UNIVERSITY OF GOTHENBURG

SCHOOL MATERIAL RESOURCES AND STUDENT READING ACHIEVEMENT IN THE UNITED ARAB EMIRATES

PISA 2018 data through the lens of Ecological Systems Theory

Linda C. Gogliotti, M.A.

Master's thesis:	30 credits
Programme/course:	L2EUR (IMER) PDA184
Level:	Second cycle
Term/year:	Spring 2020
Supervisor:	Aimee Lee Haley, PhD
Examiner:	
Report nr:	

Abstract

Aim: This study aimed to 1) identify differences in overall student reading achievement in the UAE based on student gender, 2) determine if there is school material resource inequity in the UAE based on school location or school type, and 3) measure the effects of school material resources on student reading achievement in the UAE while controlling for other variables.

Theory: The theoretical framework used for this study is Bronfenbrenner's Ecological Systems Theory including the Process-Person-Context-Time model.

Method: Statistical analyses using UAE's PISA 2018 data included t-test, principal component analysis, PLUM ordinal regression, and two-level hierarchical linear modeling.

Results: This study joins a small but growing amount of research focused on using data such as PISA's to better understand the UAE educational system and perhaps to help further its reforms. Findings include: Girls outperformed boys in reading achievement in the UAE. As for material resource inequity based on school type, co-ed schools were more likely than either of the single-sex school types to report that their school's capacity to enhance learning and teaching using digital devices was sufficient. However, co-ed schools were also more likely to report that their school's capacity to provide instruction was hindered due to the quantity/quality of the material resources when compared to either of the single-sex schools. While almost all of the relationships in the regressions were statistically significant (school type, emirate, and school urbanization level), both models did a poor job of fitting the data. Finally, of the two school material resources indices used in this study, the school materials index was not related to reading achievement, but the school digital devices index was a significant predictor of student reading achievement. Every unit increase in the index corresponded to a 9.069 increase in the predicted reading achievement score. Although the included variables reduced the variance, some unaccounted variance remains.

Foreword

Life takes us many places. This study was born of a blending of more abstract and theoretical concerns combined with more practical, classroom-based concerns. As a result of the time I spent in the UAE, investigating this topic felt like a logical next step since I like engaging in practical, useful, and hopefully meaningful tasks; I think this study is all three. In combining my research interests along with my practitioner (classroom-based) concerns, this study is a blending of my past experiences, present reality, and future possibilities.

To those to whom I am forever grateful –

Much love, health, and happiness to everyone –

Linda Gogliotti

Gothenburg, Sweden

June 2020

Table of Contents

List of Tables

List of Abbreviations and Scientific Symbols

B	beta value
CI	confidence interval
Df	degrees of freedom
EPF	education production function
EST	Ecological Systems Theory
HH	His Highness
HLM	hierarchical linear modeling (also known as multilevel modeling)
IEA	International Association for the Evaluation of Educational Achievement
ILSA	international large-scale assessment
MoE	Ministry of Education
OECD	Organisation for Economic Co-operation and Development
p	significance level (e.g., $p < .001$)
PCA	principal component analysis
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PPCT	Process-Person-Context-Time
PV	plausible value
RM	Rasch Model
SD	standard deviation
SE	standard error
SES	socioeconomic status
SEM	structural equation modeling
SMR	school material resource(s)
SPSS	Statistical Package for the Social Sciences (= IBM SPSS Statistics)
SRA	student reading achievement
t	t-statistic
TIMSS	Trends in International Mathematics and Science Study
UAE	United Arab Emirates

Introduction

As part of the United Arab Emirates' (UAE) Vision 2021, HH Sheikh Mohammed bin Rashid Al Maktoum, Vice President and Prime Minister of the UAE and ruler of Dubai, set the National Agenda aim of creating a first-rate education system (UAE, 2018). As one measure of that, the UAE aspires to rank as one of the top countries on both the Trends in International Mathematics and Science Study (TIMSS) and Programme for International Student Assessment (PISA) assessments ("The UAE Reforms Education System as Part of Vision 2021," 2017) despite ranking below the Organisation for Economic Co-operation and Development (OECD) average in all of the main subject areas tested (mathematics, reading, and science) (Westley, 2017).

Though both are international large-scale assessments (ILSAs), TIMSS and PISA are administered by different organizations, each with its own aims. Briefly, TIMSS is administered by the International Association for the Evaluation of Educational Achievement (IEA), "an international cooperative of national research institutions, government research agencies, scholars, and analysts working to evaluate, understand, and improve education worldwide" (IEA). "By linking research, policy, and practice to assess and measure how well education systems are preparing children for the future," (IEA), the IEA "aims to help its members understand effective practices in education and develop evidence-based policies to improve education" (IEA). On the other hand, PISA, the data used for this study, is administered by the OECD, an international organisation that works to shape policies that foster prosperity, equality, opportunity, and well-being for all and for better lives (OECD, 2019d). Together with governments, policy makers, and citizens, they work on establishing evidence-based international standards and finding solutions to a range of social, economic and environmental challenges, including improving economic performance and creating jobs (OECD, 2019d). With a focus on economic preparedness and competitiveness, PISA assesses "knowledge and skills experts in the participating countries and economies consider to be most important for students' full participation in knowledge-based societies" (OECD, 2018) with a view to better preparing the world of tomorrow (OECD, 2019d).

This study is a result of the two years I spent working for the UAE's Ministry of Education (MoE). In September 2016, I went to the UAE as part of the first major wave of 'Western' teachers hired to help the MoE reach their Vision 2021 education aim. Working in a variety of capacities, I witnessed changes from inside the classroom as well as from MoE headquarters in Dubai and Abu Dhabi. From curriculum development to teacher training, I was part of the widespread changes being carried out. Though changes are ongoing, the purpose of this study is to see where the UAE ranks as measured by its PISA 2018 reading achievement scores and to investigate a belief amongst some teachers about school resource inequity as related to student gender[1].

In the UAE, each year a specific focus is decreed; e.g., HH Sheikh Khalifa bin Zayed Al Nahyan, President of the UAE, proclaimed 2019 the Year of Tolerance. Coinciding with the introduction of the new curriculum, 2016 was the Year of Reading. To that end, many MoE initiatives were geared towards increasing students' reading ability in line with the preparation of an integrated national literacy strategy and the enactment of the National Reading Law (Warner & Burton, 2017). PISA's 2018 major domain was reading, just a couple of years after the UAE's Year of Reading. Though PISA assesses reading ability each assessment, the last time reading was the major domain prior to 2018 was 2009. Related to reading, new textbooks which supported the revised curriculum started to be introduced in UAE public schools from 2016 (Warner & Burton, 2017). These textbooks would go on to be revised during the 2016-2018 period (and thereafter); however, the changes started to be implemented on a wider scale within the UAE from 2016. (n.b. Dubai/the Northern Emirates worked

---

[1] For the purposes of this paper, 'gender' equals 'sex'.

separately from Abu Dhabi some of this time.) This new/revised curriculum may have impacted reading achievement, and the 'intervention' began during the three-year period between the PISA 2015 and PISA 2018 assessments. Hence the timely nature of this study.

As the UAE continues to make changes to its educational system, analyzing the effect(s) school material resources (SMR), which may influence student achievement (Chiu & Khoo, 2005; Woessmann, 2016), have on reading achievement may aid future decision making. However, given the number and breadth of changes taking place concurrently as well as method limitations, drawing causal links will not be possible. Although changes in achievement cannot be attributed to specific causes, understanding the context within which the PISA test was administered, and educational reforms made, may provide additional insights into any progress or lack thereof as measured by the main variables of interest, which for this study are SMR and student reading achievement (SRA). Though educational reform is a long-term goal (particularly given cultural changes that must accompany such an undertaking), a shorter-term assessment of achievement is worth looking at in my opinion.

To that end, this study uses data from 2018, the latest assessment cycle of PISA, and statistical analyses in combination with Bronfenbrenner's Ecological Systems Theory (EST) to investigate how SMR affect overall SRA with a focus on students attending single-sex schools in the UAE. The targeted nature of this inquiry coupled with a lack of UAE-specific research gives rise to the research gap this study seeks to address. The aims then are to:

- identify differences in overall SRA in the UAE based on student gender.

- determine if there is SMR inequity in the UAE based on school location or school type.

- measure the effects of SMR on SRA in the UAE while controlling for other variables.

Potential confounding variables are considered as they relate to SMR and SRA. These include student-level variables, gender and socioeconomic status (SES), as well as school-level variables, location (emirate and urbanization level), type of school (based on the gender of the students attending), teacher quality, and the percentage of students from socioeconomically disadvantaged homes.

Particularly now as "a substantive number of the initiatives address the effectiveness of the educational system with a focus on a return on investment" (Tamim & Colburn, 2019, p. 162), having research-backed data (albeit processed through the non-economic theoretical lens of EST) in order to make informed decisions, e.g., about allocation policies or choice of materials, seems an imperative as the UAE carries on with its educational reforms. After all, in most contexts, money is limited, presumably even more so in developing countries engaged in widescale educational reforms. The UAE is such a country. Therefore, return on investment in the form of achievement gains seems desirable (Della Sala, Knoeppel, & Marion, 2017).

Background – UAE Context

The understanding of context is essential; not just cultural but historical as well. Although focus is on the future, an understanding of the UAE's current educational system as well as the historical changes that have led up to this point are essential. Since the UAE is a young nation with a developing educational system, the UAE is not a well-known context in educational research, and as a result, not a lot of UAE-specific research is available, particularly independent (non-governmental) research (see *Education in the United Arab Emirates*, 2019). The first organized modern school in what is now the UAE goes back to just 1930 in Sharjah, one of the seven emirates, with primary education becoming mandatory for Emiratis after 1971, the year the country was founded (Alhebsi, Pettaway, & Waller,

2015, p. 4). Unlike other more established countries, the UAE, having just celebrated its 48th National Day in 2019, and its educational system do not have lengthy histories on which to look back.

In terms of UAE-specific research, due to the widespread educational reforms taking place in the country, even UAE-context research – the limited amount there is – published relatively recently (e.g., within the past 10 years) can be rendered outdated in some respects and must be treated with caution. For example, starting from 2008 but particularly in 2010, 2016, and 2017 (Gallagher, 2019b, p. 3; Tamim & Colburn, 2019, p. 172 & 174; Taylor Gobert, 2019, p. 113), large-scale reforms went into effect vis-à-vis curriculum, staffing, and organizational structure (e.g., the new textbooks and the hiring en masse of 'Western' teachers in 2016, and the merging of Abu Dhabi and Dubai educational authorities in 2017 for countrywide unification). Thus, even a source that could be considered foundational, such as Ridge's (2009) dissertation which looked at the education of boys in the UAE, does not in some respects reflect the current situation in the UAE just a decade later, so profound have some of the changes been.

Nonetheless, one source that proved invaluable in understanding the current as well as historical changes in the UAE education sector was Gallagher's (ed.) *Education in the United Arab Emirates* (2019), which brought together articles on a wide range of topics related to education in the UAE. Initially skeptical of the book's independence given the editor's affiliation with a federal university in the UAE and the other contributing authors' connections to various other UAE institutions some of which are also federal, it is included in the corpus on account of its comprehensive coverage of a wide range of education-related topics focused on the UAE as well as its recent (2019) publication. Most importantly, academic independence was shown in the writings as the topics and findings were rather pointed. For example, oft-repeated themes were lack of information sharing (Kippels & Ridge, 2019, p. 37), and the rapid rate of change (Gallagher, 2019a, p. 140), often devoid of follow-up analyses (Dickson, Fidalgo, & Cairns, 2019, p. 107). In addition, commented on frequently was a lack of research (Dickson et al., 2019, p. 107; Gallagher, 2019c, p. v), particularly publicly available education research (Kippels & Ridge, 2019, p. 37). These challenges were addressed head on. Worth noting is that, despite their best efforts, even some of the authors were limited at times by lack of information (Tamim & Colburn, 2019, p. 169).

Putting this study into its cultural and historical country-specific context, it is necessary to briefly address each of the themes mentioned above. The changes in the UAE – not just limited to the education sector but focused here now – have been described in dramatic if not hyperbolic terms of change. Words such as 'revolution', 'transformation', 'unprecedented', and 'unparalleled' have been used to describe the UAE government's educational reforms (Gallagher, 2019b, pp. 6-7) as well as the development and growth of the country itself. While some terms may be more figurative than literal – perhaps even leaning towards exaggeration (Gallagher, 2019b, p. 7) – the use of such adjectives speaks to the number and breadth of changes the UAE government is attempting to make to its educational system (Gallagher, 2019a, p. 140) as well as the fast pace and cyclical manner of those changes (Gallagher, 2019a, p. 127; Tamim & Colburn, 2019, pp. 168-169).

As ambitious an undertaking as these wide-ranging reforms are, their implementation has not been without criticism. Many have noted the situation as "constantly evolving" (Gallagher, 2019a, p. 127) with changes that are rapid if not too so (Tamim & Colburn, 2019, p. 169). Lack of information sharing with the public (Tamim & Colburn, 2019, p. 169) along with lack of publicly available policy information (Kippels & Ridge, 2019, p. 51; Tamim & Colburn, 2019, p. 169 & 172) have had their effect. Dickson (2013) confronts these frustrations in her research, addressing them from the perspective of students. Combined with the many changes, lack of information sharing has left some students feeling frustrated and angry. As Dickson (2013) explains, they wanted their voices – expressing mainly displeasure – to be heard when asked about the educational reforms that had taken place (p. 280). It seems the opportunity for this was lacking, which is unfortunate as "greater

communication between education stakeholders would increase buy-in of future policies, including those around curriculum, and ultimately support the implementation of new restructuring and reforms" (Kippels & Ridge, 2019, p. 50). The perceptions of such students speak to the chaotic and disruptive nature of the changes. While all change can be difficult, the manner in which some of the changes have been made may have gone far in exacerbating an already tumultuous time. That was the state of the UAE educational system during this study's timeframe. As Gallagher (2019c) cautions in the book preface for which she is editor, "(I)n this rapidly evolving educational context, where new initiatives are proposed, new policies are enacted, new agencies are formed, and existing agencies are repurposed on a frequent basis, the volume presents a panoramic snapshot of the state of contemporary education" (p. vi) in the UAE.

Finally, as noted previously, this study is attempting to help address a research gap. When there is educational research in the UAE, it is small-scale and independent; large-scale educational research is rare (Gallagher, 2019c, p. v). The UAE would benefit from more publicly available education data and research (Kippels & Ridge, 2019, p. 51). For example, a major focus of the UAE government is innovation with technology in the classroom being a big part of that. Despite this, "there is a dearth of published academic work on the actual implementation of technology in UAE schools" (Dickson et al., 2019, p. 107). While this may be due to the country's young educational system and an even younger technology sector as the authors posit (Dickson et al., 2019, p. 107), it seems addressing this gap sooner rather than later could go far if only in enhancing change management. By utilizing ILSA data, this thesis is one of not so many UAE-specific studies to join the discussion.

Motivation for Study

While researching this subject, I came across a quote that in many ways speaks to the very heart of this study: "(T)eachers care about difficult-to-measure variables such as the availability of materials, and the quality of administrative support" (Johnson, 1990, as cited in Murnane, 1995, p. 318). Like most teaching practitioners, I too care about availability of materials, and that is where this study originates – my interest as a practitioner based on what I experienced in the field. During my time in the UAE, schools (girls' as well as boys') lacked materials (e.g., textbooks and hard/software) at times due at least in part to the constant changes brought on by the educational reforms. Further, while working briefly as a secondary school teacher in two all-girls' government-run schools in two different emirates, I came to find out that some of things I experienced were common to many teachers. (Ridge (2009) outlines some of these in her dissertation *Privileged and penalized: The education of boys in the United Arab Emirates*; see UAE Context.) For example, based on my tenure in the UAE, there is a belief amongst some teachers that girls' schools are not allocated the same resources as boys' schools (e.g., photocopiers and paper). Many of the schools in the UAE, particularly government-run ones, are segregated by sex (Kippels & Ridge, 2019, p. 40). Therefore, this study analyzes the connection between student gender, SMR, and SRA. Despite general societal gender equality in the UAE (particularly in comparison to other countries in the MENA[2] region) and the UAE government's strong support for women in the workplace, there may be an underlying assumption that investment in girls' education is not as necessary; it may not pay off as much as for boys. Although Ridge offers information to discount this assumption, such a disparity – whether real or imagined – may have consequences, including possibly affecting achievement.

Research Questions

The research questions (RQs) this study seeks to answer are:

RQ1:   How do girls compare to boys in overall student reading achievement in the UAE?

---

[2] MENA = Middle East and North Africa

RQ2:   In the UAE, how do school location and school type relate to school material resources?

RQ3:   In the UAE, how do school material resources relate to student reading achievement with other variables being controlled?

Structure

Following on from this introduction, the remaining chapters include: the literature review organized around student/reading achievement and SMR; a discussion of EST as the theoretical framework; the method section which includes information about the data, variables, analytical approach, study limitations, validity and reliability matters, and ethical considerations; presentation of results for RQs 1-3; a discussion section focused on the findings as they relate to EST; and finally the last section, Conclusions and Recommendations, which focuses on future research possibilities and considerations when using ILSA data and findings from such studies for comparative purposes.

Literature Review

The literature review is organized around the main variables of interest – SRA and SMR. Within each of these sections, other variables that may impact these variables (and which are controlled for in this study) are included.

Student/Reading Achievement

When researching student achievement for this study, topics focused on were student gender, student SES (at the student and school levels), teacher quality, and school location (emirate and urbanization level) as these are used in this study. When choosing sources for inclusion in the corpus, emphasis was placed on student reading achievement when possible.

Gender and Achievement

There is general agreement in the literature that girls outperform boys in reading. Chung (2018, p. 53) found an average reading advantage of 30 points for girls in all countries analyzed when doing a cross-national analysis of PISA 2015 data. As an OECD partner country, the UAE was not included. Similarly, using PISA 2000 data, Chiu and Khoo (2005, p. 587) found girls scored +24 points in reading compared to boys (M = 470.85, 48.80 min. – 854.69 max.) (Chiu & Khoo, 2005, p. 586). Using PISA 2009, Reilly (2012) found that girls outperform boys in reading literacy across all nations (p. 6), noting statistically significant differences more pronounced at both tails of distribution with each favoring girls (p. 10). Tsai, Smith, and Hauser (2018) also found gender gaps favoring girls in all six of the countries they studied (three East Asian and three Western) using PISA 2012.

The same is true for the UAE. Kippels and Ridge (2019) state there is a pronounced reverse gender gap in education in the UAE (p. 49), noting a 50-point difference when referring to PISA 2015 reading results (p. 50). Although many teachers in the GCC[3] say their students' weakest skill is reading (Taylor Gobert, 2019, p. 117), students in the UAE achieved the highest score of all Arab countries on PIRLS[4] 2016 (Gallagher, 2019b, p. 2). Despite this, the UAE still ranks below the OECD averages on all PISA measures, including reading (Westley, 2017). Regarding general gender achievement differences in the UAE, Russell (2012) notes that girls in the UAE had better educational outcomes (p. 93), and boys were more likely to drop out and not pass exams (p. 84).

This gender gap is neither for all countries nor all subjects, however. Using four PISA assessment results (2000 – 2009), Stoet and Geary (2015) found that girls outperformed boys in overall

---

[3] GCC = Gulf Cooperation Council; comprised of six countries: Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the UAE

[4] PIRLS = Progress in International Reading Literacy Study

achievement (reading, mathematics, and science) in 70% of the countries, with girls falling behind in just 4% of the countries (p. 137). While the reading gap has increased over the past decade (Stoet & Geary, 2015, p. 149), they note that the PISA data for overall achievement follows a pattern. While girls outperform boys, the largest gender gap favoring girls exists at the lowest achievement levels. At the highest achievement levels, there is a reverse gender gap in developed countries which favors boys (Stoet & Geary, 2015, p. 148). Finally, Marks (2008), using PISA 2000 data, found that gender gaps in reading and math were "highly correlated" (p. 106), and "the magnitude of the gaps reflects the implementation and success or otherwise of policies designed to improve girls' educational outcomes" (p. 106). In other words, reducing the mathematics gender gap will likely increase the reading gender gap (Marks, 2008, p. 106).

As shown above, there is an overwhelming consensus that girls outperform boys in reading achievement. While the magnitude of the gender gap may vary along the achievement continuum, perhaps dependent on a country's status of development (as will be discussed later) and/or measures of economic equality as some research indicates (e.g., Chiu, 2018), most of the corpus is in agreement that girls outperform boys in reading achievement.

Student Background and Achievement
This section focuses on prior research related to SES, including parents' education and books in the home, as they relate to student achievement.

Socioeconomic Status / Parents' Education
Analyzing the effects school resources have on achievement in Finnish senior secondary schools, Häkkinen, Kirjavainen, and Uusitalo (2003) found that parents' education (family background) along with the grade point average in comprehensive school (earlier achievement) were the strongest explanatory variables for student achievement (p. 329).

Not only can one's SES influence one's own achievement, but so too can it affect one's classmates' achievement. In a study which used Bronfenbrenner's EST and PIRLS data, Chiu and Chow (2015) found that classmates' SES and home resources had more of an effect on a student's reading achievement than other classmate characteristics (e.g., attitudes toward reading). In addition, high-SES classmates benefited high-SES students more than low-SES students (Chiu & Chow, 2015, p. 163). Following on from EST, the theoretical perspective chosen for this paper, children are influenced by their immediate environments. At school, this would include classmates/peers whose individual SES could impact that of other children, and vice versa.

Books in the Home
Attempting to explain the variance of reading achievement of Hong Kong pupils in the PIRLS 2011 study, Cheung et al. (2017) used structural equation modeling (SEM) to propose and test a model which was able to explain 34% of the reading achievement variance. "Parental background acts as the fundamental factor that exerts an indirect effect on reading motivation, reading self-efficacy, and reading achievement of students via books at home and early reading abilities" (Cheung et al., 2017, p. 889). It is worth noting that Hong Kong often ranks as one of the top performers on both PISA and PIRLS.

Teacher Quality and Achievement
In his synthesis of more than 800 meta-analyses related to achievement, Hattie (2009) found that quality of teaching is one of the most important determinants of learning. Two studies whose findings support the importance of teachers in student achievement are Ma and Crocker (2007) and Ning, Van Damme, Gielen, Vanlaar, and Van Den Noortgate (2016), both of which are detailed in Achievement and School Material Resources. Finally, in summarizing what was known about school effects' influence on achievement up to that time, Gustafsson (2003) wrote: "The results indicate that among the resource factors, teacher competence is the single most powerful factor in influencing student

achievement, and the effect sizes seem to be substantially larger than those associated with class size" (p. 103). Further, "Given the strength of effects associated with teacher competence, it would seem that investments in teacher competence would have a higher likelihood of paying off in terms of student achievement than would other investments" (Gustafsson, 2003, p. 104).

Location and Achievement

Considering student achievement and location as they relate to this study, Chung (2018) found an 'urban advantage', whereby students from economically developed countries scored higher in reading on average. This advantage was moderated by gender (p. 72). However, what may be surprising is that female students from more economically developed societies tended to score approximately five points lower than males for each dollar increase in GDP per capita (Chung, 2018, p. 72) and six points lower than the average gender gap (Chung, 2018, p. 69). Stoet and Geary (2015) also described the PISA pattern as differing based on a country's status of development with no gender gap (or a closed one) in most developing countries, but a reverse gender gap favoring boys at the highest achievement levels in developed countries (p. 148).

Regarding the UAE, in her mixed-methods dissertation, Russell (2012) built upon Ridge's (2009) work by using teacher interviews and student questionnaires at four schools in Ras Al Khaimah (RAK), one of the seven emirates that comprise the UAE, as well as country-wide MoE data to examine gender, academic achievement, and meanings of schooling in RAK. She found that school location by emirate may impact outcomes, but no difference was detected based on urban vs. rural (Russell, 2012, p. 93). This may be because in the UAE K-12 education is managed at both the federal and emirate level (Kippels & Ridge, 2019, p. 39), and "like much else in education in the UAE it varies depending on the sector and the emirate" (Gallagher, 2019a, p. 140). To test these findings, both school emirate and urbanization level are variables in this study. Since most single-sex schools at the secondary school level in the UAE are public (government-run) schools, school sector has been considered indirectly.

Summary

Despite the individual nature of student achievement, research has established some common determinants, or predictors, including student gender, SES, parents' education, books in the home, and teacher quality. Less well established in the literature is the effect school location might have; however, as that is one of this study's foci, it is included here as well. Now, however, we turn to a discussion of SMR.

School Material Resources

After a brief historical summary of SMR research and a description of SMR in the UAE educational context, education production functions (EPFs) as they relate to SMR research are looked at. Finally, SMR are discussed in relation to achievement, gender, and location.

Historical Summary

Overall, prior research on SMR has resulted in mixed results and conflicting findings often leading to contradictory conclusions being reached. Several sources provide a thorough recounting of this history. Understanding the historical development of the school resource literature, including but not limited to prior research's contradictory findings, is desirable for getting a more complete understanding of the field[5]. Some of the conflicting results can be attributed to early, well-cited research (e.g., Hanushek's) that used less-advanced EPFs, which look at school and student inputs and a measure of student output, to analyze data (e.g., see Gustafsson, 2003). Other reasons include the

---

[5] For an overall summary of the history of 'resource' research, see Chudgar and Luschei (2009), Gustafsson (2003), or Odden, Borman, and Fermanich (2004). For a clearly presented and concise summary of the school resources do/not matter debate, see Della Sala et al. (2017).

definition/operationalization of this variable varying widely in the literature because other words (e.g., school effects and school resources) are often used interchangeably for this area of research. Following on from them, a variety of different measures are used, including expenditure per student, measures related to teachers, SES variables, country economic (e.g., GDP) and equality measures, and percentage of female students. As a result, it is quite difficult if not incorrect to try and compare findings and conclusions. Chudgar and Luschei (2009) offer possible explanations as to why there are so many discrepancies in the literature. While noting the difficulties in measuring relevant school variables, they say better measures of school resources are needed (Chudgar & Luschei, 2009, p. 648).

UAE Context

Many readers may not be familiar with the UAE, the country itself let alone its educational system. Therefore, describing the UAE educational context in terms of SMR in an effort to more fully understand and appreciate this study and its findings seems essential.

According to Ridge (2009), 100% of the female teachers in her study reported having to pay for their photocopying compared to just 50% of males (p. 124). Several explanations were offered as to why this might be; e.g., female teachers' greater discretionary income as they are not usually head of household and the more positive learning environments created in girls' schools with decorations and the like. Perhaps the belief amongst some teachers that girls' schools are not allocated the same resources as boys' schools is a result of this difference. Although Ridge offers information to discount this belief, such a disparity – whether real or imagined – can have negative consequences, including affecting achievement.

As a further example of what motivated this study into SMR in the UAE, public schools in the UAE blamed late delivery of course books (amongst other things) for their failure in final exam results in the first trimester of 2017 (Taylor Gobert, 2019, p. 122), and the three main barriers cited to the successful implementation of CALL[6] in teaching reading to children were the lack of availability of resources, lack of hardware, and lack of suitable programs (Dickson et al., 2019, p. 98).

Concerning the quantity of resources at secondary schools in the UAE, Ridge (2009) found that they are much the same (abstract). It is unclear if this is at a policy level only or if it transcends to implementation/allocation as well, or if that is at the federal level only and then it is dependent on emirate, because, in what seems to contradict that point, she notes that boys feel more negatively about their school resources than girls even though, as in the example given, they had more computers than girls (Ridge, 2009, pp. 122-123). The UAE government's 'gender-neutral' stance, which she states has actually benefited girls (Ridge, 2009, abstract) notwithstanding, again, there may be an underlying assumption that investment in girls' education is not as necessary; it may not pay off as much as for boys. This could materialize through the allocation of resources to schools be it based on location or the gender of the schools' students. Hence this study. While Ridge (2009) notes that there are, in fact, differences between boys' and girls' schools at the secondary-school level, she says they, like the government's gender-neutral stance, also benefit girls (abstract).

While there is a lot of literature related to SMR, there is very little related to the UAE specifically. Therefore, this study is an opportunity to make a valuable contribution to the debate surrounding SMR in an effort to help fill the UAE research gap, especially now amid the country's wide-ranging educational reforms.

---

[6] CALL = computer-assisted language learning

Education Production Functions and School Material Resources Research

The current study does not use an EPF[7]; however, along with economic theory, much of the early school resource research did. Therefore, it is important to discuss EPFs for historical and comparative purposes. Although they are still used today, the focus here is on understanding the limitations of EPFs in an effort to qualify prior results that used them. Of particular interest are the researchers who explained why they did not use EPFs in their studies and the types of analyses they chose instead; or, if they used an EPF, how their EPF was different to ones used in previous research (how they compare or are improvements upon earlier EPFs). In recounting and summarizing the history of EPF school resource research complete with its conflicting findings and competing types of analyses, several corpus sources give much coverage to EPFs and explain them well. They include Chudgar and Luschei (2009)[8], Della Sala et al. (2017), Gustafsson (2003), and Odden et al. (2004), all of which will be discussed further.

In order of publication date as statistical methods have advanced, we start with Gustafsson. In summarizing the effects school resources had had on educational results up until that time, Gustafsson (2003) wrote: "(T)here is reason to believe that every single study that has been conducted has omitted variables that should have been included in order to obtain unbiased estimates of the effects of resource variables" (p. 83). Not only might these variables have been non-randomly omitted (e.g., entry achievement level and resource history), but inappropriate variables or multiple measures of the same resource variable may have been used (Gustafsson, 2003, p. 83). A further critique of EPFs is that they do not investigate the intervening educational process by simply looking at input/output (Gustafsson, 2003, p. 84). Since the EPF tends to disregard the multilevel nature of educational data (Gustafsson, 2003, p. 84), "we need a more solid foundation of research than is furnished by the educational production function studies" (Gustafsson, 2003, p. 85). In effect, EPFs simplify the process too much, thereby limiting the meaningfulness of the results, if not perhaps rendering them useless (Monk, 1992, as cited in Gustafsson, 2003, p. 84).

Odden et al. (2004) also discuss the history of EPF research. They did not use an EPF in their research as an EPF was not helpful because district-level variables, which are averages, cannot be considered conclusive (Odden et al., 2004, p. 20). They did not use such data because, as they note, using readily available data – not necessarily the most important data – can cause the misspecification of an EPF (Odden et al., 2004, p. 19). If, e.g., factors that were not the most important in determining student performance are used, the effects of school and teachers will be underestimated. And if the school and teacher variables are district averages, much of their variation across students will be eliminated, thereby reducing the ability to find an association (Odden et al., 2004, p. 19). Resources as defined by them included expenditures per pupil, and school and class sizes.

Finally, Della Sala et al. (2017) thoroughly cover the debate surrounding educational resources and student achievement. Their conclusion is that input/output frameworks like EPFs limit the potential to account for educational resources' unique effects on achievement because they do not accommodate mediating and moderating variables (Della Sala et al., 2017, p. 198). "Production equations are limited to the degree that they model only the quantitative contributions of resources while leaving aside more qualitative aspects of how resources are deployed in the classroom" (Della Sala et al., 2017, p. 188). EPFs limit results because they can only account for a single dependent variable. Using school- and district-level variables to account for the variation, they are unable to fully depict the effects of resources on student achievement because those variables do not measure variations at the

---

[7] See Theoretical Framework for information about EST and the rationale for why it was chosen to guide the analysis and interpretation of this study's data. EPFs are discussed here in an historical and comparative capacity only.

[8] See Location and School Material Resources for more information about Chudgar and Luschei.

classroom level. In addition, there are difficulties in attaining precise measurements of variables (Della Sala et al., 2017, p. 187).

As illustrated above, research into school effects/school (material) resources has a fairly long history. Often, the work of Hanushek is cited. Hanushek (1996) himself acknowledged that per-pupil expenditure studies (a common EPF variable) do not analyze resources at the classroom level but rather they aggregate the data at the school district level (p. 406). And it was Fuller (1987), when reviewing 60 multivariate studies, who suggested that perhaps it is time to abandon the EPF metaphor because researchers know little of why things do/not work (e.g., how material resources are managed) (p. 288). See Location and School Material Resources for more on his reasoning. As methods of analysis have evolved and school resource effects become more nuanced, this topic may still hold many unanswered questions beyond just an unresearched context like the UAE.

Achievement and School Material Resources

Focusing now on SMR and their effects on SRA, Archibald (2006), using 2002-2003 InSite data for third to sixth graders in the US, looked at school-level resources, not aggregated district-level spending as is often used in such research (p. 34). School resources was measured by per-pupil expenditure, which was broken down into four categories, one of which was instruction (= pupil-use technology, software, instructional materials, supplies, etc.) (Archibald, 2006, p. 40). Expenditures for instruction and instructional support (school-level per-pupil spending) were positively related and statistically significant for reading (Archibald, 2006, p. 34). The coefficient of variation for per-pupil spending for instruction = 0.15 (Archibald, 2006, p. 41). Similarly, Marks, Cresswell, and Ainley (2006) found that material resources – defined as wealth and educational resources (= dictionary, desk, textbooks, calculators, etc.) – have a substantial impact on student achievement in a small minority of countries (p. 105); e.g., Brazil, Mexico, Portugal, and the US (Marks et al., 2006, p. 122). Using PISA 2000 (reading was the major domain), when controlling for material resources, the average effect of socioeconomic background on reading achievement declined by six points (17%) across countries (Marks et al., 2006, p. 115 & 122).

Another study using PISA 2000 observed that better equipment with instructional material (measured as strongly or not at all lacking) was associated with superior student performance (Fuchs & Wößmann, 2007, p. 461). Although Fuchs and Wößmann (2007) determined "the importance of institutions for the cross-country variation in test scores seems to be greater than that of resources" (p. 460), they did note that, particularly for quality of instructional material, the effects are "positively related to student performance once family-background and institutional effects are extensively controlled for" (p. 451). They used educational expenditure per student. Similarly, Chiu (2018) observed that the greater availability of resources in richer countries can substitute for educational resources at home, thereby reducing disparity and increasing girls' reading advantage (p. 49) because wealth can buy resources which can directly or indirectly raise learning (p. 58). In an earlier study Chiu was involved in, Chiu and Khoo (2005), using PISA 2000 data, found that students scored higher when they had more resources in their country, family, or school (p. 594). More resources (students scored higher when there were sufficient teaching materials (Chiu & Khoo, 2005, p. 594)), less distribution inequality (richer countries showed more equal distribution of resources (Chiu & Khoo, 2005, p. 595)), and less privileged student bias (PSB) were all linked to higher student performance (Chiu & Khoo, 2005, p. 594). Their study used PISA's EDUSHORT index (index of education/teaching material shortage) amongst others (distribution variables and PSB), and each school resource had a small but cumulative effect size of 10% (Chiu & Khoo, 2005, p. 594). Finally, focusing on ICT[9] usage, Skryabin, Zhang, Liu, and Zhang (2015) looked at ICT usage and student

---

[9] Although ICT was not defined in this article, according to *Cambridge Dictionary*, ICT, which stands for information and communication technology, commonly includes computers and other electronic equipment used to store and send information ("ICT," 2020). ICT usage, then, would be the usage of this equipment.

achievement (reading, math, and science) but found that the relationship is still inconsistent (p. 52). For example, using TIMSS 2011 and PIRLS 2011 for fourth graders and PISA 2012 for eighth graders, ICT usage was positive and significant for the fourth graders in all three subjects (reading, math, and science) even after controlling for SES and gender but negative for the eighth graders in all three subjects (Skryabin et al., 2015, p. 54).

Though not related specifically to reading, Greenwald, Hedges, and Laine (1996), in response to Hanushek (1996), found that student achievement is related to resource availability (p. 411). They recommend that policies should ensure sufficient resources and incentives to spend those resources 'wisely' should be in place as their findings demonstrate that money and the resources the money buys do matter to the quality of a child's education (Greenwald et al., 1996, p. 415). How "wisely" is defined is unclear, however. Greenwald et al. (1996) and Hanushek (1996) are older sources, but Ning et al. (2016) came to a similar conclusion using EPF and multilevel linear modeling with PISA 2009 data (see two paragraphs below). They determined that how resources are used is more important than the amount spent in high-income countries (Ning et al., 2016, p. 527). Looking at it from another perspective but drawing similar conclusions, Hanushek and Woessmann (2017) compiled a survey of economists' work related to school resources. Based on their analyses of the prior research, shortage of material tends to be negatively associated with student outcomes, e.g., when using principal-reported shortage of material indicators; one notable exception was that the availability of computers at school was not related to student outcomes in multivariate analyses (Hanushek & Woessmann, 2017, p. 161). According to them, resources in general are a cause and a consequence of student achievement or of unobserved factors related to student achievement (Hanushek & Woessmann, 2017, p. 162). Many economists use EPFs in their research so this may have impacted their findings, and the importance of the unobserved factors affecting student achievement cannot be understated.

In comparison to those studies, several found weak or limited results. For example, Hanushek and Luque (2003) determined that "the overall strength of resources in obtaining better student performance appears rather limited" (p. 497); however, certain countries stood out because they had significant effects (Hanushek & Luque, 2003, p. 497). Important to note is that they were looking at math and science scores and school resources was defined differently than for this study. Another example is Woessmann (2016) who, using results from different years of PISA, TIMSS, and PIRLS as examples, looked at various variables such as expenditure per student, location (town/city/large city), and shortage of instructional materials (large/none). Based on his analyses of EPF research, expenditure per student and, surprisingly, class size appear to have little effect on student achievement (Woessmann, 2016, p. 27). In his estimation, the contribution of school resources is quite limited, but the predictive power of the model used varies across countries (Woessmann, 2016, p. 27). Again, different cultural, political, etc. factors can influence a model.

Jürges and Schneider (2004) came to similar conclusions when looking at TIMSS data. Although "lack of financial resources is often thought to impede high-quality teaching" (Jürges & Schneider, 2004, p. 373), they argued that shortage of instruction material and shortage of computer hardware do not appear to have a sizeable effect on the distribution of math scores even if teachers report their teaching is limited by such (p. 369). When there is no shortage at all of instruction material, students' scores were about four points higher than the average (Jürges & Schneider, 2004, p. 369). While Jürges and Schneider (2004) used neighboring countries presumably to try and make the comparisons more similar (p. 358), in a comparison of two distant and rather dissimilar places (Shanghai and Finland), Ning et al. (2016) found that 7% of the differences in reading achievement were attributed to school-level variables (p. 522). In Finland, the quality of educational materials significantly predicted reading achievement, but it did not in Shanghai (Ning et al., 2016, pp. 522-523). Even with the 0.52-point advantage quality of educational materials gave Finland (Ning et al., 2016, p. 526), qualified teachers had a bigger impact for both locations (Ning et al., 2016, p. 526). In terms of transformation power, quality of educational materials was a small but significant advantage for

Shanghai, but again, teachers were higher (Ning et al., 2016, p. 527). The takeaway was that how resources are used (e.g., by skilled teachers) is more important than the amount spent on them in high-income countries (Ning et al., 2016, p. 527).

Compared to the positive or negative impacts mentioned above, some research found no effect of school resources on student achievement or mixed results. For example, based on their model, Della Sala et al. (2017), using US Dept. of Education elementary school data, discovered a non-significant relationship between schools' instructional conditions and student achievement measures (p. 199). Their instructional condition variable includes a percentage of expenditures for instruction. Another is Van Hek, Kraaykamp, and Pelzer (2018) who found no effect of school materials on gender differences in reading performance (p. 12). For further details, see Gender and School Material Resources. Finally, while Jürges and Schneider (2004) did not find that a shortage of instructional materials affected math scores, Ma and Crocker (2007) suggest something even more positive about lack of resources – it may have a positive effect on achievement[10]. In looking at school material and instructional resources in Canada's 10 provinces using PISA 2000 data, Ma and Crocker (2007) found that overall SMR had positive statistically significant effects on reading achievement in seven of them; three provinces had statistically significant negative effects. For school instructional resources, four provinces had a statistically significant positive effect on reading achievement, whereas six had statistically significant negative (p. 101). Similar to Canadian provinces, this study looks at differences between emirates in the UAE.

Gender and School Material Resources

The corpus is not in agreement as to whether there are gendered effects from school resources. For example, in a quasi-experimental study using the German longitudinal ELEMENT dataset of reading and math ability for fourth to sixth graders and the German-I-Plus 2003 data, Legewie and Diprete (2012) found that boys are more sensitive to school resources that create a learning-oriented environment (p. 463). Their study focused on peer socioeconomic composition as the school resource variable. Yet, the authors argue that their theoretical argument can apply to all kinds of school resources that create a learning-oriented environment despite their findings being limited to the variable tested (Legewie & Diprete, 2012, p. 481). As one might expect, they suggest future studies that use other school-based resources (Legewie & Diprete, 2012, p. 481). This study seeks to do just that. As Legewie and Diprete researched the Germany context, this study investigates the UAE context.

Also using schools' socioeconomic composition as a component of their school resources variable, Van Hek et al. (2018) found no effect of school materials on gender differences in reading performance (p. 12). Using PISA 2009, they used school materials as an indicator of school resources and as a control variable (Van Hek et al., 2018, p. 17); they "considered the lack, shortage, or inadequacy of instructional materials (e.g., textbooks), computers, internet, library staff, and library materials" (Van Hek et al., 2018, p. 9). Their school resources variable was defined as schools' socioeconomic composition, proportion of girls, and proportion of highly educated teachers (Van Hek et al., 2018, p. 8). Their conclusion was that "it depends on the country context whether and how schools' socioeconomic composition affects girls' and boys' reading scores" (Van Hek et al., 2018, p. 15) . (Again, due to widely varying variable operationalization, making comparisons can be challenging. Some of these variables are, however, discussed under Student/Reading Achievement; e.g., SES.) Finally, using economic equality as a variable, Chiu (2018) found that in countries with greater economic equality, which may or may not apply to the UAE, parity of resources affects girls more than boys and increases the reading gap (p. 60). As Greenwald et al. (1996) also stated, wealth

---

[10] For an interesting perspective on how lack of school resources may be an advantage, see Ma and Crocker (2007) under Location and School Material Resources.

can buy resources (p. 415). Chiu (2018) also observed that wealth can buy resources which can directly or indirectly raise learning (p. 58).

Finally, some research shows that boys are more impacted by school resources than girls (Legewie & Diprete, 2012). Thus, although the initial focus of this study is on a perceived inequity in resources for girls' schools, similar to some other countries, rising gender inequalities in educational performance in the UAE which have boys not carrying on to higher education as much as girls make this study relevant as there are consequences for society as a whole (e.g., in the labor market and family life) (Van Hek et al., 2018, p. 3). Particularly in a country like the UAE where there is often the cultural expectation that a man will provide for his family/families, the trajectory of young men in education, and following on from there, the workforce, is of interest.

As concluded by Van Hek et al. (2018, p. 15), it may well depend on the country-context as to whether and how boys' and girls' reading scores are affected. Therefore, it is worth understanding how SMR influence student achievement in the UAE based on gender. Given the lack of UAE context-specific research as well as the conflicting findings in prior research, this study aims to determine the effect(s) in the UAE context.

Location and School Material Resources
For the purposes of this study, two aspects of school location are of interest – emirate and urbanization level. Though there are rural locales in all of the emirates, some of the emirates (namely Abu Dhabi, the capital, and Dubai) are wealthier than the others and thus may have more resources to expend regardless of rural or urban setting. Also looked at here is the comparison between developing and developed countries. Although the school resources variables in other corpus research were sometimes defined in ways different to how the variable has been defined for this study (SMR = textbooks, computer hard/software, digital devices, etc.), many of the studies included in this section used more similarly defined variables to this study than are included elsewhere. Further, as the focus is on a country's level of economic development (developing vs. developed), one can see if countries are affected in dis/similar ways in terms of various resource variables. Despite the weak effect SMR may have on student achievement, the effect may be stronger in developing countries as discussed below. In this respect, the UAE is somewhat of a paradox. Although the UAE is a developing country, unlike many other developing countries, it is a wealthy one. Because of this, findings in the literature that have applied to developing often resource-poor countries may not apply to the UAE. While UAE-context literature was prioritized, due to its scarcity, the search for literature had to be expanded. Even so, still difficult to find were studies about SMR and (reading) achievement from/in developing countries, particularly non-EPF research from suitable sources (e.g., peer-reviewed journals).

Simmons and Alexander (1978) looked at 10 previous studies that used similar analyses and some of the same variables as the others; e.g., per pupil expenditure and class size. They observed that "determinants of student achievement appear to be basically the same in both developing and developed countries" (Simmons & Alexander, 1978, p. 355). Of note, two studies found a positive and statistically significant effect for textbook availability in primary grades (Simmons & Alexander, 1978, p. 351). In response to Simmons and Alexander's findings, Heyneman (1980) contradicted them saying the determinants of school achievement are not basically the same for developing/developed countries (p. 406). One difference, according to Heyneman (1980), is "the variation in impact of economic status and school influences" (p. 403). Suggesting that "perhaps the most consistent correlate with achievement is the availability of textbooks and other reading materials" (p. 406), Heyneman (1980) further agreed with Simmons and Alexander that individual school variables which predict achievement are not consistent enough to support single-minded investment policy decisions (p. 406). Heyneman (1984) again addressed the role of educational research in developing countries when he focused on the differences between developing and developed countries and the availability of and investment in material resources. His focus was return on investment, which he found was

higher in developing countries (Heyneman, 1984, p. 298). Additionally, the percentage of explained variation due to school varies amongst countries (Heyneman, 1984, p. 300).

Supporting earlier findings related to positive and significant effects for textbooks, Heyneman, Jamison, and Montenegro (1984) looked at the impact of the Textbook Program in the Philippines. Using Grade 1 and 2 achievement tests and student questionnaires, they found the intervention was more effective on impoverished children (Heyneman et al., 1984, p. 146). Textbooks had the most pronounced effects on science scores of low-status students (Heyneman et al., 1984, p. 146), and from this the assertion that physical resources can affect significantly the quality of educational output in developing countries was made (Heyneman et al., 1984, p. 147). In contrast to Heyneman et al.'s findings but similar to Simmons and Alexander's, Hanushek and Luque (2003), using TIMSS 1995 data, determined that it "does not appear to be the case that outcomes related to school resource differences are more positive in poorer countries or in countries that begin with lower levels of resources" (p. 497). Therefore, the view that school resources are relatively more important in poor countries was not supported (Hanushek & Luque, 2003, p. 481). That said, certain countries stood out as having significant effects (Hanushek & Luque, 2003, p. 497). Like with many things then, it may depend on the context.

Two comprehensive reviews of studies are worth mentioning separately for a number of reasons. The first is Fuller's review of 60 multivariate studies conducted in the "Third World"[11] (Fuller, 1987, p. 255). The positive impact of instructional materials especially those related to reading and writing was consistent across several studies (Fuller, 1987, p. 276). Out of 24 analyses, Fuller (1987) confirmed a positive effect for instructional materials (= texts and reading materials, desks) in 16 (p. 258). Compared to math and science, school effects on reading were not as consistent (Fuller, 1987, p. 287). Despite a good deal of evidence suggesting schools exercise more influence in the 'Third World' than in industrialized countries, Fuller (1987) said this claim should be considered tentative (p. 287). Beyond the review's findings, this article was important in my opinion for other reasons. First, as previously mentioned, Fuller (1987) suggested that perhaps it is time to abandon the EPF metaphor because researchers know little of why things do/not work (e.g., how material resources are managed) (p. 288). Further, he recommended that researchers specify the conditions under which their findings hold; e.g., school effects in the 'Third World' seem stronger in rural areas and among lower-income pupils (Fuller, 1987, p. 288). Presumably, this was done to make the findings more meaningful and accurately interpreted. Fuller (1987) goes on to state that researchers should ask more interesting questions instead of doing national surveys to allow broad inferences (p. 288). One must wonder then what he would think about ILSAs like PISA. One reason he said the claim of school exercising more influence in the 'Third World' should be tentative is because specifying students' background characteristics is difficult. Demographic questions in assessments like PISA may make that easier.

In a second review worth mentioning, Fuller and Clarke (1994) attempted to explain achievement in developing countries. Despite Fuller's earlier review, it was only in this article that the term 'developing countries' was defined. As noted by the authors, developing countries are not industrialized; they are low-income countries that might be called part of the Third World, as well as somewhat more affluent nations (Fuller & Clarke, 1994, p. 144). An important takeaway, in my opinion, was the notion that cultural conditions must be taken into account. In terms of explaining achievement in developing countries, concerning textbooks, of 26 primary-school analyses, 19 had significant effects. The percentage was lower for the secondary-school analyses (7 of 13 had significant effects) (Fuller & Clarke, 1994, p. 126). According to the authors, the effect sizes for basic inputs like textbooks tend to be larger when the base line level is lower (Fuller & Clarke, 1994, p. 135) so perhaps secondary schools are better supplied than primary schools.

---

[11] "Third World" is not a term I would use now. However, that is the title of the article and the term the author used. The paper was written more than 30 years ago so we can see how language has changed.

Included here for another critical reason, Chudgar and Luschei (2009) make a distinction between relative vs. overall importance of school resources in poor countries (p. 630). This may help explain some of the discrepancies in the literature. Their conclusion is that maybe school does not exceed the importance of family in any country but the relative role of schools in poor countries may exceed the relative role in rich ones (Chudgar & Luschei, 2009, p. 630).

Finally, unlike the preceding studies which looked at the differences between developing and developed countries, Ma and Crocker (2007) compared Canadian provinces. In their study which analyzed PISA 2000 data, SMR and school instructional resources were examined in terms of their effects on reading achievement, the major domain of that year's assessment. They also included urban vs. rural location as one of the school context variables. Both of the resource variables were statistically significant provincial variables (Ma & Crocker, 2007, p. 99), and the effects were statistically significant as well (Ma & Crocker, 2007, p. 87). SMR had positive effects on reading achievement in seven of the 10 provinces; schools with more material resources performed better than schools with fewer. The remaining three provinces had negative effects. For school instructional resources, just four of the provinces had positive effects, meaning schools with more school instructional resources performed better than schools with fewer. Six of the provinces had negative effects (Ma & Crocker, 2007, p. 101). What makes this study particularly interesting is the possible explanation for the negative findings. The authors speculate that lack of these resources may have become an advantage for Quebec (Ma & Crocker, 2007, pp. 103-104). The authors suggest that it is teachers not materials that matter in regard to achievement (Ma & Crocker, 2007, p. 104). While they are not calling for a reduction in materials, if schools are not able to accommodate both teachers and materials, schools should give priority to teachers rather than resources because the findings suggest that it is teachers – not materials – that matter more in terms of achievement (Ma & Crocker, 2007, p. 104). This study includes a teacher quality variable.

Summary

Based on a systematic review of the literature, it is clear there is still a lot that is 'unsettled' in the SMR discussion. Ironically, much of what is agreed upon relates to the level of contradiction that lies within this debate; not surprising for a topic that encompasses competing interests and understandings, and which uses widely varying variable definitions and analytic processes on different subject areas studied by students of different ages in vastly different places to name but a few of the differences. While there is general agreement in the corpus that girls outperform boys in reading, it is less clear what role if any SMR (as defined for this study) can contribute to that achievement. While some may discount the effect of SMR saying it is weak or that they have less of an effect than other variables, findings from more recent research do not discount their effect(s) entirely. Further complicating matters is what effect SMR might have in a developing country like the UAE that is, in fact, not poor as is often the case.

There seems to be agreement amongst many – though not all – that EPFs do not fully capture the complexity of school resources or adequately measure school resource effects on students' achievement in the classroom. EPF research often uses variables that, while easy to acquire, are not classroom-/student-level. Aggregated data gives a general impression that does not speak to the specific classroom experience. PISA can get closer with school-level data as well as teacher data, but still there are limitations. Not many of the corpus sources utilized direct student input; notable exceptions being Heyneman et al. (1984), Ridge (2009), and Russell (2012). As was mentioned, it is often difficult to measure some of the variables, so it is often simply more pragmatic to utilize data that is easier to get and measure. Though school resource research started out using EPFs, and still does to some extent, there seems to be agreement that more and different types of analyses are needed to get more meaningful if not accurate results. How to go about that and what variable(s) should be included are, however, where the agreement starts to fade.

Theoretical Framework

Although much of the early SMR research used EPFs, due to previously mentioned limitations[12], an alternative theoretical lens which more fully accounts for the unique and interconnected 'worlds' of children was chosen. That theoretical framework is Ecological Systems Theory (EST), which was developed by American psychologist Urie Bronfenbrenner. However, this study makes use of a latter version of his theory which, building upon the original theory, includes the Process-Person-Context-Time (PPCT) model that he and his colleagues developed. Sometimes referred to as the bioecological theory or bioecological model of development (Rosa & Tudge, 2013, p. 251), this version of his theory with the PPCT model contains the most mature concepts of EST (Eriksson, Ghazinour, & Hammarström, 2018, p. 421) and was the last before Bronfenbrenner's passing.

Ecological Systems Theory

EST is used to study the interactions of children and their environments (ecosystems) to see how the interactions influence children's growth and development. Onwuegbuzie, Collins, and Frels (2013) identified EST as "a theoretical framework that incorporates virtually all research that represents the social, behavioral, and health sciences" (p. 4). Additionally, since EST is about relationships, philosophy, policy, and practice may be linked and viewed through it (Onwuegbuzie et al., 2013, p. 4). According to the theory, the student's development will be influenced by ecological system factors (e.g., a student's interaction with his/her family, classmates, school, etc.) (Dooley, 2018). These factors are part of a five-level nested model comprised of the following concentric circle ecosystems:

Level 1 – Microsystem (Immediate Setting)
Bronfenbrenner (1979) defines a microsystem as "a pattern of activities, roles, and interpersonal relations experienced by the developing person in a given setting with particular physical and material characteristics" (p. 22). These are the individual's immediate environments which provide him/her with the "initial and most direct opportunities to learn about and interact with the world" (Dooley, 2018, p. 498), including but not limited to home, neighborhood, school, and friends' homes.

Level 2 – Mesosystem (Interconnections Among Systems)
Described by Bronfenbrenner (1979) as "the interrelations among two or more settings in which the developing person actively participates (such as, for a child, the relations among home, school, and neighborhood peer group: for an adult, among family, work, and social life)" (p. 25), Level 2, or mesosystem, involves the interconnections amongst and the interactions between the microsystems (Dooley, 2018).

Level 3 – Exosystem (Community)
As defined by Bronfenbrenner (1979), an exosystem is "one or more settings that do not involve the developing person as an active participant, but in which events occur that affect, or are affected by, what happens in the setting containing the developing person" (p. 25). Even though there is no direct engagement, the individual can still be impacted negatively or positively and in a powerful way (Dooley, 2018, p. 498). Examples can include politics, a parent's/partner's job responsibilities, etc.

Level 4 – Macrosystem (Cultural Values)
Bronfenbrenner's macrosystem refers to "consistencies in the form and content of lower-order systems (micro-, meso-, and exo-) that exist, or could exist, at the level of the subculture or the culture as a whole, along with any belief systems or ideology underlying such inconsistencies" (Bronfenbrenner, 1979, p. 26). These evolve over time (Onwuegbuzie et al., 2013, p. 5) and influence

---

[12] See Education Production Functions and School Material Resources Research in the literature review for more information.

"how, when, and where individuals engage in their everyday lives" (Dooley, 2018, p. 498). Included in a macrosystem are "societal belief systems, cultural norms, ideologies, policies, or laws that indirectly influence the person" (Onwuegbuzie et al., 2013, p. 5) along with SES, ethnicity, and religious affiliation (Dooley, 2018, p. 498).

Level 5 – Chronosystem (Historical Influences and Time Dimension)

The chronosystem refers to the cumulative experiences one has in his/her lifetime, including life transitions and shifts (Dooley, 2018, p. 500). However, these changes over time include not only those "within the person but also in the environments in which that person is found" (Eriksson et al., 2018, p. 420).

Essential to EST is that "(a)ll of the systems are interconnected within and between settings; thus, they are inseparable from one another and greatly impact the individual at the center of the system" (Dooley, 2018, p. 498). Further, both the individual and the individual's environments are always changing and evolving (Dooley, 2018, p. 498). The impacts of each of these systems on the individual child will vary as each person's individual characteristics will interact with and impact relationships he/she has in combination with the social environment that surrounds him/her (Dooley, 2018, p. 498).

Process-Person-Context-Time (PPCT) Model

As a further refinement of EST, the PPCT model builds upon Bronfenbrenner's reconceptualization of the microsystem (Eriksson et al., 2018, p. 420), which now includes proximal processes. As Bronfenbrenner's theory evolved, later transformations and elaborations of his theory placed more emphasis on processes and the role of the biological person. According to Bronfenbrenner, proximal processes involve "reciprocal interaction between the developing individual and other (significant) persons, objects and symbols in his/her immediate environment" (Eriksson et al., 2018, p. 420). Bronfenbrenner considered them the most powerful predictor of human development – the "engines of development" (Rosa & Tudge, 2013, p. 251) – and "wanted to show how individual characteristics, together with aspects of the environment, influence proximal processes" (Eriksson et al., 2018, p. 420). He posited that "proximal processes operate within microsystems and involve interaction with three features of the immediate environment: persons, objects and symbols" (Eriksson et al., 2018, p. 420).

Compared to the original theory, the PPCT model tries to rule out further why different developmental outcomes vary between individuals by considering each of the model's components – process, person, context, and time (Eriksson et al., 2018, p. 420). Of particular importance to this study is the inclusion of 'person' which addresses "how individual characteristics influence proximal processes, such as assessing how age, gender, temperament, intelligence, etc. influence these activities and interactions" (Eriksson et al., 2018, p. 420). Although the original theory had already placed individuals at the center of the interconnected systems, the PPCT model tries to more fully account for individual differences in outcomes (Eriksson et al., 2018, p. 420). In considering the role of biology, the PPCT model takes into consideration personal characteristics, such as age, gender, and personal appearance (Rosa & Tudge, 2013, p. 253), of which gender is one of this study's variables.

As for the other elements of PPCT, briefly, process includes "regularly occurring activities and interactions with significant persons, objects and symbols in the developing individual's lives" (Eriksson et al., 2018, p. 420); context incorporates the first four interrelated systems (levels) previously outlined in EST, but could also include evaluating the influences of different exo- or macrosystems on the proximal processes of interest (Eriksson et al., 2018, pp. 420-421); and, finally, time, which "should be longitudinal (to evaluate the influence of proximal processes, as they are mutually influenced by person characteristics and context, on the developmental outcomes of interest) and should take into account what is occurring, in the group being studied, at the current point of historical time" (Tudge, Mokrova, Hatfield, & Karnik, 2009, p. 202). While all four elements need not

be included in every study that uses PPCT, Bronfenbrenner was straightforward in his position of the model's use: "A study involving the PPCT model should focus on proximal processes, showing how they are influenced both by characteristics of the developing individual and by the context in which they occur and showing how they are implicated in relevant outcomes" (Tudge et al., 2009, p. 207).

Application of Theory to Study[13]

As mentioned earlier, what prompted this investigation was the belief some teachers in the UAE had about SMR inequity. As Bronfenbrenner (1979) wrote: "If men define situations as real, they are real in their consequences" (p. 23). While I would amend this to read 'a person', the sentiment is the same. Whether or not SMR, as in the case of RQ3, affect SRA is almost immaterial as the consequences such a belief can have – even when not based on reality – are real regardless. Despite this if not because of it, a major aim of this study is to discover the relationship between SMR and SRA in the UAE in order to address the belief head on – to help dispel it if findings suggest it is false or to incite discussion if they suggest it is true.

Based on Bronfenbrenner's earlier guidance, using the last transformation of his theory which includes the PPCT model, application will focus on the elements of person and context and the accompanying proximal processes as they relate to the variables of interest in this study. RQ3 incorporates all of the same variables as were used for RQs 1 and 2; however, for RQ3, as mentioned in the introduction, variables are organized according to level – student and school – for inclusion in the HLM. To that end, this study blends two of EST's research study levels – micro-research studies (Level 1) and macro-research studies (Level 4) (Onwuegbuzie et al., 2013, p. 5) – by focusing on student-level and school-level factors and their impact on SRA. Each of the variables will be looked at individually as they relate to the first three elements of the PPCT model – process, person, and context. However, for the sake of explanation as related to this study, the model elements will be looked at in the following order: person, context, and process.

Person

In Bronfenbrenner's original theory, the individual (child) was at the center of the interconnected systems. However, in later developments of his theory, the role of the person is incorporated more explicitly, and individual attributes are focused on more clearly. Thus, using EST in combination with the PPCT model, student gender would operationalize the 'person' element. The degree of impact, e.g., of school material resources on SRA, will vary based on these individual background characteristics.

Context

Based on Bronfenbrenner's original theory, all of the school-level variables in addition to the student-level SES variable are theorized to operationalize either the microsystem (Level 1) or the macrosystem (Level 4). As a reminder, the microsystem (Level 1) is the individual's immediate environments, such as home, neighborhood, school, and friends' homes. The macrosystem (Level 4) includes the belief systems and cultural norms of a given society that indirectly influence a person, and SES is included. These school-level variables include the two school material resource indices, the two school location variables, and the teacher quality variable.

Although most of the school-level variables would be included in Level 1, it could be argued that the school-level variables of emirate and urbanization level operationalize Level 4. As previously mentioned, a key aspect of EST is the interconnected nature of the various settings. In the case of this

---

[13] See Analytical Approach – Hierarchical Linear Modeling for further details of how the study design relates to EST. For additional consideration of EST's levels and elements, see also Conclusions and Recommendations / EST Framework.

study, lower-level systems (i.e., microsystems) will be influenced by higher-level systems (i.e., macrosystems). Using the school location variables of emirate and urbanization level as examples, though they are part of the student's immediate environment, they are impacted by macrosystem concerns. For instance, "(H)aving an understanding that a school is located in a working-class community that generally values when all members of a family are active in raising younger children may make it easier for teachers to incorporate the broader community's values and norms into the students' education" (Dooley, 2018, p. 499). Because all contexts impact the student, the outside forces are inseparable from the school context (Dooley, 2018, p. 499). Therefore, "contexts that affect the student outside of school will also play an important role within the school environment" (Dooley, 2018, p. 499). By acknowledging the broader ecological framework in which the student exists, the teacher might begin to understand the larger spectrum of forces outside of school that may be affecting the student's behavior (Dooley, 2018, p. 499). "(S)chools are embedded within the broader community; thus, the impact of their contexts cannot be separated from one another" (Dooley, 2018, p. 499).

Another societal factor that influences students and their environments and which is included in Level 4 (macrosystem) is SES. While it could be argued that student-level SES is an individual student attribute (and thus in the realm of individual or person), according to EST, SES is included in the macrosystem. In many ways, this is a very appealing placement in my opinion. While still evaluative, a person is not rendered inherently rich or poor by aligning a SES value label (e.g., poor) with the person him/herself. Instead of being a reflection of the individual person, it is more a measure of the society in which one lives. An individual is only poor (or middle-income, etc.) if that is how society views the person. An individual is poor in comparison to others and based on how the society of which he/she is a part, or compared to, measures SES. This is not a static or terminal assessment of the individual, however. Someone who is poor in one context may well be deemed rich in another (or vice versa). This may follow on from the notion of 'subculture' (Dooley, 2018, p. 498); e.g., related to wealth, this could mean rich, poor, middle-income, etc.

In this study, school type, which is a school-level variable, is based on gender. Therefore, since gender is an individual ('person') attribute, one might argue that school type should not be a Level 1 factor according to EST. That said, similar to SES, an alternative placement might be Level 4 due to the values and norms associated with gender roles in society, particularly relevant perhaps for the single-sex schools. Since the school type factor includes co-ed as well as single-sex schools, for the purposes of this study, school type is theorized to operationalize the students' school microsystem (Level 1).

The interplay of the school setting – the investigation of children within another system in which they spend time (i.e., school when away from home) (Onwuegbuzie et al., 2013, p. 5) – along with individual student attributes (e.g., background/personal characteristics) and macrosystem influences are expected to interact and impact student learning and development.

Process
Although the notion of process was already established in the original version of EST, it was emphasized further as the first element of the PPCT model with its focus on proximal processes. In Plomin and McClern's *Nature, nurture, and psychology*, Bronfenbrenner and Ceci (1993) offer two propositions which describe proximal processes and the manner in which they operate:

> Proposition 1: (H)uman development takes place through processes of progressively more complex reciprocal interaction between an active evolving biopsychological human organism and the persons, objects, and symbols in its immediate environment. To be effective, the interaction must occur on a fairly regular basis over extended periods of time. Such enduring forms of interaction in the immediate environment are referred to henceforth as proximal processes. (cited in Rosa & Tudge, 2013, p. 252)

Proposition 2: The form, power, content, and direction of the proximal processes that affect development vary systematically as a joint function of the characteristics of the developing person and the environment (both immediate and more remote) in which the processes are taking place and the nature of the developmental outcomes under consideration. (cited in Rosa & Tudge, 2013, p. 252)

Based on these propositions, the processes of interaction and development are described as follows: As active participants in ever-changing environments, each unique being develops and changes through interactions with persons, objects, and symbols in his/her immediate environment. The interactions are of a reciprocal nature influenced by characteristics of the person as well as the environments in which the proximal processes transpire.

Though each of the variables mentioned previously has been designated as belonging to a certain level or levels, as evidenced by the discussion above, it would be a mistake to think they act in isolation of one another or of the remaining levels/elements of EST/the PPCT model. Early on in the outlining of his theory, Bronfenbrenner made clear this interconnectivity of settings and the process-like nature of that mutuality. Specifically, his first definition reads: "The ecology of human development involves the scientific study of the progressive, mutual accommodation between an active, growing human being and the changing properties of the immediate settings in which the developing person lives, as this process is affected by relations between these settings, and by the larger contexts in which the settings are embedded" (Bronfenbrenner, 1979, p. 21). Note that the individual child (person) is at the center of Bronfenbrenner's original theory. His first proposition follows on from above: "In ecological research, the properties of the person and of the environment, the structure of environmental settings, and the processes taking place within and between them must be viewed as interdependent and analyzed in system terms" (Bronfenbrenner, 1979, p. 41). Because of this, if we are to interpret this study's findings through the lens of EST, taking the student's individual characteristics along with school and societal factors into mutual consideration seems essential.

For all of the reasons outlined, EST is a suitable if not profitable choice of theoretical framework for this study. Using EST as the theoretical lens, it is expected that the independent variables will influence the dependent variable because the students' microsystems are impacted by microsystem and macrosystem influences (Onwuegbuzie et al., 2013). For children, school is a microsystem (one of their immediate environments), and the interaction between this microsystem and other systems (e.g., the macrosystem as in this study) can influence students' development in combination with their individual qualities. SRA may be negatively impacted by a lack of material resources at home or at school. However, the degree to which a particular student's reading achievement is affected will also depend on individual characteristics of the child [e.g., as related to this study, his/her gender and SES (which includes parents' educational level(s) and books in the home)] as well as interactions with other environments in which he/she exists (e.g., home and school). Further, as noted earlier, not only can one's own SES influence achievement but so too can one's classmates' (see Socioeconomic Status / Parents' Education). Chiu and Chow (2015) found that classmates' SES affects a student's reading achievement more than some other classmate characteristics. Following on from EST, children are influenced by their immediate environment as well as surrounding environments. At school, this would include classmates/peers whose individual SES could impact that of other children, and vice versa.

Method

In this section, the data is first described, and then the operationalization of variables explained[14].

---

[14] See Table A1 for a list of variables and their descriptions.

Data

This study makes use of cross-sectional survey data, details of which follow below.

PISA Data

The UAE has participated in the OECD's PISA since 2012 as a partner (non-OECD) country[15]. For this study, the UAE's PISA 2018 school and student data sets[16], which include data for all seven emirates, were sorted, merged, and used for analyses. The choice of PISA data was intentional, not merely one of convenience. In comparison to the IEA's PIRLS (an alternative ILSA that also assesses reading), PISA is more current. Whereas PIRLS is administered every five years (the most recent being 2016), PISA is administered every three years (the most recent being 2018). Further, PISA's schedule aligns more closely with 2016, the UAE's Year of Reading and when new textbooks started to be introduced into classrooms in the UAE.

Based on OECD guidance, data collected through the school questionnaire has to be analyzed at the student level. Once the student and school data files are merged, "school data can be analysed as any student-level variables since the school-level variables are now considered as attributes of students" (OECD, 2009, p. 146).

Two of the school-level variables in this study use principal data; a principal's perceptions of his/her school's capacity to provide instruction being hindered due to the quantity and quality of educational materials as well as perceptions of the school's capacity to enhance learning and teaching using digital devices. While teacher data could also have been useful to get a fuller understanding of this topic, it would have been too much for this study (see Conclusions and Recommendations).

Participants

According to PISA 2018's UAE Country Report, "Some 600 000 students completed the assessment in 2018, representing about 32 million 15-year-olds in the schools of the 79 participating countries and economies. In the United Arab Emirates, 19,277 students, in 760 schools, completed the assessment, representing 54,403 15-year-old students (92% of the total population of 15-year-olds)" (Avvisati, Echazarra, Givord, & Schwabe, 2019, p. 9). For a breakdown of the 54,403 students in the UAE, see Table 1 which has the numbers of students and valid percentages excluding any missing, if applicable.

Sampling

PISA employs a two-stage sampling procedure – first the schools are chosen and then the students. "After the population is defined, school samples are selected with a probability proportional to size. Subsequently, 35 students are randomly selected from each school" (OECD, 2009, p. 144). Larger schools have a higher probability of being selected than small schools; however, students in larger schools have a smaller within-school probability of being selected than students in small schools (OECD, 2009, p. 53). Once chosen, students are randomly given test booklets comprised of different questions. This clustered sampling has consequences for the way statistical analyses are done as most programs like Statistical Package for the Social Sciences (SPSS) make the assumption that simple random sampling was used (OECD, 2009, p. 39). These consequences for the analysis of PISA data include important considerations, including those of weights and plausible values (OECD, 2009, p. 45). See Analytical Approach for information on both.

---

[15] Dubai participated in 2009 and then a year later the rest of the emirates participated in 'PISA 2009+'. The results were merged for a composite UAE ranking. Thereafter, the entire country participated in PISA assessments (Westley, 2017).

[16] PISA 2018 data can be found at: https://www.oecd.org/pisa/data/2018database/

	N	%
Table 1:		
Descriptive statistics		
	<u>N</u>	<u>%</u>
Female	27,572	50.7
Male	26,831	49.3
Abu Dhabi	20,631	37.9
Dubai	14,581	26.8
Sharjah	10,232	18.8
Ajman	2,951	5.4
Umm Al Quwain	576	1.1
Ras Al Khaimah	3,360	6.2
Fujairah	2,072	3.8
A town or smaller (up to 100,000 people)	17,553	34.3
A city (100,000 to about 1,000,000 people)	15,539	30.3
A large city (over 1,000,000 people)	18,155	35.4
Boys' school	6,288	15.6
Girls' school	5,025	12.5
Co-ed school	28,955	71.9

Variables (Definitions and Operationalization)[17]

This section is divided into two parts: student-level variables and school-level variables. Student-level variables include SRA, student gender, and student SES (comprised also of parents' education and books in the home). School-level variables are SMR, school digital devices, school location (including emirate and urbanization level), school type (based on student gender), teacher quality, and the percentage of students from socioeconomically disadvantaged homes.

Student-level Variables

It is well established in the literature that certain variables, such as gender and SES, affect student achievement. While gender is a primary focus of this study, the latter variable is not. However, it is included to account for its influence on student achievement, thereby allowing for a more precise measure of the main variables of interest. The definition and operationalization of each of the variables used in this study is described below.

Student Reading Achievement

As reading achievement was the major domain for PISA 2018 and it is the outcome (dependent) variable for RQs 1 and 3, it is looked at in detail below.

PISA's 2018 Reading Framework

Over the years, the PISA reading framework has evolved, most recently with the 2018 revision[18]. Despite the changes, in terms of longitudinal comparison of rankings, "scores in subsequent cycles of PISA are calibrated so as to be directly comparable to those in previous cycles" (OECD, 2019a, p. 16). New forms of reading that have emerged since the framework was last updated in 2009 were incorporated into the revised PISA 2018 reading framework; in particular, digital reading, using "interactive exercises with several texts to be read in a simulated web environment" ("Reading

---

[17] See also Table A1.

[18] For a summary of the main changes in the reading framework 2000–2015, see Annex Table 2.A.1 (OECD, 2019a, p. 67).

Literacy," n.d.), and the growing diversity of material available in both print and digital forms (OECD, 2019a, p. 15). Further, two new subscales that describe students' literacy with single-source and multiple-source texts were developed (OECD, 2019a, p. 17). Overall, PISA uses a variety of text types (e.g., static/dynamic, continuous/non-continuous, etc.) and situations to assess reading achievement (OECD, 2019a, p. 15).

In relating PISA's reading framework to the theoretical framework used for this study, EST, by incorporating more and different forms of reading as well as a variety of text types and situations, the reading framework more closely mirrors what students will experience outside the classroom, in the 'real world'. In addition to helping prepare students for participation in the working world (as mentioned earlier, a focus of the OECD), by assessing their knowledge and skills from a broader spectrum of tasks and materials, a more comprehensive measure of their ability may be achieved not to mention taking into consideration the range of skills each student will uniquely possess.

Definition
The 2018 definition of reading literacy as taken from PISA's reading framework is: "An individual's capacity to understand, use, evaluate, reflect on and engage with texts in order to achieve one's goals, develop one's knowledge and potential, and participate in society" (OECD, 2019a, p. 27).

Adaptive Testing Approach
With computer-based assessment, PISA is able to implement adaptive testing. For PISA 2018, the major domain of reading "adopted an adaptive approach, whereby students were assigned units based on their performance in earlier units" (OECD, 2019a, p. 14). "Adaptive testing enables higher levels of measurement precision using fewer items per individual student" (OECD, 2019a, p. 53) as student are presented with items that are aligned to their ability level (OECD, 2019a, p. 53). "Adaptive testing has the potential to increase the resolution and sensitivity of the assessment, most particularly at the lower end of the distribution of student performance. For example, students who perform poorly on items that assess their reading fluency will likely struggle on highly complex multiple text items" (OECD, 2019a, p. 53). For a country like the UAE, which ranks below the OECD average for reading, having ability measured more precisely at the lower end of the performance spectrum should be of benefit.

Plausible Values[19]
This study's reading achievement variable, S_READACH, is PISA's overall SRA plausible value 1 (PV1READ). For PISA overall (including all participating countries and economies), the range was 0-887.692; for the UAE, the range was 84.050-814.074 with a mean of 431.051. According to Wu and Adams (2002, as cited in OECD, 2009), "The simplest way to describe plausible values is to say that plausible values are a representation of the range of abilities that a student might reasonably have" (p. 96). Not a direct estimate but rather an estimated probability distribution, PVs are a range of possible values (for an ability) with an estimated probability for each of those values (OECD, 2009, p. 96). Wu and Adams (2002, as cited in OECD, 2009) explain that PVs are "random draws from this (estimated) distribution" (p. 96). While this may seem imprecise to some, in terms of error, PISA is not a high-stakes test for individual advancement; the goal is to assess the knowledge or skills of a population. Therefore, "reducing error in making inferences about the target population is more important than the goal of reducing errors at the individual level" (OECD, 2009, p. 94). For information about using PVs in statistical analyses, see Analytical Approach / Plausible Values.

---

[19] See Validity and Reliability / The Rasch Model for an explanation of the scaling of cognitive data on which PVs are based.

Student Gender
S_GENDER is PISA's student questionnaire question ST004Q1D01T which asks if a student is female or male; 1 = Female, and 2 = Male.

Student Socioeconomic Background
This study looks at SES not only at the individual student level but at the school level as well (see School-level Variables for a discussion of the latter). At the student level, the S_gmSES index was used to analyze SES' effects on SRA. Components of the index include parents' education and occupational status as well as home possessions, which includes books in the home. The definition and operationalization of each of these variables is described below.

Socioeconomic Status
The S_gmSES index is PISA's ESCS variable, the index of economic, social and cultural status (grand-mean centered). It is derived from three variables related to family background: parents' highest level of education (PARED), parents' highest occupational status (HISEI), and home possessions (HOMEPOS – ST011, ST012, and ST013), including books in the home (OECD, 2019c, p. 216). PARED = HISCED recoded into estimated number of years of schooling (see 'Parents' Education'). "The rationale for using these three components is that socioeconomic status is usually seen as based on education, occupational status and income. As no direct income measure is available from the PISA data, the existence of household items is used as proxy for family wealth " (OECD, 2009, p. 472).

"In PISA 2018, ESCS is computed by attributing equal weight to the three standardised components" (OECD, 2019c, p. 217). "(T)he three components were standardised across all countries and economies (both OECD and partner countries/economies), with each country/economy contributing equally" (OECD, 2019c, p. 217). "(T)he final ESCS variable was transformed, with 0 the score of an average OECD student and 1 the standard deviation across equally weighted OECD countries" (OECD, 2019c, p. 217).

Parents' Education
PISA's HISCED index is the highest education of parents based on students' responses to questions ST005-ST008 regarding their parents' education. "In PISA 2018, to avoid issues related to the misreporting of parental education by students, students' answers about post-secondary qualifications were considered only for those students who reported their parents' highest level of schooling to be at least lower secondary education" (OECD, 2019c, p. 216). Parental education indices (separate for father and mother) were constructed by recoding educational qualifications using UNESCO's[20] International Standard Classification of Education (ISCED) 1997. The categories are as follows:

> (0) None; (1) ISCED 1 (primary education); (2) ISCED 2 (lower secondary); (3) ISCED Level 3B or 3C (vocational/pre-vocational upper secondary); (4) ISCED 3A (upper secondary) and/or ISCED 4 (non-tertiary post-secondary); (5) ISCED 5B (vocational tertiary); and (6) ISCED 5A, 6 (theoretically oriented tertiary and post-graduate). (OECD, 2009, p. 458)

For the index, the higher ISCED level of either parent was used.

Parents' Occupational Status
PISA's HISEI index, highest occupational status of parents, corresponds to the higher international socio-economic index of occupational status (ISEI) score of either parent or to the only available parent's ISEI score. "Occupational data for both the student's father and the student's mother were obtained from responses to open-ended questions" (OECD, 2019c, p. 216). Those responses were then coded to the 2008 version of the four-digit international standard classification of occupation (ISCO)

---

[20] UNESCO = United Nations Educational, Scientific and Cultural Organization

codes and mapped to the 2008 version of ISEI (OECD, 2019c, pp. 216-217). "(H)igher ISEI scores indicate higher levels of occupational status" (OECD, 2019c, p. 217).

HOMEPOS is a summary index of all household and possessions items as measured in the PISA student questionnaire. In response to questions ST011Q01TA-ST011Q12TA and ST011Q16TA-ST011Q19TA, "students reported the availability of 16 household items at home, including three country-specific household items that were seen as appropriate measures of family wealth within the country's context" (OECD, 2019c, p. 217). Items such as one's own room, a quiet place to study, works of art, and a link to the internet were included. In addition, students reported the amount of possessions in ST012Q01TA-ST012Q03TA and ST012Q05TA-ST012Q09TA. Items included cars, electronic devices, musical instruments and rooms with a bath or shower. Finally, question ST013Q01TA measures how many books a student has in his/her home. Categories range from 1 = 0-10 books to 6 = More than 500 books.

School-level Variables

The following school-related variables were used in this study: two indices related to SMR, school location, school type, teacher quality, and percentage of students from socioeconomically disadvantaged homes. The definition and operationalization of each of the variables are described below.

School Material Resources

Definition
'School material resources' as it has been operationalized for this study is not well represented in the literature. The focus here is on school **material** resources, defined as including, e.g., textbooks, computers, hard/software, digital devices, other IT equipment, and Internet.

Composite Index on Quantity and Quality of Educational Material
Two questions were used to create a composite index called SCH_MATERIALS_round. Both questions are based on responses in the school questionnaire which measured school principals' perceptions of potential factors hindering school instruction (OECD, 2017, p. 326). They are:

SC017	Is your school's capacity to provide instruction hindered by any of the following issues?
Q05NA	A lack of educational material (e.g. textbooks, IT equipment, library or laboratory material).
Q06NA	Inadequate or poor quality educational material (e.g. textbooks, IT equipment, library or laboratory material).

Response choices were: 1 = Not at all, 2 = Very little, 3 = To some extent, or 4 = A lot. Note that these are the reversed response choice values; i.e., the original value for 'a lot' was 1. The values were recoded to have the higher values equal a 'high' response (Pallant, 2016, p. 87); in this case, that means instruction being hindered a lot due to a lack of or inadequate or poor-quality educational material. Having reversed the scale, it also aligns with how the other index for school material resources was graded (see Digital Devices Index). The index was rounded to aid in the interpretation of results.

Along with the questions above, there were two other questions that together comprised a PISA-made index; however, they were eliminated because they asked about a lack of or inadequate or poor-quality physical infrastructure (e.g., building, grounds, heating/cooling, lighting and acoustic systems) which was not of interest to this study.

Digital Devices Index

PISA 2018 added a digital component to the school questionnaire asking the following four-point Likert scale (1 = Strongly disagree to 4 = Strongly agree) questions:

SC155	To what extent do you agree with the following statements about your school's capacity to enhance learning and teaching using digital devices?
Q01HA	The number of digital devices connected to the Internet is sufficient.
Q02HA	The school's Internet bandwidth or speed is sufficient.
Q03HA	The number of digital devices for instruction is sufficient.
Q04HA	Digital devices at the school are sufficiently powerful in terms of computing capacity.
Q05HA	The availability of adequate software is sufficient.

While the SCH_MATERIALS_round index take a deficit approach to school material resources (i.e., a school's capacity to provide instruction being hindered), the SCH_DDEVICES_round composite index, which is comprised of five of 11 related questions, takes an enrichment approach (i.e., a school's capacity to enhance learning and teaching). Questions SC155Q6-11HA were not included based on the factor analysis loadings (see Principal Component Analysis for details), and because they deal more with the teachers' preparation, skills, and incentives to use the devices rather than the materials themselves. Although this study includes a teacher quality variable, it is not the primary focus. If it had been, these questions may have been used. The index was rounded to aid in the interpretation of results.

School Location

Measures of school location include the emirate in which the school is located as well as an urbanization level variable based on the population of the school's locale.

Emirate

The SCH_EMIRATE variable is the grouped STRATUM variable with each of the seven emirates represented: 1 = Abu Dhabi, 2 = Dubai, 3 = Sharjah, 4 = Ajman, 5 = Umm Al Quwain, 6 = Ras Al Khaimah, and 7 = Fujairah.

Urbanization Level

The SCH_URBANIZATION variable is the recoded SC001Q01TA school questionnaire question. Based on the population of the location of the school, 1 = A town or smaller (up to about 100,000 people), 2 = A city (100,000 to about 1,000,000 people), and 3 = A large city (with over 1,000,000 people).

School Type

The SCH_TYPE variable is based on PISA school questionnaire questions SC002Q01TA and SC002Q02TA which measure the total school enrollment; Q01 is the number of boy students and Q02 is the number of girl students. This study focuses on single-sex schools. Therefore, coding for school type could only be done when it was clearly indicated based on the enrollment numbers that there was single-sex enrollment; e.g., if there was a value for either of the questions but the other question's value was missing, the school was excluded. Table 2 shows how the groups were formed.

Table 2: School type categorization			
	SC002Q01TA (boys)		SC002Q02TA (girls)
Boys' school	>=1	and	0
Girls' school	0	and	>=1
Co-ed school	>=1	and	>=1

Teacher Quality

The SCH_gmTQUAL index is PISA's PROATCE index (grand-mean centered). It is based on school questionnaire SC018Q02TA, which asks for the number of teachers fully certified by the appropriate authority. The index, then, is the proportion of all fully certified teachers. "The credentials defined for 'full' certification depend on school systems, but they may also depend on whether a teacher received a credential from a teacher-education programme, accumulated a minimum number of hours of student-teaching, passed an exam, or some combination of these. In some countries, there is no such certification" (OECD, 2019b, p. 110).

The comparison of ratios related to the proportion of certified teachers is made using overall ratios; the proportion of fully certified teachers is computed by dividing the total number of fully certified teachers in the target population by the total number of teachers in the target population (OECD, 2019b, p. 245). "The overall ratios are computed by first computing the numerator and denominator as the (weighted) sum of school-level totals, then dividing the numerator by the denominator" (OECD, 2019b, p. 245). The number of part-time teachers was weighted by 0.5 and the number of full-time teachers was weighted by 1.0 (OECD, 2019b, p. 219).

Although some prior research uses teacher quality variables as a measure of school resources, this study does not. Rather, due to this study's focus on physical material resources (objects as opposed to human beings), teacher quality is kept separate to assess its impact on SRA. Further, like many places, there were frequently teacher shortages when I was in the UAE. However, whether a school got a teacher depended on factors such as location, the teacher's gender (e.g., cultural norms make it far more difficult for male teachers to work in girls' schools than female teachers to work in boys' schools, but neither is ideal), subject taught, grade, etc. Therefore, having a teacher (which is essential for measuring teacher quality) can have less to do with monetary investment than it does when measuring the quantity or quality of material resources as defined and measured in this study.

School Percentage of Students from Socioeconomically Disadvantaged Homes

The SCH_gmDISADVANTAGED variable is PISA's SC048Q03NA which asks principals to estimate the percentage of students in their schools who come from socioeconomically disadvantaged homes (grand-mean centered). The range is from 0 to 100. "A socio-economically disadvantaged (advantaged) student is a student in the bottom (top) quarter of the PISA index of economic, social and cultural status (ESCS) in his or her own country/economy" (OECD, 2019c, p. 21). To represent SES, because it is not a variable that can be observed directly, researchers usually employ an indictor or a combination of several indicators to represent it (Xie & Ma, 2019, p. 850). In measuring the percentage of students who fall into the socioeconomically disadvantaged category, this question is acting as a school-level measure of SES. Just as student-level SES can impact student achievement, this variable was included to measure its effect on SRA.

Missing Values

The OECD handled all of the missing values for the PISA variables included in the data sets by coding for valid skip, not applicable, invalid, no response, and system missing. PISA-created indices ESCS and PROATCE had their missing values accounted for as well. Additionally, for ESCS, "values for students with missing PARED, HISEI or HOMEPOS were imputed with predicted values plus a random component based on a regression on the other two variables. If there were missing data on more than one of the three variables, ESCS was not computed and a missing value was assigned for ESCS" (OECD, 2019c, p. 217). Even before that, in order to reduce the number of missing values when creating the HISEI index, "an ISEI value of 17 (equivalent to the ISEI value for ISCO code 9000, corresponding to the major group 'Elementary Occupations') was attributed to pseudo-ISCO codes 9701, 9702 and 9703 ('Doing housework, bringing up children', 'Learning, studying', 'Retired, pensioner, on unemployment benefits')" (OECD, 2019c, p. 217).

However, some missingness had to be dealt with for both of the indices created by the author for this study (SCH_MATERIALS_round and SCH_DDEVICES_round) and the author-created variable SCH_TYPE. Therefore, for the indices, SPSS' RMV command (replace missing values with the series mean) option was used to replace missing values so an overall mean for each variable could be calculated. This included the following PISA questions: SC017Q05NA and SC017Q06NA for SCH_MATERIALS_round and ST155Q01HA – ST155Q05HA for SCH_DDEVICES_round. For SCH_TYPE, missing values were recoded -999.

Grand-mean Centering
The model used for RQ3 included three continuous variables which were all grand-mean centered (S_gmSES, SCH_gmTQUAL, and SCH_gmDISADVANTAGED); meaning, the grand mean is subtracted from all values of the variable (Heck, Thomas, & Tabata, 2011, p. 113). The result is that the sample mean of the predictor variable is redefined to be zero (0.0) (i.e., the sample mean is rescaled to 0.0) (Heck et al., 2011, p. 113). This was done because it can be more difficult to interpret the meaning of an interaction that involves continuous variables that are left in their natural metrics (Heck et al., 2011, p. 114). Using SES as an example, no one has a SES score of 0. Therefore, to help make the interpretation more meaningful, by grand-mean centering the variables, the intercept can be interpreted as the expected value of Y when all the predictors are at their mean values (0) (Heck et al., 2011, p. 113); in other words, when all explanatory variables are equal to zero (i.e., the expected variances for the 'average' individual) (Hox, 2002, cited in Heck et al., 2011, p. 114). "This has the effect of creating a metric for determining how a 1 SD increase in a predictor changes the dependent variable" (Heck et al., 2011, p. 113).

Analytical Approach

Using the PISA 2018 cross-sectional cohort student and school data sets, groups were compared, and variable effects measured using the following:

Plausible Values
PISA materials offer plenty of guidance in regard to using PVs in statistical analyses. Although it is recommended to use five PVs to ensure consistency between results published by the OECD and results published in scientific journals or national reports (OECD, 2009, p. 46), it is stated that "(U)sing one plausible value or five plausible values[21] does not really make a substantial difference on large samples. During the exploratory phase of the data, statistical analyses can be based on a single plausible value" (OECD, 2009, p. 46).

Therefore, for this master's-level study, one PV only is used. This decision was based on the following: 1) "Working with one plausible value instead of five will provide unbiased estimate of population parameters[22] but will not estimate the imputation error that reflects the influence of test unreliability for the parameter estimation. With a large dataset, this imputation error is relatively small" (OECD, 2009, p. 43); 2) the UAE's PISA 2018 sample is large (n = 19,277); and 3) the challenges associated with using macros and replicates, which are needed when using more than one PV. For these reasons, the decision was made to use only one PV.

The Rasch Model
The Rasch Model (RM) is based on item response theory. It is essential to briefly discuss the RM as it relates to student achievement as measured by PISA. "The Rasch Model is designated as a one-parameter [item response] model because item characteristic curves only depend on the item difficulty" (OECD, 2009, p. 92). The RM has become fundamental in educational surveys (OECD,

---

[21] From 2015, PISA calculates 10 PVs.
[22] "A parameter is any summary number, like an average or percentage, that describe the entire population" ("S.1 Basic Terminology," 2020).

2009, p. 92) because it addresses two conflicting demands related to their administration – limited student-level testing time and the broad coverage of the assessment domain (OECD, 2009, p. 89); the latter needed to ensure validity. Related to measurement error (see Validity and Reliability / Measurement Error), after a long testing time, students' results start to be affected by fatigue which can bias the outcomes of the surveys (OECD, 2009, p. 78). Further, in taking away from teaching time, school administrators could refuse to free their students for the very long testing period that would be required if the whole item battery were administered. This would reduce the school participation rate, which in turn might substantially bias the outcomes of the results (OECD, 2009, p. 78).

To address these issues, students are assigned a subset of test items – not the whole item battery – by use of a rotated booklet design (OECD, 2009, p. 89) so "only certain subsamples of students respond to each item" (OECD, 2009, p. 78). Then the "cognitive data in PISA are scaled with the Rasch Model and the performance of students is denoted with plausible values (PVs)" (OECD, 2009, p. 118). "One of the important features of the Rasch Model is that it will create a continuum on which both student performance and item difficulty will be located and a probabilistic function links these two components" (OECD, 2009, p. 81). "If some link items are guaranteed, the Rasch Model will be able to create a scale on which every item and every student will be located" (OECD, 2009, p. 92). "The item difficulty and the student ability are linked by a logistic function. With this function, it is possible to compute the probability that a student succeeds on an item" (OECD, 2009, p. 92).

> The Rasch Model uses the number of correct answers and the difficulties of the items administered to a particular student for his or her ability estimate. Therefore, a Rasch score can be interpreted independently of the item difficulties. As far as all items can be located on the same continuum, the Rasch model can return fully comparable student ability estimates, even if students were assessed with a different subset of items. Note, however, that valid ascertainment of the student's Rasch score depends on knowing the item difficulties. (OECD, 2009, p. 90)

Weights
Weighting is a way to adjust for sampling bias. "Weighting a sample should make it more representative of the population it is designed to represent so that reliable estimates can be made from the sample to the population" (De Vaus, 2001, p. 151). All SPSS analyses for this study were done using the W_FSTUWT variable to weight the data. The W_FSTUWT variable is the full student sampling weight variable that must be used when analyzing student-level data (OECD, 2009).

T-test
For RQ1, an independent-samples t-test was conducted to compare the reading achievement scores for females and males. "An independent-samples t-test is used when you want to compare the mean score, on some *continuous* variable, for *two* different groups of participants" (Pallant, 2016, p. 244, emphases in original). The independent variable must be categorical, and the dependent variable must be continuous. In running the t-test, the purpose was to test the probability that the two sets of reading achievement scores (for females and males) came from the same population (Pallant, 2016, p. 245).

Principal Component Analysis
The purpose of employing principal component analysis (PCA) was to see which questions naturally grouped together. Five items were subjected to PCA using IBM SPSS Statistics Version 26. The five PISA questions taken from the school questionnaire were SC155Q01HA – SC155Q05HA (see Digital Devices Index for the wording of the questions). They ask about the school's capacity to enhance teaching and learning using digital devices. The reason these questions were included was because they reflect digital devices themselves, not a teacher's ability to use them, for example. These five questions deal with the quantity and capacity of the physical devices as compared to a teacher's

ability and motivation to use and learn how to use the digital devices, on which the other six questions focus. For this reason, the other six questions were excluded. While the other items may certainly be important, they are not the focus of this study. Similar to the composite School Materials Index which measures perceptions related to the quantity and quality of the school materials themselves (not a teacher's ability to use the materials), this study focuses on the materials themselves. As was noted in the literature review, the definition of 'school material resources' varies widely in the literature. As defined for this study, the other questions are not as relevant. Finally, there is also a separate teacher quality variable to assess this related measure.

Prior to performing PCA, the suitability of data for factor analysis was assessed. Inspection of the correlation matrix revealed all coefficients were .3 and above (.484 - .726) (Pallant, 2016, p. 187). The Kaiser-Meyer-Olikin value was .854, exceeding the recommended value of .6 (Pallant, 2016, p. 187) and Bartlett's Test of Sphericity reached statistical significance at $p < .001$ (Pallant, 2016, p. 187), supporting the factorability of the correlation matrix ($> .3$).

PCA revealed the presence of one component with an eigenvalue exceeding 1, explaining 70.079% of the variance. An inspection of the screenplot revealed a clear break after the first component. Using Catell's scree test, it was decided to retain the one component for further investigation (Pallant, 2016, p. 193). Because only one component was extracted, the solution could not be rotated. See Figure A1 for the screeplot and Table 3 for the unrotated loadings from the component matrix.

Table 3: Table of unrotated loading from component matrix	
	Component 1
SMEAN(SC155Q01HA)	.873
SMEAN(SC155Q05HA)	.856
SMEAN(SC155Q03HA)	.848
SMEAN(SC155Q02HA)	.846
SMEAN(SC155Q04HA)	.757

PLUM Ordinal Regression

For RQ2, Polytomous Universal Model (PLUM) ordinal regressions were performed to assess the impact of three factors on both dependent variables – SCH_MATERIALS_round and SCH_DDEVICES_round. Since both outcome variables used Likert scale responses, ordinal regression was most appropriate as it does not ignore the ordering of variables which is a key characteristic of 4-point scale ordinal variables (Muijs, 2011, p. 165). As for how PLUM ordinal regression is different, "(I)nstead of considering the probability of an individual event …, it considers the probability of that event and all events that are ordered before it" (Muijs, 2011, p. 165). The method is "based on probabilities of reaching thresholds of the dependent depending on the response to the independent variable" (Muijs, 2011, p. 166).

Hierarchical Linear Modeling

For RQ3, two-level multilevel regression modeling, also known as hierarchical linear modeling (HLM), was employed. From the corpus, Odden et al. (2004), in explaining the merits of HLM, say it is particularly well suited for analyzing the variables in an integrated multilevel model because it takes into account the nested nature of the data (e.g., students in classrooms in schools) (p. 21). Although classroom-level variables were not included in this study, HLM makes for more precise estimates of school and classroom effects since achievement can be explained as a function of classroom or school while taking into account variance within those classrooms or schools at the student level. It can also estimate between- and within-group (e.g., within and between classrooms) variance at the same time (Odden et al., 2004, p. 21). In their own study, Odden et al. (2004)

suggested some standard HLM models that account for the nested nature of variables (p. 5). They asserted that using HLM in combination with the model proposed could produce substantial and solid findings about the impact of many student-, classroom- and school-level factors on student achievement (Odden et al., 2004, p. 25).

The aim of RQ3 is to determine the effect(s) SMR have on SRA in the UAE in order to discover which of prior research's findings hold true for the UAE. To that end, a two-level multilevel regression model was constructed. The SMR variables were SCH_MATERIALS_round and SCH_DDEVICES_round. The other controlled variables were S_GENDER, S_gmSES, SCH_TYPE, SCH_EMIRATE, SCH_URBANIZATION, SCH_gmTQUAL, and SCH_gmDISADVANTAGED. Due to the 'nested' nature of the data (i.e., students in schools), the variables used as potential predictors in this study were divided into two levels – student and school. The two student-level variables were S_GENDER and S_gmSES. The remaining variables were all school-level variables. For more information, see the results section for RQ3.

Due to the large sample size and nested nature of the data, maximum likelihood estimation[23] (ML) was chosen instead of restricted maximum likelihood (REML). "REML is often preferred when the sample size is small" (Raudenbush & Bryk, 2002, as cited in "Statistical Computing Workshop: Using the SPSS Mixed Command," 2020). Further, "comparing nested models should only be done when using ML estimation" (Heck et al., 2011, p. 78).

Given EST's focus on process and interdependency, HLM lends itself to use in a study such as this one. The interconnected nature of the predictor variables can be seen clearly in the model. Level 1 variables can affect Level 1 variance as well as Level 2. Considering both student- and school-level variables allows the researcher to see the levels functioning in concert. For these reasons – the influences of the ecosystems on a child's experiences and development, EST was chosen in combination with HLM.

Limitations

As with all research, this study has limitations that must be considered when evaluating its findings and conclusions. These include: the use of HLM and method-related limitations (see Validity and Reliability for the latter).

Hierarchical Linear Modeling

When discussing the limitations of EPFs, Della Sala et al. (2017) explained that while "multilevel models are better suited to disentangle the effects of resources on achievement" (p. 189), "HLM does not account for latent variables (e.g., independent variables with similar characteristics)" (p. 189). For their study, they used Mplus for SEM, an alternative software program to using PISA's SPSS macros. While SEM using Mplus may provide some additional advantages, that is beyond the scope of this thesis.

Validity and Reliability

This study made use of PISA data. Due to the nature of cross-sectional survey data, students' prior achievement could not be controlled for. In addition, while analyses can establish whether or not a relationship exists, causal attribution is not possible; why a change did or did not occur cannot be asserted. Further, while multiple variables were included for analysis, extraneous variables (other variables not analyzed in this study) cannot be ruled out. Finally, this study measures SRA for

---

[23] "Maximum likelihood estimation is a method that determines values for the parameters of a model. The parameter values are found such that they maximise the likelihood that the process described by the model produced the data that were actually observed" (Brooks-Bartlett, 2018).

students in the UAE. Had this study chosen math or science as the dependent variable, prior research suggests the results would be different. And given the importance of culture, results for students in this context may not be generalizable to other contexts.

For a thorough breakdown of steps the OECD takes to ensure PISA's validity and reliability, see the PISA 2018 Assessment and Analytical Framework (OECD, 2019a) or the Pisa Data Analysis Manual SPSS (OECD, 2009). Both are available from the OECD's website (www.oecd.org/pisa/). Also, the Analytical Approach section addresses issues of validity and reliability as well.

Measurement Error
As mentioned when describing the Rasch Model, there is substantial measurement error due to broad education measures which might be affected by the students' mental and/or physical dispositions on the day of the assessment as well as testing conditions which might affect the results (OECD, 2009, p. 96). Further, "(i)n education, the measurement error is not always independent of the proficiency level of the students. It may be smaller for average students, and larger for low and high achievers, depending on the test average difficulty" (OECD, 2009, p. 96). That said, as mentioned under Plausible Values, PISA's goal is not assessing individual student-level achievement but rather making inferences about the knowledge and skills of a population.

Ethical Considerations

In line with the Swedish Research Council (*Good Research Practice*, 2017) and the updated regulations to the Ethics Review Act on research relating to people that went into effect from January 1, 2020, measures were taken to safeguard the anonymity of personal data. These included the "collection, registration, organization, storage, processing or changing, recycling, gathering, use, disclosure through transmission, dissemination or other provision of data, compilation or matching, blocking, obliteration or destruction" (O. Franck, personal communication, November 19, 2019) of coded or encrypted personal data that can be traced directly or indirectly to a physical person who is currently alive regardless of whether or not I personally have access to the information (O. Franck, personal communication, November 19, 2019). Though PISA data is freely available for download on the OECD's website, data was downloaded, stored, and accessed from a password-protected laptop. There was no external download or storage of data.

In addition to the measures taken above, issues related to the use of big data/ILSA data as well as using this study's results for comparative purposes will be addressed in the final section of this paper.

Presentation of Research Results[24]

Each RQ and its results will be taken in turn.

RQ1

How do girls compare to boys in overall student reading achievement in the UAE?

Results
An independent-samples t-test was conducted to compare the reading achievement scores for females and males. There was a significant difference in scores for females ($M = 459.192$, SD = 100.917) and males ($M = 402.132$, $SD = 117.990$; $t (52657.537) = 60.543$, p < .001, two-tailed). The magnitude of

---

[24] In the spirit of full transparency, all of the SPSS syntax used to run this study's analyses is available for review. Contact the author at LGOGLIOTTINAZELL@gmail.com for further details.

the differences in the means (mean difference = 57.060, 95% *CI:* 55.212 to 58.907) was moderate/medium (eta squared = .063; Cohen's d = 0.520)[25].

Key Finding
Similar to many other countries and consistent with prior UAE research, girls outperformed boys in reading achievement in the UAE.

RQ2

In the UAE, how do school location and school type relate to school material resources?

Results
The results from each regression will be looked at in turn. See Tables A2 and A3 for the crosstabulations.

SCH_MATERIALS_round
PLUM ordinal regression was performed to assess the impact of three factors on the likelihood that the respondents would report that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material. The model contained three independent variables: school emirate, school urbanization level, and school type. A chi-square test was conducted to look at improvement in prediction compared to the baseline model. The full model containing all predictors was statistically significant, X^2 (10, n = 40,226.548) = 2822.052, p < .001, indicating that the model fit better than the baseline model with no predictors. Both Pearson (.001) and Deviance (.001) were significant indicating the observed and expected counts were different (respondents did not give the answer predicted based on the model), so the model does not fit well. In other words, the values predicted by the model were not close to the actual ones observed in the population. Although Muijs (2011) says a good model fit is less likely with a large sample (p. 170), since Cox and Snell R square (.07) and Nagelkerke R squared (.08) also indicate the improvement the predictors gave the model is poor (Muijs, 2011, p. 165), the conclusion is that, while the model fits the data significantly better than a model with no predictors at all, it still does not fit the data well (Muijs, 2011, p. 170).

All the relationships examined were found to be statistically significant at p < .001 (the exception being Umm Al Quwain at p < .047; however, this is still within the .05 limit for significance). Table 4 lists all the parameter estimates.

Starting with school type as that is a focus of this study, the results were:

School Type
Using co-ed schools as the reference group, boys' schools (-.814) were less likely to report that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material. Girls' schools were even more less likely with a coefficient of -.867. Therefore, co-ed schools were more likely to report that their school's capacity to provide instruction was hindered when compared to either of the single-sex schools.

Emirate
When compared to the reference group (Fujairah), schools in Abu Dhabi (-1.004) were less likely to report that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material. The coefficients for schools in the rest of the emirates were as follows: Sharjah (-.817), Ajman (-.653), Ras Al Khaimah (-.451), Dubai (-.324), and Umm Al Quwain (.283). This means, when compared to Fujairah, schools in all of the other emirates – with the exception of Umm Al Quwain – were less likely to report that their school's

---

[25] See Appendix B for calculations.

capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material. Umm Al Quwain was the only emirate reporting a beta value indicating schools in that emirate were more likely to report that their school's capacity to provide instruction was hindered. Of potential interest, though not the smallest emirate in terms of land size, Umm Al Quwain is the least populous emirate in the UAE ("United Arab Emirates Population (Live)," n.d.).

Urbanization Level

Compared to the reference group (a large city with over 1,000,000 people), schools located in towns or smaller with up to about 100,000 people (-.194) were less likely to report that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material. If the school was located in a city with 100,000 to about 1,000,000 people, the coefficient was -.130, a bit less likely than schools in towns or smaller. Put another way, schools in large cities with over 1,000,000 people were more likely to report that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material when compared to the other two categories.

	B	SE	Wald	Df	P	CI lower	CI upper
Abu Dhabi	-1.004	.068	215.966	1	.001	-1.138	-.870
Dubai	-.324	.072	20.173	1	.001	-.466	-.183
Sharjah	-.817	.071	131.166	1	.001	-.957	-.677
Ajman	-.653	.082	64.138	1	.001	-.813	-.493
Umm Al Quwain	.283	.143	3.935	1	.047	.003	.562
Ras Al Khaimah	-.451	.079	32.820	1	.001	-.606	-.297
Fujairah	0	.	.	0	.	.	.
A town or smaller	-.194	.030	40.581	1	.001	-.254	-.134
A city	-.130	.029	20.558	1	.001	-.186	-.074
A large city	.	.	.	0	.	.	.
Boys' school	-.814	.029	767.792	1	.001	-.872	-.756
Girls' school	-.867	.033	703.733	1	.001	-.931	-.803
Co-ed school	.	.	.	0	.	.	.

Table 4:
Parameter estimates for instruction hindered by educational material

SCH_DDEVICES_round

PLUM ordinal regression was performed to assess the impact of three factors on the likelihood that the respondents would report that their school's capacity to enhance learning and teaching using digital devices was sufficient. The model contained three independent variables: school emirate, school urbanization level, and school type. A chi-square test was conducted to look at improvement in prediction compared to the baseline model. The full model containing all predictors was statistically significant, X^2 (10, n = 40,268.145) = 2412.346, p < .001, indicating that the model fit better than the baseline model with no predictors. Both Pearson (.001) and Deviance (.001) were significant indicating the observed and expected counts were different (respondents did not give the answer predicted based on the model), so the model does not fit well. In other words, the values predicted by the model were not close to those observed in the population. Although Muijs (2011) says you are less likely to get to get a good model fit with a large sample (p. 170), since Cox and Snell R square (.06) and Nagelkerke R squared (.07) also indicate the improvement the predictors gave the model is poor (Muijs, 2011, p. 165), the conclusion is that, while the model fits the data significantly better than a model with no predictors at all, it still does not fit the data well (Muijs, 2011, p. 170).

Table 5: Parameter estimates for school's capacity to enhance learning and teaching using digital devices							
	B	SE	Wald	df	P	CI lower	CI upper
Abu Dhabi	-.397	.060	43.935	1	.001	-.514	-.280
Dubai	.521	.063	68.707	1	.001	.398	.644
Sharjah	-.442	.063	49.608	1	.001	-.565	-.319
Ajman	.476	.073	42.326	1	.001	.333	.619
Umm Al Quwain	.734	.117	39.324	1	.001	.504	.963
Ras Al Khaimah	-.100	.070	2.032	1	.154	-.238	.038
Fujairah	.	.	.	0	.	.	.
A town or smaller	.236	.028	71.080	1	.001	.181	.291
A city	.417	.026	259.719	1	.001	.367	.468
A large city	.	.	.	0	.	.	.
Boys' school	-.462	.029	255.574	1	.001	-.519	-.406
Girls' school	-.341	.032	112.568	1	.001	-.404	-.278
Co-ed school	.	.	.	0	.	.	.

Table 5 lists all the parameter estimates. All but one of the relationships examined were found to be statistically significant. In decreasing order of coefficient size starting with school type, the results were:

School Type
Using co-ed schools as the reference group, boys' schools (-.462) were even more less likely than girls' schools (-.341) to report that their school's capacity to enhance learning and teaching using digital devices was sufficient. In other words, co-ed schools were more likely than either of the single-sex school types to report that their school's capacity to enhance learning and teaching using digital devices was sufficient.

Emirate
When compared to the reference group (Fujairah), schools in Umm Al Quwain (.734), Dubai (.521) and Ajman (.476) were more likely to report that their school's capacity to enhance learning and teaching using digital devices was sufficient. On the other hand, when compared with Fujairah, schools in Sharjah (-.442) and Abu Dhabi (-.397) were less likely to report that their school's capacity to enhance learning and teaching using digital services was sufficient. The relationship for Ras Al Khaimah was not significant.

Urbanization Level
Compared to the reference group (a large city with over 1,000,000 people), both schools in cities with 100,000 to about 1,000,000 people (.417) and schools located in towns or smaller with up to about 100,000 people (.236) were more likely to report that their school's capacity to enhance learning and teaching using digital devices was sufficient. Put another way, schools located in places other than large cities were more likely than schools in large cities to report that their school's capacity to enhance learning and teaching using digital devices was sufficient.

Key Findings
In total, all relationships but one for digital devices (schools in Ras Al Khaimah) were found to be statistically significant; however, both models did a poor job of fitting the data despite being significantly better than the no-predictors model.

Combining both of the indices, the following findings were key:

As the reference group, co-ed schools were most likely to report that their school's capacity to enhance learning and teaching using digital devices was sufficient. However, they were also most likely to report that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material.

When compared to the reference group, schools in Umm Al Quwain were the only ones to report that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material. Yet, in regard to digital devices, when compared to the reference group, schools in Umm Al Quwain were one of only three of the seven emirates to report that their school's capacity to enhance learning and teaching using digital devices was sufficient; Dubai and Ajman were the other two.

Compared to the reference group, schools in large cities were most likely to report that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material, as well as least likely to report that their school's capacity to enhance learning and teaching using digital devices was sufficient.

RQ3

In the UAE, how do school material resources relate to student reading achievement with other variables being controlled?

Results

As mentioned before, the aim of RQ3 was to determine the effect(s) SMR have on SRA in the UAE; to discover which of prior research's findings hold true for the UAE. To that end, a two-level multilevel regression model was constructed. HLM estimates three effects – the intercept[26], the between-school variation in intercepts, and the variation in individual scores within schools. To do this, the model is broken down into three parts: the Null model, the Level 1 model, and the Level 2 model. The Null model does not have any predictors and measures variability in intercepts within and between schools. The Level 1 model is the individual level random intercept model, and it includes student-level predictors. Finally, the Level 2 model is the group level random intercept model, and it adds school-level predictors. Each of these levels will be looked at in detail.

Null Model

The Null model partitions variance in the outcome into its within- (Level 1) and between-groups (Level 2) components (Heck et al., 2011, p. 73), which helps determine how much variance lies between schools in the sample (Heck et al., 2011, p. 73). As a result, the Null model "provides an estimated mean achievement score for all schools" (Heck et al., 2011, p. 73).

The results of the Null, or no-predictors, model suggest that the development of a multilevel model is warranted because intercepts vary significantly across schools (Wald Z[27] = 18.359, p < .001), and the ICC suggests that about 48% of the total variability in reading scores lies between schools[28].

The ICC, or intraclass correlation, "is the ratio of between-groups variance to the total variance" (Heck et al., 2011, p. 74); stated another way, it is the initial variability in SRA observed between schools. In addition, the ICC "provides a sense of the degree to which differences in the outcome Y exist between Level 2 units" (Heck et al., 2011, p. 79), so "it helps answer the question of the

---

[26] The intercept is the expected mean value of Y (dependent variable) when X (independent variable) equals a certain value (e.g., zero or a mean value) (Grace-Martin, 2020).

[27] The Wald Z test provides a Z statistic summarizing the ratio of the estimate to its SE. A Z statistic gives an idea of how far a data point is from the mean. It measures how many standard deviations above or below the population mean a raw score is ("Z-Score: Definition, Formula and Calculation," 2020).

[28] See Appendix B for calculation of ICC.

existence or nonexistence of meaningful differences in outcomes between the Level 2 units" (Heck et al., 2011, p. 79). The higher the ICC, the more homogenous the units (i.e., there exists substantial variability between schools). If the ICC is quite small (~0.05), there would be little advantage to conducting a multilevel analysis as the higher-level grouping does not affect the estimates in any meaningful way (Heck et al., 2011, p. 74). "In these cases, a single-level analysis conducted at the individual level would suffice" (Heck et al., 2011, p. 74).

The residual parameter describes the variance due to individuals within groups (variance within schools). Results suggest there is significant variance to be explained within groups (Wald Z = 162.999, p < .001). The intercept parameter indicates that the intercepts vary significantly across the sample of schools (i.e., variance between schools).

Level 1 Model
Since the intercept was not 0 (see 'Null Model'), and there was a large F score (24,270.142), defined as further away from 1, this suggests SES is significantly related to reading achievement (Heck et al., 2011, p. 84). Compared to the Null model (414.454 = grand mean), the intercept adjusted for gender and SES now becomes 397.278, a reduction of 17.176. This is the average school mean adjusted for gender and SES.

The output suggests that the addition of the within-group predictors, gender (factor) and SES (covariate), reduces the residual (within-group) variability from 6830.188 in the Null model to 6621.637 in the Level 1 model. After calculating the change in residual (the within-groups portion), the result suggests that student gender and SES background account for about 3.05% of the within-school variability in student scores[29]. The within-school predictor also affects that residual variability in intercepts at the school level. And after calculating the reduction in variance estimate for the within-school and between-schools portions of the model, the result of 21.06% suggests that within-group gender and SES account for around 21% of the between-groups variability in SRA; stated another way, 21% of the variation in means across schools can be attributed to differences in the SES of students in those schools.

Despite a 5.09% ICC reduction after controlling for gender and SES (43.24% down from 48.34%), the data suggest there is still significant variability to be explained both within schools (Wald Z = 160.612, p < .001) and between schools (Wald Z = 18.059, p < .001). The Wald Z test suggests that, even after controlling for student gender and SES within schools, a statistically significant amount of variation in outcomes still remains both within and between schools (Heck et al., 2011, p. 86). Therefore, additional predictors were added to see if they might help to explain this residual variability in intercepts.

Level 2 Model
Building upon the Level 1 model, the Level 2 model added seven school-level predictor variables to try to explain more of the variability in intercepts across schools. In the order all variables were entered (both levels combined), the factors were gender, school emirate, school urbanization level, and school type; and the covariates were the school materials index, the school digital devices index, teacher quality, student SES, and the school percentage of students from disadvantaged homes. The idea behind employing HLM is to examine the change in residual variance that occurs by adding predictors (Heck et al., 2011, p. 121). Starting with the intercept-only (Null) model, "(t)his serves as a baseline against which to evaluate subsequent reduction in variance at each level as other variables are subsequently added to the model" (Heck et al., 2011, p. 121). It is worth noting that, while Level 1 variables can explain (reduce) variance at both Level 1 and at Level 2, variables added at higher levels do not affect the variance present at lower levels (Heck et al., 2011, p. 121).

---

[29] See Appendix B for Level 1 calculations.

The estimates of the variance components suggest that the Level 1 and Level 2 predictors reduce the variance component at the school level substantially (i.e., from 6390.026 in the one-way ANOVA to 2261.894 in the Level 2 model). In comparing Level 1 and Level 2 reductions in variance, the addition of the Level 2 variables explains 43.54% more variance than the Level 1 variables alone. The percentage of explained variance rose from 21.06% at Level 1 to 64.6% at Level 2 (the amount of variance accounted for at Level 2)[30]. This means 64.6% of the variance in Y can be predicted from the independent variables. See Table 6 for a summary of residual and intercept changes for each of the models. Despite a reduction in unexplained variance, the remaining intercept variance is still significant (Wald Z = 12.313, p < .001), which indicates that even after adding all the predictors to the model, there is still variance in intercepts that could be explained across schools by adding additional variables to the model (Heck et al., 2011, p. 86).

Table 6:
Summary of estimates of covariance parameters for reading achievement

	B	SE	Wald Z	p	CI lower	CI upper
Null model						
Residual	6830.188	41.903	162.999	.001	6748.551	6912.813
Intercept (school ID)	6390.026	348.064	18.359	.001	5742.987	7109.965
Level 1 model						
residual	6621.637	41.227	160.612	.001	6541.324	6702.936
intercept (school ID)	5044.097	279.316	18.059	.001	4525.309	5622.360
Level 2 model						
residual	6675.974	56.921	117.284	.001	6565.337	6788.475
intercept (school ID)	2261.894	183.696	12.313	.001	1929.049	2652.169

The addition of the school-level variables changes the intercept to 403.760, up from Level 1's 397.278. The intercept variance is estimated as 2261.894, so the estimate of the SD is 47.559 (square root of the intercept variance estimate) (Seltman, 2018, p. 371). This tells us that for any given group, the individual subjects will have personal intercepts that are up to 47.559 higher or lower than the group average about 68% of the time, and up to 95.118 higher or lower about 95% of the time (Seltman, 2018, p. 371). "This suggests that there are important unmeasured explanatory variables for each subject that raise or lower their performance in a way that appears random because we do not know the value(s) of the missing explanatory variable(s)" (Seltman, 2018, p. 372).

Results for each of the predictors are broken down below. See Table 7 for all of the results[31].

Independent Variables
Of the two SMR indices, only the digital devices index was statistically significant; meaning, after controlling for student gender and individual SES within schools, this variable affected reading achievement.

---

[30] See Appendix B for Level 2 reduction of variance calculation.
[31] The estimates in Table A7 show the relationships between the independent variables and the dependent variable; they tell the amount of increase or decrease in the dependent variable that is predicted for every unit increase in the independent variables, holding all other variables constant. If an independent variable is not significant, the corresponding coefficient for that variable is not significantly different statistically from 0 using alpha of .05 since its p-value is greater than .05 (so not statistically significant at the .05 level).

School Digital Devices

In assessing the relationship between school digital devices and reading achievement, for every unit increase in 'school digital devices', there is a 9.069 increase in the predicted reading achievement score, holding all other variables constant. In other words, the more a school reports that its capacity to enhance learning and teaching using digital devices is sufficient (range 1-4), the higher the predicted reading achievement scores for its student would be; approximately 9 points higher for each unit increase in a school's reported perceived capability of enhancing learning and teaching using digital devices. At $p = .026$ level, we can be more than 97% confident that this increase in reading achievement scores will range from 1.100 to 17.037.

School Material Resources

The relationship between school materials ($p = .118$) and reading achievement was not statistically significant ($p > .05$). Therefore, the data suggests that this variable which measures perceptions of the school's capacity for teaching and learning being hindered by the quantity and quality of its educational materials does not affect reading achievement (there is not any effect on reading achievement).

Control Variables[32]

The relationships between reading achievement and two of the control variables – school emirate ($p = .110 – .805$) and teacher quality ($p = .966$) – were found to be statistically insignificant ($p > .05$). However, the remainder of the relationships were significant ($p < .05$); therefore, the data suggests that they do, in fact, affect reading achievement even after controlling for student gender and individual SES within schools.

Student Gender

For every unit increase in a boy's reading achievement score, there will be a 41.230 increase in the predicted reading achievement scores for a girl, holding all other variables constant. In other words, for boys, for each unit increase, their predicted reading achievement scores would be 41.230 points lower than for females, holding all other variables constant. At the $p < .001$ level, we can be more than 99% confident that girls' scores will be between 37.70 and 44.76 points higher than boys' scores for each unit increase.

School Urbanization Level

Holding all other variables constant, students studying at schools in the reference group (large cities with over 1,000,000 people) will have higher reading achievement scores than students at schools in the other two school urbanization level categories. Students at schools in towns or smaller will score 31.822 points less when compared to the reference group. This is similar to the number of fewer points the students at schools in cities would earn (-30.191).

School Type

Comparing school types, for every unit increase in the co-ed schools (the reference group), there would be decreases in the predicted reading achievement scores for the other two school types, holding all other variables constant. The decrease is more pronounced for boys' schools (-75.634) than for girls' schools (-35.966).

Student SES

As for student SES, for every unit increase, there would also be a 12.046 increase in the predicted reading achievement score, holding all other variables constant. As student SES increases, so does their reading achievement.

---

[32] Although not the focus of this study, due to the lack of UAE-specific research available, results for the control variables are included if only as a basis for comparison with future studies.

Finally, for every unit increase in a school's percentage of students from socioeconomically
disadvantaged homes, there would be a corresponding .436 decrease in the predicted reading
achievement score, holding all other variables constant.

Table 7:
Level 2 estimates of fixed effects

	β	SE	Df	t	p	CI lower	CI upper
Intercept	403.760	19.864	366.409	20.326	.001	364.698	442.822
Female	41.230	1.802	25574.625	22.881	.001	37.698	44.762
Male	.	.	0	.	.	.	.
Abu Dhabi	-9.339	11.642	389.193	-.802	.423	-32.227	13.550
Dubai	3.827	12.433	388.311	.308	.758	-20.616	28.271
Sharjah	-3.210	13.007	382.935	-.247	.805	-28.784	22.365
Ajman	-22.300	15.244	375.573	-1.463	.144	-52.275	7.675
Umm Al Quwain	-31.558	19.679	423.052	-1.604	.110	-70.239	7.123
Ras Al Khaimah	-11.669	12.891	401.311	-.905	.366	-37.012	13.674
Fujairah	.	.	0	.	.		
A town or smaller	-31.822	8.012	352.643	-3.972	.001	-47.580	-16.063
A city	-30.191	7.692	341.569	-3.925	.001	-45.320	-15.061
A large city	.	.	0	.	.	.	.
Boys' school	-75.634	8.479	376.157	-8.920	.001	-92.307	-58.961
Girls' school	-35.966	8.608	372.306	-4.178	.001	-52.893	-19.040
Co-ed school	.	.	0	.	.	.	.
School Materials	5.027	3.204	345.482	1.569	.118	-1.274	11.328
School Digital Devices	9.069	4.052	357.541	2.238	.026	1.100	17.037
Teacher Quality	.610	14.099	365.750	.043	.966	-27.115	28.334
Student SES	12.046	.703	27842.808	17.142	.001	10.669	13.423
% Disadvantaged SES (School)	-.436	.142	364.046	-3.065	.002	-.715	-.156

Key Findings

To summarize the main findings, principals' perceptions of their school being hindered by the
quantity or quality of school materials were not found to be significant predictors of SRA. This
relationship was not significant statistically. However, the measure of school digital devices was
found to be significant. Therefore, schools' perceptions of their capacity to enhance learning and
teaching using digital devices were found to be predictors of SRA. On the scale of 1-4 used to
measure those perceptions, every unit increase corresponded to a 9.069 increase in the predicted
reading achievement score, holding all other variables constant.

As for some of the control variables, although they were not the focus of this study, there were some
interesting findings for the UAE, some of which support earlier research. For example, the higher the
student-level SES, the higher the reading achievement scores; similarly, the higher the percentage of

disadvantaged students, the lower the reading achievement scores. Boys' reading achievement is lower than girls, particularly lower if the boy is attending a boys' school. In terms of location, while emirate was not found to be a predictor of reading achievement scores (it was not statistically significant), schools in large cities have the highest reading achievement compared to schools in cities and towns or smaller. Surprisingly, teacher quality was not significant and, therefore, not a predictor of SRA according to this model and population.

For this study, 'school material resources' was operationalized using two indices – school materials and school digital devices. Based on this study's results, the data suggest that, while the school materials index was not related to reading achievement scores, the school digital devices index was. That said, although variance was reduced by the included variables, there is still variance unaccounted for; meaning, other variables can/should be included to reduce the variance further.

Discussion

In the following section, the theoretical framework is applied to the results.

Application of Theory to Results

As the focus was on SMR and their effect on SRA, the main findings from RQ3 are that, when modeled using HLM and all the included predictors, only the SCH_DDEVICES_round index impacted SRA. The failure of the SCH_MATERIALS_round index to influence SRA in a significant way may, as some corpus research mentioned[33], be associated with the fact that teachers find ways to compensate for lack of materials. In this case, inferior or insufficient quality materials do not impact SRA; however, the ability to enhance learning does have a positive impact on SRA. It follows on from EST that the direction/degree to which a student's reading achievement is affected by the independent variables (if at all) will also depend on the individual characteristics of the child in combination with the interactions with other environments in which he/she exists. In this way, the impact(s) will vary in relation to the child.

Related to the acquisition of such devices, the findings support what has already been well established in the literature that student SES and school-level SES (as measured by the percentage of students from socioeconomically disadvantaged homes) both have an impact on SRA. If students are coming from homes that are socioeconomically disadvantaged, it is likely they will lack at least some resources because their parents cannot afford them. In the case of reading and of the PISA SES index, this may mean, e.g., digital devices or books in the home. In this case, school – another of their immediate environments – may be a place that can supplement that deficiency. Similarly, if a school has a high percentage of students coming from disadvantaged homes, the importance of ensuring they have the digital devices needed to enhance learning seems clear. As mentioned previously, individual student SES can also impact classmates' achievement, so it is for the good of the entire school that it is considered. While school's perceptions of being hindered by the quantity and quality of educational materials was not found to be a statistically significant relationship with SRA, the school's perceptions of being able to enhance learning by using digital devices was. For schools whose capacity to enhance teaching and learning using digital devices is not sufficient, there is an opportunity to raise SRA by investing in such devices as SRA was positively impacted by the sufficient presence/availability of these school material resources.

While this study started off from the premise there may be material resource inequity disadvantaging female students, it is worth considering the current state of boys' education and achievement in the UAE and elsewhere. The other student-level factor, gender, was also shown to affect SRA. Similar to

---

[33] See Ma and Crocker (2007) under Location and School Material Resources as well as Ning et al. (2016) in Achievement and School Material Resources.

many other countries and consistent with prior UAE research, girls outperformed boys in reading achievement. According to EST, male children are more likely to be affected by environmental change than females (Bronfenbrenner, 1979, p. 225). These sex differences were also found in differences between the different school types (which are based on the gender of students attending the school). As Van Hek et al. (2018) suggested, gendered effects may well depend on the country context; therefore, taking the context into consideration is essential when using EST as the theoretical lens.

Returning to the quasi-experimental study Legewie and Diprete (2012) conducted using the German longitudinal ELEMENT dataset of reading and math ability for fourth to sixth graders and the German-I-Plus 2003 data, they found that boys are more sensitive to school resources that create a learning-oriented environment (p. 463). Their study focused on peer socioeconomic composition as the school resource variable; therefore, not the same SMR as used in this study. However, as both gender and SES were included in this study's model, it is worth considering the effects of other SMR, as they suggest. The authors actually argue that their theoretical argument can apply to all kinds of school resources that create a learning-oriented environment even though their findings are limited to the variable tested (Legewie & Diprete, 2012, p. 481). As for the UAE context, investment in digital devices may be one way to address the lower level of achievement for boys. As measured in this study, this would include Internet access, software, and digital devices (e.g., online readers).

Although teacher quality was not found to be a predictor of SRA PVs when modeled with the other variables included, as this conflicts with a lot of prior research, I think it is an area worthy of additional investigation. Teacher quality (as measured by competence) has been found to be one of the most important and consistent factors in student achievement (e.g., see Gustafsson, 2003), so I think these results must be taken with some caution.

Going back to the notion of material resource inequity in the schools, RQ2 addressed this. First, findings from the SCH_MATERIALS_round index, which measures a principal's perceptions that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material:

When compared to Fujairah, schools in all of the other emirates – with the exception of Umm Al Quwain – were less likely to report that their school's capacity to provide instruction was hindered. Therefore, Umm Al Quwain was the only emirate whose schools were more likely to report that their school's capacity to provide instruction was hindered by a lack of educational material or inadequate or poor-quality educational material. Likewise, schools in large cities with over 1,000,000 people were more likely to report that their school's capacity to provide instruction was hindered when compared to the other two categories. Finally, co-ed schools were more likely to report having their instruction hindered when compared to either of the single-sex schools.

Regarding the SCH_DDEVICES_round index, which measures a principal's perceptions that their school's capacity to enhance learning and teaching using digital devices was sufficient:

Abu Dhabi, Dubai, and Sharjah are the three largest emirates and home to the majority of the population. Conversely, Umm Al Quwain is the least-populated emirate and Ajman is the smallest. Even so, Umm Al Quwain, Dubai, and Ajman were all more likely to report that their school's capacity to enhance learning and teaching using digital devices was sufficient when compared to Fujairah (the reference group). Abu Dhabi, home to the capital, and Sharjah both were less likely to report the same thing when compared to Fujairah. This runs counter to my expectations as Abu Dhabi is also the wealthiest of the emirates. When compared to schools in large cities, schools located in places other than large cities were more likely to report that their school's capacity to enhance learning and teaching using digital devices was sufficient. Again, this runs counter to my expectations as large cities are often perceived as having more money and being more developed than smaller

locales. This may be because outfitting so many schools in a large city is harder to do due to the sheer quantity of resources needed to supply all the schools. Finally, co-ed schools were more likely than either of the single-sex school types to report that their school's capacity to enhance learning and teaching using digital devices was sufficient. This may be because many co-ed schools are private and thus charge tuition (fees) which may be used to purchase such devices. In comparing all three school types, girls' schools were a bit more likely than boys' schools to report that their school's capacity to enhance learning and teaching using digital devices was sufficient, but less likely than co-ed schools.

Conclusions and Recommendations

From the start, the lack of UAE-context research was made clear. To that end, this study joins a small but growing amount of research focused on understanding how data such as PISA's can be used to better understand the UAE educational system and perhaps to help further its reforms. Similar to many other countries and consistent with prior UAE research, this study found that girls outperformed boys in reading achievement in the UAE. Less clear from the findings is the association between SMR and the school location and school type variables as the regression models did a poor job of fitting the data despite nearly all the relationships being statistically significant. Results were mixed. While co-ed schools were more likely than either of the single-sex school types to report that their school's capacity to enhance learning and teaching using digital devices was sufficient, they were also more likely to report that their school's capacity to provide instruction was hindered due to the quantity/quality of the material resources when compared to either of the single-sex schools. What can be taken away from this study is that, of the two SMR indices used, the school digital devices index was a significant predictor of SRA. Every unit increase in the index corresponded to a 9.069 increase in the predicted reading achievement score. While the included variables reduced the variance, some unaccounted variance remains which warrants further investigation. In terms of future directions, these findings as well as the questions raised bring forth a number of issues worth considering, including the study's limitations, the use of ILSA data for comparative purposes, and future UAE-specific research, some of which could incorporate variables from other levels of Bronfenbrenner's EST as well as expand on those included here. Each of these areas is looked at below.

Study Limitations

As with all research, this study has limitations. In terms of data, principal data was used for the school-level variables. In the future, PISA UAE teacher survey data or other ILSA teacher data such as the OECD's Teaching and Learning International Survey, or TALIS, could be used to gain further insights about the UAE context. Although principals may be out of touch with the reality 'on the ground', teachers' impressions may be based solely on their own classrooms, not the entire school. While the knowledge on which principals form their opinions may be based on a wider range of information (e.g., across all subjects and in all classrooms) thus giving them a broader sense of the issues than that based on an individual teacher's classroom, including perspectives from the classroom level seems valuable as well. As previously mentioned, HLM lends itself to such an inquiry.

In addition to teacher data, incorporating student data could offer additional insights. As Dickson explained (see Background – UAE Context), the students in her study wanted their voices to be heard. Following on from EST, Bronfenbrenner hypothesized: "The developmental potential of a setting is enhanced to the extent that there exist direct and indirect links to power settings through which participants in the original setting can influence allocation of resources and the making of decisions that are responsive to the needs of the developing person and the efforts of those who act in his behalf" (Bronfenbrenner, 1979, p. 256). As educational bodies are working on behalf of the children they serve, it must not be forgotten that "(t)he developmental potential of a setting varies inversely

with the number of intermediate links in the network chain connecting that setting to settings of power" (Bronfenbrenner, 1979, p. 256).

Finally, this study with its cross-sectional data represents but a snapshot of the UAE educational system. This point should not be forgotten, particularly given the pace of change in the UAE.

ILSA Data

Now with the rabid interest in ILSAs, using data like PISA's is not uncommon, especially for comparative purposes. "The UAE overwhelmingly wishes to be regarded as a modern nation state and as such engages in policies and rhetoric that it believes will signal modernity to the international community" (Ridge, 2009, p. 162). This includes "policy borrowing" (Ridge, 2009, p. 164). One strategy the UAE has used in trying to become one of the top PISA countries has been to attempt to imitate what top-performing countries do presumably in an effort to emulate their success. The choice of policy or educational reform seems to stem from the desire for the UAE to be regarded as 'modern' by the international community (Ridge, 2009, p. 168). This can result in simply transplanting other countries' efforts and initiatives in the UAE context.

I believe there are many hazards associated with policy borrowing that does not take into full consideration the context. Forgetting the cultural component can hinder progress. In my opinion, the cultural component is an important factor being forgotten as to how the UAE uses PISA data/results and in their choice of initiatives to implement (see Lindblad, Pettersson, & Popkewitz, 2018). Noting the importance of taking contextual factors into consideration, Hanushek and Woessmann (2017), when discussing using the same EPF for more than one country, suggested that "There may be reverse causality, and unobserved country differences—e.g., cultural traits or institutional and political factors—may be correlated with both inputs and outcomes" (p. 158). This realization illustrates the point that cultural effects should not be discounted. Although I cannot dictate or ensure how findings from this study will be used or interpreted, in my capacity as the study author, I bring these issues to the fore so they may be considered by readers.

Finally, while valuable and increasingly sophisticated and complex, as mentioned previously, I trust there are perspectives being lost and/or missed despite the range of input we do have access to (e.g., principals, teachers, and parents). Even including teacher and student perspectives, due to the nature of large quantitative survey research data, supplementing such data with data of a more qualitative nature could potentially add more and different levels of understanding and shades of meaning.

UAE-specific Research

One way to address the cultural component would be to engage in more UAE-specific research. Bronfenbrenner discusses "the importance in educational research of investigating systematically the changes in activity that occur from one grade to the next, from one school to another, and for pupils from different socioeconomic, ethnic, and cultural groups within a given educational setting" (Bronfenbrenner, 1979, p. 248). Since there is so little UAE-specific research available now, it behooves researchers to determine what effects various factors and initiatives have in that context. Particularly as the UAE is in the midst of largely (re)defining its educational system, now is the time to investigate since so much money and time are being invested in the reforms. Not only are the consequences great for the students individually, but they are for the entire country as well. While the students' academic achievement is certainly paramount, also are the potentially negative consequences to self-identity, confidence, self-esteem, and drive that might result because of what transpires in the classroom. One could argue these might be even more important due to their possible effects on achievement in addition to the well-being of the child. Bronfenbrenner's EST offers one such framework for motivating choices and evaluating progress.

EST Framework

Two elements of the PPCT model (person and context), which includes two levels of Bronfenbrenner's original theory (1 and 4), were measured in this study; however, the other levels and elements could provide a number of possible future avenues of study, as could expanding some of the others already used. For example, moving out the nested levels, as part of the greater educational system, any one school is affected by decisions made at the ministerial level. As Bronfenbrenner (1979) used the term, the MoE is a power setting and schools are directly affected by MoE decisions. Following EST, if one were to continue focusing on student achievement, exo-research studies (Onwuegbuzie et al., 2013, p. 5) based on EST's Level 3 exosystem could be carried out "whereby one or more persons, groups, or other living organisms are examined within systems by which he/she/they/it might be influenced but of which he/she/they/it does not play an active role" (Onwuegbuzie et al., 2013, p. 5). Even if students and perhaps even schools do not play an active role in the MoE decision-making process, they are impacted by decisions made at this level.

Moving back to EST's Level 4, as a reflection of greater society, the MoE, as is the case of the UAE, along with individual schools, each in its own varying context, will be influenced by societal norms and values. To address these influences, macro-research studies (Onwuegbuzie et al., 2013, p. 5) could be undertaken "whereby one or more persons, groups, or other living organisms are studied within the larger cultural world or society surrounding him/her/them/it" (Onwuegbuzie et al., 2013, p. 5). As noted earlier, culture is an important consideration when trying to implement changes. From choice of initiatives to manner of implementation, cultural norms and values can be critical to successful implementation and adoption. As Dickson (2013) mentions and as has been seen in recent developments in the UAE, there are concerns about outsiders (particularly non-Arabs) coming and making changes to the UAE's educational system, along with the loss of Arabic language and identity as a result of the implementation of English as the medium of instruction in schools and the foreign culture influences that can accompany such change (focusing on the negative effects from cultures that are markedly different/more liberal than Emirati culture).

Finally, looking at the changes to the UAE educational system from an historical perspective, the education students are receiving now is considerably different to that their elders received. Although EST originally included a time element, it was not until a further development that Bronfenbrenner included the fifth and final system, chronosystem, which looks at historical influences on development as well as the time dimension. Although children do not yet have extensive life histories, even so, in a relatively short period of time, major life events and transitions may occur that impact them; e.g., parents divorcing, a move, etc. Expanding the time element further, the PPCT model broadens the concept of time "to include what happens over the course of both ontogenetic and historical time" (Rosa & Tudge, 2013, p. 254). Bronfenbrenner (1995) stated "The individual's own developmental life course is seen as embedded in and powerfully shaped by conditions and events occurring during the historical period through which the person lives'' (cited in Rosa & Tudge, 2013, p. 254). Bronfenbrenner in collaboration with Morris described time as having three levels: microtime, mesotime, and macrotime (Rosa & Tudge, 2013, p. 254). According to them, macrotime "focuses on the changing expectation and events in the larger society, both within and across generations" (Bronfenbrenner & Morris, 2006, cited in Rosa & Tudge, 2013, p. 254). And finally, in a continuation of PPCT's Proposition 2, Bronfenbrenner and Morris (1998) include "the social continuities and changes occurring over *time* through the life course and the historical period during which the person has lived" (cited in Tudge et al., 2009, p. 200, emphasis in original).

Particularly in the context of the UAE, the socio-cultural context seems rather important given the fundamental shift that is taking place in the country by way of the organization and running of their educational system. Traditionally, the UAE has been a nomadic, Bedouin culture. As previously noted, organized learning in 'modern' schools is a relatively recent phenomenon, for all students but particularly for girls as education for boys goes back further via religious schools tasked with

educating boys on traditional subjects such as Islamic Studies and Arabic[34]. At present, the country is still designated a developing country, albeit a high-income one (UN, 2020). As the UAE continues to move from a developing to a developed country, the school experiences of each PISA cohort may be much different than previous ones. In fact, the school experiences of this cross-section of PISA participants might be markedly different than even just 2015's assessment cohort. Even greater, generationally, students' experiences in the UAE now will be much different than their parents' if not fundamentally different to that of most of their grandparents' due to the establishment and development of the modern educational system. Therefore, consideration of time in this context seems most worthwhile, thereby making application of the theory even more clear.

Further contemplation of students' other microsystems (e.g., home) in relation to these other systems may allow classroom practice, educational policies, and cultural norms and values to be seen as all functioning in concert. Any discussion of culture will naturally segue into studying factors influencing the other system levels when using EST as the theoretical lens.

Conflict of Interest

Despite being employed previously by the UAE's MoE, there is no conflict of interest. This study was done in a fully independent capacity. Beyond completing my degree, I had no vested interest in any particular results or outcomes. Before completing my employment contract in 2018, MoE management was told that I might investigate something related to the UAE educational system as part of my impending master's program. This information was neither encouraged nor discouraged.

References

Alhebsi, A., Pettaway, L. D., & Waller, L. (2015). A History of Education in the United Arab Emirates and Trucial Sheikdoms. *The Global eLearning Journal, 4*(1), 1-6.

Archibald, S. (2006). Narrowing in on Educational Resources That Do Affect Student Achievement. *Peabody Journal of Education, 81*(4), 23-42. doi:10.1207/s15327930pje8104_2

Avvisati, F., Echazarra, A., Givord, P., & Schwabe, M. (2019). PISA 2018 Country Note United Arab Emirates. In *Programme for International Student Assessment (PISA) Results from PISA 2018*. Paris: PISA, OECD Publishing.

Bronfenbrenner, U. (1979). *The ecology of human development : experiments by nature and design*. Cambridge, Mass.: Cambridge, Mass. : Harvard Univ. Press.

Brooks-Bartlett, J. (2018). Probability concepts explained: Maximum likelihood estimation. *Towards Data Science.* Retrieved from https://towardsdatascience.com/probability-concepts-explained-maximum-likelihood-estimation-c7b4342fdbb1

Cheung, W. M., Lam, J. W. I., Au, D. W. H., So, W. W. Y., Huang, Y., & Tsang, H. W. H. (2017). Explaining student and home variance of Chinese reading achievement of the PIRLS 2011 Hong Kong. *Psychology in the Schools, 54*(9), 889-904. doi:10.1002/pits.22041

Chiu, M. M. (2018). Contextual Influences on Girls' and Boys' Motivation and Reading Achievement: Family, Schoolmates, and Country. In P. Orellana García & P. Baldwin Lind (Eds.), *Reading Achievement and Motivation in Boys and Girls: Field Studies and*

---

[34] For a history of the UAE educational system, see Alhebsi et al. (2015) or *Education in the United Arab Emirates* (2019).

Methodological Approaches (Vol. 15, pp. 49-63). Cham, Switzerland: Springer International Publishing AG, part of Springer Nature.

Chiu, M. M., & Chow, B. W.-Y. (2015). Classmate Characteristics and Student Achievement in 33 Countries: Classmates' Past Achievement, Family Socioeconomic Status, Educational Resources, and Attitudes Toward Reading. *Journal of Educational Psychology, 107*(1), 152-169. doi:10.1037/a0036897

Chiu, M. M., & Khoo, L. (2005). Effects of Resources, Inequality, and Privilege Bias on Achievement: Country, School, and Student Level Analyses. *American Educational Research Journal, 42*(4), 575-603. doi:10.3102/00028312042004575

Chudgar, A., & Luschei, T. F. (2009). National Income, Income Inequality, and the Importance of Schools: A Hierarchical Cross-National Comparison. *American Educational Research Journal, 46*(3), 626-658. doi:10.3102/0002831209340043

Chung, H. (2018). *Gender Equality and the Gender Achievement Gap: A Crossnational Study Using PISA 2015*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (ProQuest No. 13803937)

De Vaus, D. A. (2001). *Research design in social research*: London : SAGE.

Della Sala, M. R., Knoeppel, R. C., & Marion, R. (2017). Modeling the Effects of Educational Resources on Student Achievement: Implications for Resource Allocation Policies. *Education and Urban Society, 49*(2), 180-202. doi:10.1177/0013124516643757

Dickson, M. (2013). School improvements in Abu Dhabi, United Arab Emirates: Asking the 'expert witnesses'. *Improving Schools, 16*(3), 272-284. doi:10.1177/1365480213501060

Dickson, M., Fidalgo, P., & Cairns, D. (2019). The 'S' and 'T' in STEM: Integrating Science and Technology in Education in the UAE. In K. Gallagher (Ed.), *Education in the United Arab Emirates Innovation and Transformation* (pp. 95-111). Singapore: Springer Nature Singapore Pte Ltd.

Dooley, K. (2018). Ecological Systems Theory. In E. B. Braaten (Ed.), *The SAGE Encyclopedia of Intellectual and Developmental Disorders* (pp. 499-500). Thousand Oaks: Thousand Oaks: SAGE Publications.

Education in the United Arab Emirates. (2019). (K. Gallagher Ed.). Singapore: Springer Nature Singapore Pte Ltd.

Eriksson, M., Ghazinour, M., & Hammarström, A. (2018). Different uses of Bronfenbrenner's ecological theory in public mental health research: what is their value for guiding public mental health policy and practice? *Social Theory & Health, 16*(4), 414-433. doi:10.1057/s41285-018-0065-6

Fuchs, T., & Wößmann, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data. *Empirical Economics, 32*(2-3), 433-464. doi:10.1007/s00181-006-0087-0

Fuller, B. (1987). What School Factors Raise Achievement in the Third World? *Review of Educational Research, 57*(3), 255-292.

Fuller, B., & Clarke, P. (1994). Raising School Effects while Ignoring Culture? Local Conditions and the Influence of Classroom Tools, Rules, and Pedagogy. *Review of Educational Research, 64*(1), 119-157.

Gallagher, K. (2019a). Challenges and Opportunities in Sourcing, Preparing and Developing a Teaching Force for the UAE. In K. Gallagher (Ed.), *Education in the United Arab Emirates Innovation and Transformation* (pp. 127-145). Singapore: Springer Nature Singapore Pte Ltd.

Gallagher, K. (2019b). Introduction: Education in the UAE - Context and Themes. In K. Gallagher (Ed.), *Education in the United Arab Emirates Innovation and Transformation* (pp. 1-18). Singapore: Springer Nature Singapore Pte Ltd.

Gallagher, K. (2019c). Preface. In K. Gallagher (Ed.), *Education in the United Arab Emirates Innovation and Transformation* (pp. v-vi). Singapore: Springer Nature Singapore Pte Ltd.

Good Research Practice. (2017). Swedish Research Council.

Grace-Martin, K. (2020). Interpreting the intercept in a regression model. Retrieved from https://www.theanalysisfactor.com/interpreting-the-intercept-in-a-regression-model/

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). Interpreting Research on School Resources and Student Achievement: A Rejoinder to Hanushek. *Review of Educational Research, 66*(3), 411-416. doi:10.3102/00346543066003411

Gustafsson, J.-E. (2003). What do we know about effects of school resources on educational results? *Swedish economic policy review, 10*(2), 77-110.

Häkkinen, I., Kirjavainen, T., & Uusitalo, R. (2003). School resources and student achievement revisited: new evidence from panel data. *Economics of Education Review, 22*(3), 329-335. doi:10.1016/S0272-7757(02)00060-2

Hanushek, E. A. (1996). A more complete picture of school resource policies. *Review of Educational Research, 66*(3), 397-409.

Hanushek, E. A., & Luque, J. A. (2003). Efficiency and equity in schools around the world. *Economics of Education Review, 22*(5), 481-502. doi:10.1016/S0272-7757(03)00038-4

Hanushek, E. A., & Woessmann, L. (2017). School Resources and Student Achievement: A Review of Cross-Country Economic Research. In M. Rosén, K. Yang Hansen, & U. Wolff (Eds.), *Cognitive Abilities and Educational Outcomes : A Festschrift in Honour of Jan-Eric Gustafsson* (pp. 149-171). Cham, Switzerland: Springer Nature.

Hattie, J. (2009). *Visible learning : a synthesis of over 800 meta-analyses relating to achievement*: Abingdon, Oxon : Routledge.

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2011). *Multilevel and Longitudinal Modeling with IBM SPSS*: Taylor and Francis.

Heyneman, S. P. (1980). Differences between Developed and Developing Countries: Comment on Simmons and Alexander's "Determinants of School Achievement". *Economic Development and Cultural Change, 28*(2), 403-406. doi:10.1086/451180

Heyneman, S. P. (1984). Research on education in the developing countries. *International Journal of Educational Development, 4*(4), 293-304.

Heyneman, S. P., Jamison, D. T., & Montenegro, X. (1984). Textbooks in the Philippines: Evaluation of the Pedagogical Impact of a Nationwide Investment. *Educational Evaluation and Policy Analysis, 6*(2), 139-150. doi:10.3102/01623737006002139

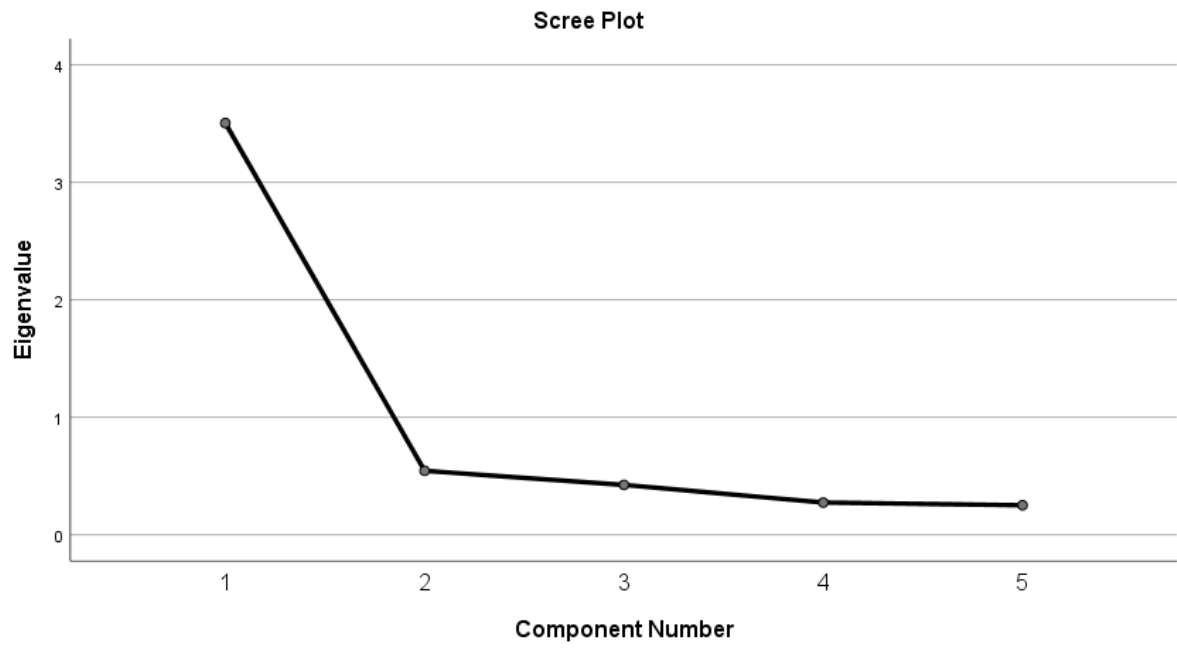IBM. (2019). IBM SPSS Statistics for Windows (Version 26.0). Armonk, NY: IBM Corp.

ICT. (2020). In *Cambridge Dictionary*. Retrieved from
https://dictionary.cambridge.org/dictionary/english/ict?q=ICT

IEA. About us. Retrieved from https://www.iea.nl/about

Jürges, H., & Schneider, K. (2004). International Differences in Student Achievement: An Economic Perspective. *German Economic Review, 5*(3), 357-380.

Kippels, S., & Ridge, N. (2019). The Growth and Transformation of K–12 Education in the UAE. In K. Gallagher (Ed.), *Education in the United Arab Emirates Innovation and Transformation* (pp. 37-55). Singapore: Springer Nature Singapore Pte Ltd.

Legewie, J., & Diprete, T. A. (2012). School Context and the Gender Gap in Educational Achievement. *American Sociological Review, 77*(3), 463-485. doi:10.1177/0003122412440802

Lindblad, S., Pettersson, D., & Popkewitz, T. S. (2018). Getting the Numbers Right: An Introduction. In S. Lindblad, D. Pettersson, & T. S. Popkewitz (Eds.), *Education by the Numbers and the Making of Society: The Expertise of International Assessments*. New York, NY: Taylor and Francis.

Ma, X., & Crocker, R. (2007). Provincial Effects on Reading Achievement. *Alberta Journal of Educational Research, 53*(1), 87-109.

Marks, G. N. (2008). Accounting for the gender gaps in student performance in reading and mathematics: evidence from 31 countries. *Oxford Review of Education, 34*(1), 89-109. doi:10.1080/03054980701565279

Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Educational Research and Evaluation, 12*(2), 105-128. doi:10.1080/13803610600587040

Muijs, D. (2011). *Doing quantitative research in education with SPSS* (2nd ed.). London: Sage Publications.

Murnane, R. J. (1995). Supply of Teachers. In M. Carnoy (Ed.), *International Encyclopedia of Economics of Education* (2nd ed., pp. 317-321). Oxford: Oxford : Pergamon.

Ning, B., Van Damme, J., Gielen, S., Vanlaar, G., & Van Den Noortgate, W. (2016). What Makes the Difference in Reading Achievement? Comparisons Between Finland and Shanghai. *Scandinavian Journal of Educational Research, 60*(5), 515-537. doi:10.1080/00313831.2015.1062413

Odden, A., Borman, G., & Fermanich, M. (2004). Assessing Teacher, Classroom, and School Effects, Including Fiscal Effects. *PJE. Peabody Journal of Education, 79*(4), 4-32. doi:10.1207/s15327930pje7904_2

OECD. (2009). PISA Data Analysis Manual SPSS® (2nd ed.). Retrieved from
https://www.oecd.org/pisa/data/

OECD. (2017). PISA 2015 Technical Report. Retrieved from https://www.oecd.org/pisa/publications/

OECD. (2018). PISA FAQ. Retrieved from https://www.oecd.org/pisa/pisafaq/

OECD. (2019a). PISA 2018 Assessment and Analytical Framework. doi:https://doi.org/10.1787/b25efab8-en

OECD. (2019b). PISA 2018 Results (Volume II): Where All Students Can Succeed. In. Paris: PISA, OECD Publishing.

OECD. (2019c). PISA 2018 Results (Volume III): What School Life Means for Students' Lives. In. Paris: PISA, OECD Publishing.

OECD. (2019d). Who we are. Retrieved from https://www.oecd.org/about/

Onwuegbuzie, A. J., Collins, K. M. T., & Frels, R. K. (2013). Using Bronfenbrenner's ecological systems theory to frame quantitative, qualitative, and mixed research. *International Journal of Multiple Research Approaches, 7*(1), 2-8. doi:10.5172/mra.2013.7.1.2

Pallant, J. (2016). *SPSS survival manual : a step by step guide to data analysis using IBM SPSS* (6th ed.). Maidenhead: Open University Press : McGraw-Hill.

Reading Literacy. (n.d.). Retrieved from https://www.pisa.tum.de/en/domains/reading-literacy/

Reilly, D. (2012). Gender, culture, and sex-typed cognitive abilities. *PLoS ONE, 7*(7), 1-16. doi:10.1371/journal.pone.0039904

Ridge, N. (2009). *Privileged and penalized: The education of boys in the United Arab Emirates.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3348361)

Rosa, E. M., & Tudge, J. (2013). Urie Bronfenbrenner's Theory of Human Development: Its Evolution From Ecology to Bioecology. *Journal of Family Theory and Review, 5*(4), 243.

Russell, C. (2012). *Gender, Academic Achievement, and Meanings of Schooling in Ras al Khaimah, United Arab Emirates.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3506367)

S.1 Basic Terminology. (2020). Retrieved from https://online.stat.psu.edu/statprogram/reviews/statistical-concepts/terminology

Seltman, H. J. (2018). *Experimental Design and Analysis*. Retrieved from http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

Simmons, J., & Alexander, L. (1978). The Determinants of School Achievement in Developing Countries: A Review of the Research. *Economic Development and Cultural Change, 26*(2), 341-357.

Skryabin, M., Zhang, J., Liu, L., & Zhang, D. (2015). How the ICT development level and usage influence student achievement in reading, mathematics, and science. *Computers & Education, 85*, 49-58. doi:10.1016/j.compedu.2015.02.004

Statistical Computing Workshop: Using the SPSS Mixed Command. (2020). Retrieved from https://stats.idre.ucla.edu/spss/seminars/spss-mixed-command/

Stoet, G., & Geary, D. C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence, 48*, 137-151. doi:10.1016/j.intell.2014.11.006

Tamim, R. M., & Colburn, L. K. (2019). In Quest of Educational Quality in the UAE. In K. Gallagher (Ed.), *Education in the United Arab Emirates Innovation and Transformation* (pp. 161-179). Singapore: Springer Nature Singapore Pte Ltd.

Taylor Gobert, M. (2019). Transformation in English Language Education in the UAE. In K. Gallagher (Ed.), *Education in the United Arab Emirates Innovation and Transformation* (pp. 113-126). Singapore: Springer Nature Singapore Pte Ltd.

Tsai, S.-L., Smith, M., & Hauser, R. (2018). Gender Gaps in Student Academic Achievement and Inequality. In H. Park & G. Kao (Eds.), *Research in the Sociology of Education* (Vol. 20, pp. 181-218). UK: Emerald Publishing Limited.

Tudge, J. R. H., Mokrova, I., Hatfield, B. E., & Karnik, R. B. (2009). Uses and Misuses of Bronfenbrenner's Bioecological Theory of Human Development. *Journal of Family Theory & Review, 1*(4), 198-210. doi:10.1111/j.1756-2589.2009.00026.x

UAE. (2018). Vision 2021 National Agenda. Retrieved from https://www.vision2021.ae/en/national-agenda-2021/list/first-rate-circle

The UAE Reforms Education System as Part of Vision 2021. (2017). Retrieved from https://fanack.com/education/uae-reforms-education-system/

UN. (2020). *World Economic Situation and Prospects*. New York: United Nations.

United Arab Emirates Population (Live). (n.d.). Retrieved from https://www.worldometers.info/world-population/united-arab-emirates-population/

Van Hek, M., Kraaykamp, G., & Pelzer, B. (2018). Do schools affect girls' and boys' reading performance differently? A multilevel study on the gendered effects of school resources and school practices. *School Effectiveness and School Improvement, 29*(1), 1-21. doi:10.1080/09243453.2017.1382540

Warner, R. S., & Burton, G. J. S. (2017). A Fertile Oasis: The Current State of Education in the UAE. Retrieved from https://www.mbrsg.ae/home/publications/research-report-research-paper-white-paper/a-fertile-oasis-the-current-state-of-education-in.aspx

Westley, J. (2017). Using PISA to Benchmark UAE Schools - Why, & Why Not ... Retrieved from https://schoolscompared.com/guides/using-pisa-benchmark-uae-schools-not/

Woessmann, L. (2016). The Importance of School Systems: Evidence from International Differences in Student Achievement †. *Journal of Economic Perspectives, 30*(3), 3-32. doi:10.1257/jep.30.3.3

Xie, C., & Ma, Y. (2019). The mediating role of cultural capital in the relationship between socioeconomic status and student achievement in 14 economies. *British Educational Research Journal, 45*(4), 838-855. doi:10.1002/berj.3528

Z-Score: Definition, Formula and Calculation. (2020). Retrieved from https://www.statisticshowto.com/probability-and-statistics/z-score/

Appendix A

Name/label	PISA variable	Description/question
Table A1: List of Variables		
S_GENDER	ST004Q1D01T (renamed)	Student gender 1 = Female, 2 = Male
S_gmSES	ESCS (renamed S_SES)	Grand-mean-centered student SES variable Student index of economic, social and cultural status (socioeconomic status)
S_READACH	PV1READ (renamed)	Overall student reading achievement plausible value 1
SCH_DDEVICES_round	SC155Q01–Q05HA	Rounded digital devices index = school index of five of 11 SC155 questions (school materials resources – digital devices) 1 = Strongly disagree 2 = Disagree 3 = Agree 4 = Strongly agree
SCH_EMIRATE	STRATUM	Location of school – recoded STRATUM variable 1 = Abu Dhabi, 2 = Dubai, 3 = Sharjah, 4 = Ajman, 5 = Umm Al Quwain, 6 = Ras Al Khaimah, 7 = Fujairah
SCH_gmDISADVANTAGED	SC048Q03NA	Grand-mean-centered SCH_DISADVANTAGED School percentage of students from socioeconomically disadvantaged homes
SCH_gmTQUAL	PROATCE (renamed SCH_TQUAL)	Grand-mean-centered SCH_TQUAL School index proportion of all teachers fully certified (teacher quality)
SCH_MATERIALS_round	SC017Q05–Q06NA	Rounded school materials index composite index of two questions (school material resources) *reverse coded values listed 1 = Not at all 2 = Very little 3 = To some extent 4 = A lot
SCH_TYPE	SC002Q01–Q02TA	Total school enrollment (number of students) (Q01 = boys/Q02 = girls) 1 = boys' school, 2 = girls' school, 3 = coed school
SCH_URBANIZATION	SC001Q01TA	Location of school – recoded SC001 variable 1 = A town or smaller (up to about 100,000 people) 2 = A city (100,000 to about 1,000,000 people) 3 = A large city (with over 1,000,000 people)
W_FSTUWT	W_FSTUWT	Final trimmed nonresponse adjusted student weight (full student sampling weight variable to be used when analyzing student-level data)

Figure A1: Screeplot for PCA (IBM, 2019)

| | | | School Materials | | | | |
			1.00	2.00	3.00	4.00	Total
School Emirate	Abu Dhabi	Count	1656	2334	4527	10300	18817
		% within School Emirate	8.8%	12.4%	24.1%	54.7%	100.0%
		% within School Materials	41.1%	45.5%	43.6%	32.7%	36.8%
	Dubai	Count	300	1162	1870	10779	14111
		% within School Emirate	2.1%	8.2%	13.3%	76.4%	100.0%
		% within School Materials	7.4%	22.7%	18.0%	34.2%	27.6%
	Sharjah	Count	1434	852	2256	5691	10233
		% within School Emirate	14.0%	8.3%	22.0%	55.6%	100.0%
		% within School Materials	35.6%	16.6%	21.7%	18.0%	20.0%
	Ajman	Count	295	103	399	1576	2373
		% within School Emirate	12.4%	4.3%	16.8%	66.4%	100.0%
		% within School Materials	7.3%	2.0%	3.8%	5.0%	4.6%
	Umm Al Quwain	Count	0	27	181	301	509
		% within School Emirate	0.0%	5.3%	35.6%	59.1%	100.0%
		% within School Materials	0.0%	0.5%	1.7%	1.0%	1.0%
	Ras Al Khaimah	Count	263	454	586	1772	3075
		% within School Emirate	8.6%	14.8%	19.1%	57.6%	100.0%
		% within School Materials	6.5%	8.9%	5.6%	5.6%	6.0%
	Fujairah	Count	83	194	558	1112	1947
		% within School Emirate	4.3%	10.0%	28.7%	57.1%	100.0%
		% within School Materials	2.1%	3.8%	5.4%	3.5%	3.8%
Total		Count	4031	5126	10377	31531	51065
		% within School Emirate	7.9%	10.0%	20.3%	61.7%	100.0%
		% within School Materials	100.0%	100.0%	100.0%	100.0%	100.0%
School Urbanization (SU)	A town or smaller	Count	1924	2487	3949	9169	17529
		% within SU	11.0%	14.2%	22.5%	52.3%	100%
		% with School Materials	47.7%	48.5%	38.4%	29.1%	34.4%
	A city	Count	1734	1281	2789	9484	15288
		% within SU	11.3%	8.4%	18.2%	62.0%	100.0%
		% within School Materials	43.0%	25.0%	27.1%	30.1%	30.0%
	A large city	Count	374	1357	3544	12880	18155
		% within SU	2.1%	7.5%	19.5%	70.9%	100.0%
		% within School Materials	9.3%	26.5%	34.5%	40.8%	35.6%
Total		Count	4032	5125	10282	31533	50972
		% within SU	7.9%	10.1%	20.2%	61.9%	100.0%
		% within School Materials	100.0%	100.0%	100.0%	100.0%	100.0%
School Type	Boys'	Count	630	897	1468	3293	6288
		% within School Type	10.0%	14.3%	23.3%	52.4%	100.0%
		% within School Materials	32.9%	24.6%	18.4%	12.3%	15.6%
	Girls'	Count	507	775	1178	2565	5025
		% within School Type	10.1%	15.4%	23.4%	51.0%	100.0%
		% within School Materials	26.5%	21.3%	14.7%	9.6%	12.5%
	Co-ed	Count	775	1973	5346	20820	28914
		% within School Type	2.7%	6.8%	18.5%	72.0%	100.0%
		% within School Materials	40.5%	54.1%	66.9%	78.0%	71.9%
Total		Count	1912	3645	7992	26678	40227
		% within School Type	4.8%	9.1%	19.9%	66.3%	100.0%
		% within School Materials	100.0%	100.0%	100.0%	100.0%	100.0%

Table A2:
Crosstabulations for School Materials Ordinal Regression

			School Digital Devices (SDD)				
			1.00	2.00	3.00	4.00	Total
School Emirate	Abu Dhabi	Count	748	2436	12917	4529	20630
		% within School Emirate	3.6%	11.8%	62.6%	22.0%	100.0%
		% within SDD	95.8%	42.2%	41.5%	27.1%	37.9%
	Dubai	Count	0	1039	6957	6585	14581
		% within School Emirate	0.0%	7.1%	47.7%	45.2%	100.0%
		% within SDD	0.0%	18.0%	22.4%	39.3%	26.8%
	Sharjah	Count	0	1651	5832	2749	10232
		% within School Emirate	0.0%	16.1%	57.0%	26.9%	100.0%
		% within SDD	0.0%	28.6%	18.8%	16.4%	18.8%
	Ajman	Count	33	173	1601	1144	2951
		% within School Emirate	1.1%	5.9%	54.3%	38.8%	100.0%
		% within SDD	4.2%	3.0%	5.1%	6.8%	5.4%
	Umm Al Quwain	Count	0	67	279	230	576
		% within School Emirate	0.0%	11.6%	48.4%	39.9%	100.0%
		% within SDD	0.0%	1.2%	0.9%	1.4%	1.1%
	Ras Al Khaimah	Count	0	312	2096	953	3361
		% within School Emirate	0.0%	9.3%	62.4%	28.4%	100.0%
		% within SDD	0.0%	5.4%	6.7%	5.7%	6.2%
	Fujairah	Count	0	97	1421	553	2071
		% within School Emirate	0.0%	4.7%	68.6%	26.7%	100.0%
		% within SDD	0.0%	1.7%	4.6%	3.3%	3.8%
Total		Count	781	5775	31103	16743	54402
		% within School Emirate	1.4%	10.6%	57.2%	30.8%	100.0%
		% within SDD	100.0%	100.0%	100.0%	100.0%	100.0%
School Urbanization (SU)	A town or smaller	Count	348	2246	9890	5070	17554
		% within SU	2.0%	12.8%	56.3%	28.9%	100%
		% within SDD	44.5%	38.9%	35.4%	30.3%	34.3%
	A city	Count	154	1181	9229	4976	15540
		% within SU	1.0%	7.6%	59.4%	32.0%	100.0%
		% within SDD	19.7%	20.5%	33.0%	29.7%	30.3%
	A large city	Count	280	2348	8830	6698	18156
		% within SU	1.5%	12.9%	48.6%	36.9%	100.0%
		% within SDD	35.8%	40.7%	31.6%	40.0%	35.4%
Total		Count	782	5775	27949	16744	51250
		% within SU	1.5%	11.3%	54.5%	32.7%	100.0%
		% within SDD	100.0%	100.0%	100.0%	100.0%	100.0%
School Type	Boys'	Count	20	986	3882	1399	6287
		% within School Type	0.3%	15.7%	61.7%	22.3%	100.0%
		% within SDD	4.8%	19.0%	18.8%	9.9%	15.6%
	Girls'	Count	224	746	2311	1745	5026
		% within School Type	4.5%	14.8%	46.0%	34.7%	100.0%
		% within SDD	53.3%	14.4%	11.2%	12.4%	12.5%
	Co-ed	Count	176	3445	14405	10928	28954
		% within School Type	0.6%	11.9%	49.8%	37.7%	100.0%
		% within SDD	41.9%	66.5%	69.9%	77.7%	71.9%
Total		Count	420	5177	20598	14072	40267
		% within School Type	1.0%	12.9%	51.2%	34.9%	100.0%
		% within SDD	100.0%	100.0%	100.0%	100.0%	100.0%

Table A3:
Crosstabulations for School Digital Devices Ordinal Regression

Appendix B

Calculations

T-test eta squared:

$60.543^2 / 60.543^2 + (27572 + 26831 – 2)$

$3665.455 / 3665.455 + 54401$

$3665.455 / 58066.455 = .063$ (6.3%)

T-test Cohen's *d*:

$(402.132 - 459.192) / 109.785885$

$-57.06 / 109.785885 = 0.519739$

Null Model ICC:

$6390.026 / 6390.026 + 6830.188$

$6390.026 / 13220.214 = .48335$ (48.34%)

Level 1 Model change in residual:

$6830.188 – 6621.637 / 6830.188$

$208.551 / 6830.188 = .0305$ (3.05%)

Level 1 Model reduction in variance estimate:

$6390.026 – 5044.097 / 6390.026$

$1345.929 / 6390.026 = .2106$ (21.06%)

Level 1 Model ICC:

$5044.097 / 5044.097 + 6621.637$

$5044.097 / 11665.734 = .4324$ (43.24%)

Level 2 Model reduction in variance estimate:

$6390.026 – 2261.894 / 6390.026$

$4128.132 / 6390.026 = 64.6\%$