely to demonstrate higher reading scores.

Importantly, early literacy characteristics focus on the learning environmentat home. Examining a series of these characteristics, Leseman and Jong (1998) rana multivariate analysis on the opportunities, instructional quality, cooperation and socio-emotional quality of children to predict early reading achievement. Researchers conducted interviews with 89 families who had 4-year-old children in inner city Netherlands. These interviews were coupled with observations on parent-child book reading interactions at the ages of 4, 5, and 6 years old. Findings revealed that home literacy was multifaceted and significantly impacted by background factors such as SES, ethnicity, and parent literacy practices. Interestingly, after controlling for vocabulary learning and home language, there was still a significant effect for opportunity, instruction quality, and cooperation quality on home literacy.

Besides literature on the language minority status and early literacy characteristics of children, there is also a body of research on the effects of SES on academic achievement. Although this relationship varies across cultures and contexts (OECD, 2006), both parental education and income are consistently associated with academic achievement. Sammons, Elliot, Sylva, Melhuish, Siraj-Blatchford and Taggart (2004) conducted a large-scale study looking at pre-school education, SES, and cognitive attainment. The group of researchers investigated a randomized control group of children who attended 141 pre-schools in the UK with a home sample of children who did not attend pre-school. Running a multilevel analysis on the home environment, children attainment, and pre-school attendance, Sammons et. al. (2004) found that SES, income, mother's education level, ethnic background, and language background all played a role in academic achievement. They concluded that pre-school is beneficial to young children, particularly if their environments are impacted by SES. In line with the Sammons et. al. (2004) findings, the current study will hone in on parental education and thenumber of books at home to represent income as variables that affect fourth grade literacy performance in Indonesia.

Although there is less research in the Indonesian context with regard to the influence of language background status, early literacy characteristics, and SES on literacy achievement, a few studies do illuminate the inequities of schooling by gender (Oey-Gardiner, 1991) where school availability appears to predict the enrollment ratios for males and females in Indonesian schools. Examining the relationship between this ratio and school availability, formal sector employment, drop-out patterns, and marriage patterns, Oey-Gardiner (1991) found that school availability was a stronger predictor for females than for males. Also in the Indonesian context, scholars have utilized PISA (Program for International Student Assessment) problems to examine the mathematical literacy of Grade VIII students, finding that using a Pendidikan Matematika Realistik Indonesia

approach allowed students to construct their mathematical knowledge to help with large-scale international assessments like PISA (Larasati & Rianasari, 2017) and PIRLS (Progress in International Reading Literacy Study).

While much of the literature has been conducted in international contexts, few studies examine how home characteristics might affect literacy outcomes in the Indonesian context. As researchers identify different factors that influence reading achievement, the present study seeks to explore the effects of language background status (NDL or DL status), early literacy characteristics (including kindergarten attendance and home literacy practices), and SES (focusing on parental education and books at home) on literacy achievement in Indonesia. Thus, the research question guiding this study is: How does language background status (DL), early literacy characteristics, and socioeconomic status (SES) affect literacy achievement of 4th grade students in the PIRLS 2011 assessment inIndonesia? Moreover, can early literacy characteristics and SES help explain the relationship between DL status and literacy achievement?

Method

Dataset

The dataset was used as a secondary analysis from the Progress in International Reading Literacy Study (PIRLS) assessment – an international assessment that is administered every year from the International Association for the Evaluation of Educational Achievement (IEA). The dataset of interest was taken from Indonesian fourth grade students who took the test in 2011. This was the last time Indonesia participated in the PIRLS assessment as they opted not to participate in 2016. In addition to reading assessment results, this study will draw on data from the student questionnaire that students completed after the assessment, answering questions about their home and school life, as well as a home questionnaire titled the "Learning to Read Survey" that was completed by parents, asking questions about language use at home, early literacy characteristics and home resources.

In the home questionnaire, both the early literacy tasks (Q6) and early reading activities (Q2) are indices composed of five and nine questions, respectively. The data from these indices were compiled according to Rasch modeling with an international mean of 10 and an international standard deviation of 2. Through Rasch modeling, each scaled score was divided into three categories: the most desirable (high), the least desirable (low), and the remaining score in the middle. For early literacy tasks, a scaled score of <8.9 was considered not well, 8.9-11.5 was moderately well, and >11.5 was very well. Similarly, for the frequency of early reading activities, a scaled score of <6.2 described never to almost never, 6.2-10.7 was sometimes, and >10.7 was often (see Table 2). From the dataset, literacy achievement (reading score) of 4th graders on the PIRLS 2011 assessmentserved as the dependent variable. A series of independent variables are included inTable 2. Moreover, to capture language use at home, the survey asks (Q21) "When talking at home with your child, what language does the child's father (or stepfather or male guardian) use? What language does the child's mother (or stepmother or female guardian) use?

Among early literacy characteristics, kindergarten attendance, early literacy tasks and early reading activities were chosen because of the literature that substantiates their influence on literacy achievement. In addition, SES factors like the number of children books at home and parents' levels of education were selected because of the established association between SES and literacy achievement in recent literature (Sammons, et. al, 2004).

Importantly, the language of the assessment and questionnaires was translated to Bahasa Indonesian. While Bahasa Indonesian is the official language of schools and the education sector with 22.8 million L1 speakers, Javanese is the language with most L1 speakers (84.3 million). There are also many other dialects that are not accounted for in the translation, including major dialects like Banjar (3.5 million), Bugis (5 million) and Musi (3.1 million; Lewis, Simons & Fennig, 2015). This diversity of language varieties may account for a large portion of the original sample who did not answer this question in the home survey as it was not offered in their language (see Table 2 Notes). Still, the IEA translates all home surveys and student assessments through two of the following methods: multiple- forward translation, back-translation, translation review by bilingual judges, and statistical review (Maxwell, 1996).

Analytical strategy

The quantitative analysis of this study proceeds with a multiple regression analysis because the study involves models that have two or more predictor variables and a single, continuous dependent variable. A regression analysis will also demonstrate how variables are related to one another, the strength of these relations, and the relative predictive power of the independent variables on the dependent variable. The models use OLS regression between literacy achievement and: (1) DL status; (2) DL status and early literacy characteristics; (3) DL status and SES; and (4) all variables. In addition, a series of independent t-tests and one-way ANOVAs are used to compare language majority and minority students on literacy achievement, testing whether each variable in the sample is significantly correlated with literacy achievement and can be extrapolated to the general population.

Findings and Discussion

Descriptive Patterns

The analytic sample in Table 2 consists of 2,725 4th grade students in Indonesia who sat the PIRLS assessment in 2011. After listwise deletion of missing, omitted and invalid variables, the 2,725 in this sample represents 59% of the original sample. The largest number of deletions were taken from the DL variables where 918 surveys had omitted the language(s) spoken by the father, and1,164 surveys had omitted the language(s) spoken by the mothers. After isolating this sample population, the mean score of reading achievement was 436.13 (SD=76.50), and language minority (or NDL) students, defined as those with both parents who did not speak the majority language, accounted for 49% of the analytic sample. With almost equal representation between the DL and NDL groups, we are also able to explore the impact of DL on literacy achievement.

The remaining independent variables were divided into two categories: early literacy characteristics and SES. These variables were used to explore how these factors might explain the relationship between DL and literacy achievement. Most students experienced favorable early literacy characteristics (see Table 2). About 72% of 4th grade students in the sample attended kindergarten. While this is an unexpectedly large percentage of students attending kindergarten, which causes us to question who the sample consists of in this dataset, the original sample of 4,644 students before listwise deletion reveals a similar 71.27% of students in Indonesia who attended kindergarten.

Besides kindergarten attendance, children were reported by parents to have a 'moderate level' of literacy prior to entering school with an average of 10.05 (SD=1.84). This was calculated from Rasch modeling between the transformed scaled cutoff values of 8.9 and 11.5. Most students appeared to be 'sometimes engaged' in early reading activities (mean = 9.40, SD = 1.73). This was calculated from Rasch modeling between the transformed scaled cutoff values of 6.2 and 10.7. Finally, the socioeconomic factors varied among Indonesian families in the analytic sample. The number of children books at home was relatively low (mean = .49 out of 3, SD=.67). Meanwhile, both fathers and mothers had similarly moderate levels of education (mean = 1.74, SD=.81 and mean = 1.71, SD=.79, respectively).

Dominant Language Status, Early Literacy, and SES

The first variable that we are trying to explain is the influence of DL status on literacy achievement. A two-tailed independent t-test was run on DL (Table 3) and yielded a statistically significant relationship (t=-8.3, DF=2,273, p<.001), suggesting that DL speakers are statistically more likely to have a higher literacy score than NDL speakers.

Among early literacy characteristics (kindergarten attendance, early literacy tasks and early reading activities) in Indonesia, each of variable as statistically significant through a two-way independent t-test for kindergarten attendance (t=-10.96, DF=2,723, p<.001). A simple correlation revealed a medium, positive relationship for early literacy tasks (r=.33, p<.001) and a weak, and positive relationship for early reading activities to literacy achievement (r=.17, p<.001). For the latter two variables, this means that as early literacy tasks and early reading activities increases, so does the reading score, and vice versa.

Lastly, SES factors like the number of children books at home and parents' levels of education were used as categorical variables in a one-way ANOVA test. Findings revealed a significant positive relationship between literacy achievement and books at home (F=28.67, DF=3, p<.001), father's level of education (F=123.89, DF=3, p<.001), and mother's level of education (F=145.70, DF=3, p<.001), and vice versa. Since all the factors for early literacy characteristics and SES had significant relationships with literacy achievement, they were included inthis study to see if they could help explain the influence of DL status on literacy achievement.

OLS Regression

From the OLS regression analysis in Table 4, there is a discrepancy between DL and literacy achievement in the 2011 4th grade PIRLS assessment (Model 1). Findings reveal that the reading score is expected to be 24.04 points higher for students with NDL status ($p<.001$). Since this is a binary variable, non-NDL students score an average of 424.46 while NDL students score an average of 448.50. As such, we accept the alternate hypothesis and reject the null hypothesis because there is a significant association between DL and reading achievement in Indonesia.

To better understand this discrepancy, three additional models were run to see how student early characteristics and socioeconomic status affected literacy achievement and DL status. The OLS regression analysis reveals that both student early characteristics (Model 2) and SES (Model 3) affect literacy achievement. As expected, kindergarten attendance and early literacy tasks were strongly associated with reading achievement, predicting an increase of 14.56 and 11.14 points in literacy achievement, respectively, with each unit of increase ($p<.001$). Surprisingly, early reading activities were not correlated with literacy achievement in this model ($p>.05$), suggesting it does not explain reading achievement as well as the other variables. This leads us to question how accurately the index of sub-questions was able to capture early reading activities. Interestingly, DL status in Model 2 remains very significant ($p<.001$) suggesting that DL, kindergarten attendance and early literacy tasks play a similar role in the influence of literacy achievement. This also suggests that early literacy characteristics do not necessarily explain the relationship between DL and readingachievement. Yet it is important to note that this model accounts for 12% of the variation in literacy achievement ($R2=.12$). Findings from this model suggest that families who send their children to kindergarten and focus on the five literacy skills measured in the PIRLS home survey (reading the alphabet, words and sentences, writing letters and words) have an increased likelihood of improving their children's level of literacy.

The number of books at home in Model 3 predicts an increase of 8.71 points ($p<.001$) per increase of books. In conjunction with this economic factor, both father's and mother's levels of education predict an increase of 12.43 and 25.50 points in literacy achievement, respectively, for every unit of increase ($p<.001$). It is important to note the very high coefficient for mother's level of educationwhich is more than double that of father's education level and triple that of the amount of children books at home. Interestingly, DL status is not significant ($p>.05$) when the regression model includes SES ($p>.05$), suggesting SES explains the influence of literacy achievement much more than DL status. In this model, SES is able to explain 15% of the variation in literacy achievement ($R2=.15$). This model reveals that there are noteworthy issues of inequality and impact of SES on literacy attainment. This suggests that children from low SES families are not given access to the same opportunities or resources as those from high SES families, which calls attention to the need for policymakers to address and protect the education (and linguistic) trajectories of these children.

These findings are all corroborated in the last model where the OLS regression includes DL status, student early characteristics and SES. Early literacytasks and the amount of children books at home remain significantly correlated with literacy achievement, though the parents' levels of education have the strongest influence on literacy achievement – and in particular, the mother's level of education which predicts an increase of 20.67 points in literacy achievementper increase in level of education ($p<.001$). In addition, DL is not significant in this model suggesting that the language of parents at home may not explain literacy attainment as much as the other factors. Furthermore, this final model accounts for almost one fifth ($R2=.19$) of the variation in literacy attainment, which is quite substantial. Looking ahead, it may be valuable for future studies to tease out other factors that help explain the effect of DL, such as parent expectations, the sociolinguistic effects of linguistic capital in society, and school efforts to assimilate language speakers to the dominant language. Unfortunately, because this was a secondary analysis, variables were limited in scope according to the questions asked by the survey.

Overall, the four models show a stepwise decrease in the AIC, suggesting all factors in the last model are important and contribute to literacy achievement. Therefore, this study suggests that there is an association between DL status, early literacy characteristics, and SES on literacy achievement. Furthermore, SES accounts for the largest explanation for literacy achievement among 4th grade students in Indonesia.

Conclusion

Findings from this study are aligned with the literature that emphasize the influence of parent education and family income on academic achievement (Haveman & Wolfe, 1995; Klebanov, Brooks-Gunn, & Duncan, 1994; Smith, Brooks-Gunn, & Klebanov, 1997). While the literature shows that the language(s) spoken by parents are an important factor in academic achievement, the current study reveals that other factors like parent education have a stronger influence on literacy attainment in Indonesia. Moving forward, it would be interesting to unpack why it is that parent education influences academic attainment in Indonesia and whether there is a correlation between the language(s) spoken by parents and their levels of education. Davis-Kean (2005) found that parents'beliefs and behaviors had a significant influence on child achievement because of the expectations that educated parents placed on their children. Davis-Kean (2005)also found that these expectations differed by racial group, which could suggest potential discrepancies bewteen dominant and non-dominant linguistic groups in Indonesia.

Secondly, it would be useful to triangulate data from this study with other quantitative datasets (international or national data) that measure literacy, particularly with data more recent than the 2011 PIRLS assessment. Moreover, the current study excluded 41% of the original population due to missing data. This poses a limitation to the generalizability of results from the analytic sample to the broader population, especially because those who omitted this answer are likely to speak a language other than the language of the survey. It would also be valuable

to include qualitative interviews of parents and students to better understand how language is used at home and better capture variation within DL status. In some countries, speaking a non-dominant language does not necessarily equate to less linguistic capital (Bourdieu, 1991), nor does it inform us of the status of speakers, whether they be immigrants, Indigenous groups, 3rd generation residents or simply members of a plurilingual country. Qualitative interviews may help address the limitations of this study.

Finally, future studies may look into analyzing literacy achievement in countries that are similarly multilingual and diverse like Papua New Guinea, or contrast them with monolingual countries like South Korea and Hong Kong. Findings from such comparative studies may reveal the different needs of countries, and inform policymakers from these countries about how resources might be allocated to best support the linguistic and educational needs of non- dominant speaking students.

References

Benson, C. & Kosonen, K. (2013). *Language issues in comparative education: Inclusive teaching and learning in non-dominant languages and cultures.* Rotterdam: Sense Publishers.

Bourdieu, P. (1977). The economics of linguistic exchanges. *Social Science Information, 16*(6), 645-668.

Caughy, M., DiPietro, J. & Strobine, D. M. (1994). Day care participation as a protective factor in the cognitive development of low income children. *Child Development, 65*(2), 457-471.

Center for Public Education [CPE]. PIRLS of wisdom: Full report. Retrieved on March 15, 2019 from http://www.centerforpubliceducation.org/Main-Menu/Evaluating-performance/PIRLS-of-wisdom-At-a-glance/PIRLS-of-wisdom-Full-report.html

CIA World Factbook. (2014). Indonesia: People and society. Retrieved from https://www.cia.gov/library/publications/the-world-factbook/geos/id.html

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology, 19*(2), 294.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth, 8*(2), 195-222.

Goldberg, C., Rueda, R. S., & August, D. (2006). Sociocultural influences on the literacy attainment of language-minority children and youth. In August, D. & Shanahan, T. (Eds.) Developing Literacy in Second-Language Learners: Report of the National Literacy Panel on Language-Minority Children and Youth (pp. 269-318). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Greenberg, J. H. (1956). The measurement of linguistic diversity. *Language*,*32*(1), 109-115.

Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature, 33*(4), 1829–1878.

Klebanov, P. K., Brooks-Gunn, J., & Duncan, G. J. (1994). Does neighborhood and family poverty affect mothers' parenting, mental health, and social support? *Journal of Marriage and the Family, 56*(2), 441–455.

Larasati, S. R. & Rianasari, V. F. (2017). Mathematical literacy profile of grade VIII students of SMP Pangudi Luthur 1 Yogyakarta using Pendidikan Matematika Realistik Indonesia approach. *International Journal of Indonesian Education and Teaching, 1*(1), 63-72.

Leseman, P. P., & Jong, P. F. (1998). Home literacy: Opportunity, instruction, cooperation and social-emotional quality predicting early reading achievement. *Reading Research Quarterly, 33*(3), 294-318.

Lewis, M. P., Simons, G. F. & Fennig, C. D. (2015). Ethnologue: Languages of the World, Eighteenth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com

Maxwell, B. (1996). "Translation and cultural adaptation of the survey instruments" in M.O. Martin and D.L. Kelly (eds.), Third International Mathematics and Science Study (TIMSS) Technical Report, Volume I: Design and Development. Chestnut Hill, MA: Boston College.

Oey-Gardiner, M. (1991). Gender differences in schooling in Indonesia. *Bulletin of Indonesian Economic Studies, 27*(1), 57–79.

Ogle, T. O., Begnum, A. C. & Scott, E. (2008). Comparisons of Home and School Characteristics of Language Minority and Language Majority Fourth-graders: PIRLS Results from Three Nordic Countries. Paper presented at Comparative and International Education Society Conference, New York.

Ogle, L. T., Miller, D. C. & Malley, L. B. (2006). Characteristics of United States Grade 4 language minority students in an international context: Findings from PIRLS 2001. Presented at the International Education Assessment Conference Proceedings.

Organisation for Economic Cooperation and Development (OECD). (2006). *Where immigrants students succeed: A comparative review of performance and engagement in PISA 2003*. Paris: OECD.

Sammons, P., Elliot, K., Sylva, K., Melhuish, E., Siraj-Blatchford, I., & Taggart, B. (2004). The impact of pre-school on young children's cognitive attainments at entry to reception. *British Educational Research Journal, 30*(5), 691-712.

Smith, J. R., Brooks-Gunn, J., & Klebanov, P. K. (1997). Consequences of living in poverty for young children's cognitive and verbal ability and early school achievement. In G. J. Duncan & J. Brooks-Gunn (Eds.), Consequences of Growing Up Poor (pp.132–189). New York: Russell Sage Foundation.

United Nations Department of Economic and Social Affairs (UNDESA). The number of international migrants worldwide reaches 232 million. Retrieved on April 6, 2015 from: http://esa.un.org/unmigration/documents/The_number_of_international _migrants.pdf

Widjaja, H. (1989). Praktek-praktek asuhan anak di kecamatan Cikerah [Child rearing practicesin the area of Cikeruh]. Unpublished manuscript, Padjadjaran University, Bandung, Indonesia.

Appendix

Table 1. Primary ethnicities in Indonesia

Ethnic groups	Population (million)	Percentage	Main Regions
Javanese	95.217	42.00	Central Java, East Java, Yogyakarta, Jakarta, North Sumatra, South Sumatra, Bengkulu, Lampung
Sundanese	30.978	15.41	West Java, Banten, Jakarta, Lampung
Malay	6.946	3.45	North Sumatra, Riau, Riau Islands, Jambi, South Sumatra, Bangka–Belitung Islands, West Kalimantan
Madurese	6.772	3.37	Madura Island, East Java
Batak	6.076	3.02	North Sumatra, Riau, Jakarta
Minangkabau	5.475	2.72	West Sumatra, Riau
Betawi	5.042	2.51	Jakarta, Banten, West Java
Bugis	5.010	2.49	South Sulawesi, South East Sulawesi, Central Sulawesi, East Kalimantan
Acehnese	4.419	2.05	Aceh
Bantenese	4.113	2.05	Banten
Banjarese	3.496	1.74	South Kalimantan, East Kalimantan
Balinese	3.028	1.51	Bali
Tionghoa	2.832	1.20	North Sumatra, Riau, Riau Islands, West Kalimantan, Jakarta, Bangka–Belitung Islands
Sasak	2.611	1.17	West Nusa Tenggara
Makassarese	1.982	0.99	South Sulawesi
Minahasan	1.900	0.96	North Sulawesi
Cirebonese	1.890	0.94	West Java, Central Java

Table 2: Definitions and Descriptive Statistics of Variables (N=2,725)

Variable	Definition and metrics	Mean	SD	Min	Max
Reading Achievement	First plausible value of achievement in reading	436.13	76.50	108.74	672.94
Language Majority Student (NDL)	Scale: 0 = no parents speak majority language to 1 = one or both parents speak majority language	.49	-	-	-
Early Literacy Characteristics					
Kindergarten Attendance	Scale: 0 = did not attend kindergarten to 1 = attended kindergarten	.72	-	-	-
Early Literacy Tasks	Mean of 5 items describing early literacy ability. Scale: <8.9 = not well 8.9-11.5 = moderately well >11.5 = very well Reliability (alpha) = .91	10.05	1.84	4.84	13.18
Early Reading Activities	Mean of 9 items describing frequency of early reading activities. Scale: <6.2 = never to almost never 6.2-10.7 = sometimes >10.7 = often Reliability (alpha) = .75	9.40	1.73	2.14	14.71
Socioeconomic Status					
Amount of Children Books at Home	Scale: 0 = very few books to 3 = many books	.49	.67	0	3
Father's Level of Education	Scale: 0 = no school to 3 = tertiary education or higher	1.68	.78	0	3
Mother's Level of Education	Scale: 0 = no school to 3 = tertiary education or higher	1.63	.77	0	3

Source: [PIRLS 2011 International Dataset, 4th grade Indonesian student sample]
Note: A total of 41% of the sample was omitted due to missing data, which has implications for the generalizability of the findings. Home characteristics are adapted from Ogle, Begnum & Scott (2008) study.

Table 3: Independent t-test results comparing language majority and minority students on literacy achievement

Group	n	Mean	SD	t	Df	p
Language Minority Students	1,402	424.46	74.04	-8.30	2,723	<0.001
Language Majority Students	1,323	448.49	77.14			

Source: [PIRLS 2011 International Dataset, 4th grade Indonesian student sample]