A Multilevel Analysis of Educational Inputs and Academic Achievement Among Primary Students in Nigeria

Amal Al Gergawi

Kellogg College

University of Oxford

Supervised by Dr. Ellie Ott

Submitted: Michaelmas 2020

September 11th, 2020

This dissertation is submitted to the Department of Education, University of Oxford, in partial fulfillment of the requirements for the degree of *Master of Science in Comparative and International Education*

Acknowledgements

First and foremost, I would like to express my deep gratitude to my supervisor, Dr. Ellie Ott, for her constructive feedback, persistent guidance, and empathetic support during the planning and development of this research. I further would like to extend my gratitude to my professor and mentor, Dr. David Johnson, for generously sharing his unpublished data and imparting his wisdom on instrument design. Access to this data would not have been feasible without the contributions of the World Bank, the Nigerian Federal Ministry of Education, as well as the wonderful enumerators, school leaders, teachers, and students involved in the initial data collection process. I would like to particularly acknowledge with a profound sense of reverence World Bank Data Analyst, Kabira Namit, for his statistical code checks and meticulous knowledge on the learning assessments. I would like to additionally pay special regards to postdoctoral research officer, Dr. Ellie Suh, for her astute recommendations on multilevel model design and to my dear friends Alyssa Yu, Hannah Melville-Rea, Laila Ujayli, Marie-Claude Hykpo, Sarah Al-Kaabi and Theros Wong for their discerning comments and unwavering support. To my fellow colleagues in the MSc in Comparative and International Education, though our time was cut short, your diverse and critical perspectives has left permanent traces on my dissertation and undoubtedly on my academic prowess. To my beloved family, thank you for your patience and affection this past year. As always, I am ineffably indebted to you.

Abstract

Global efforts to increase primary enrollment in the past few decades have fallen short of improving academic achievement within low- and middle-income countries: a conundrum the United Nations Educational, Scientific and Cultural Organization has referred to as the "global learning crisis." Nigeria has been no exception. In the wake of declining educational funds and stagnating student performance in the West-African nation, this study empirically examines the association between six educational inputs (i.e. good classroom conditions, functional desk conditions, presence of toilet/s, textbook availability, exercise book availability, and the presence of corner library/ies) and overall academic achievement (i.e. aggregation of English, Mathematics, and Science assessment scores) among fourth grade students. Using secondary data from the World Bank collected in 2019, this study pursues a multilevel model approach (two-level ordinal logistic regression) to analyze the input-achievement nexus with a sample of 3,262 students nested across 97 randomly selected primary schools. The study revealed mixed findings, but an overall inclination towards a positive association. With a 95% confidence interval, this study found that good classroom conditions, functional desk conditions, and textbook availability positively predicted higher academic achievement, but corner library/ies negatively predicted higher academic achievement. Toilet/s and exercise book availability yielded statistically non-significant findings. For further analytical nuance, the findings indicated that student biological sex did not statistically significantly predict higher academic achievement nor did educational inputs moderate this association. Ultimately, this study sheds light on how school leaders, policy makers, and multinational organizations can strategically address resource distribution for primary education in Nigeria; thereby, expanding the frontiers of educational literature within under researched contexts in the "Global South."

Table of Contents

ABSTRACT 2
LIST OF ABBREVIATIONS 5
1. INTRODUCTION 6
2. BACKGROUND 10
3. LITERATURE REVIEW 17
4. RESEARCH DESIGN AND METHODOLOGY 27
5. HYPOTHESES 37
6. RESULTS 44
7. DISCUSSION & LIMITATIONS 53
8. CONCLUSION 61
REFERENCES 63
APPENDIX A: DESCRIPTIVE ANALYSIS 81
APPENDIX B: CRONBACH'S ALPHA OUTPUTS 86
APPENDIX C: MULTILEVEL METHOD SELECTION 88
APPENDIX D: HIERARCHICAL LINEAR MODEL VIOLATION 90
APPENDIX E: GENERALIZED ORDINAL LOGISTIC REGRESSION OUTPUTS 95
APPENDIX F: ASSUMPTION TESTING FOR TWO-LEVEL ORDINAL LOGISTIC REGRESSIONS 98
APPENDIX G: MATHEMATICAL NOTATIONS OF TWO-LEVEL MODELS 106
APPENDIX H: TWO-LEVEL ORDERED LOGISTIC MODELS (1-4) 109
APPENDIX I: TWO-LEVEL ORDERED LOGISTIC MODELS (3.1-3.6) 120
APPENDIX J: INTRACLASS CORRELATION COEFFICIENT CALCULATIONS 128
APPENDIX K: MISSING VALUES 130
APPENDIX L: ENGLISH, MATHEMATICS, AND SCIENCE ASSESSMENTS 131
APPENDIX M: TEACHER INTERVIEW & SCHOOL SURVEY (EXTRACTS) 160

List of Abbreviations

AIC	Akaike's Information Criterion
BIC	Bayesian Information Criterion
BESDA	Better Education Service Delivery for All
DFID	Development for International Development
ESSPIN	Education Sector Support Programme in Nigeria Survey
GDP	Gross Domestic Product
GOLM	Generalized Ordinal Logistic Model
HLM	Hierarchical Linear Modeling
ICC	Intraclass Correlation Coefficient
KDE	Kernel Density Estimate
LR	Likelihood Ratio
MLA	Monitor Learning Assessment
MLE	Maximum Likelihood Estimation
MOLM	Multilevel Ordinal Logistic Model
OR	Odds Ratio
P-P Plot	Probability-Probability Plot
PRILS	Progress in International Reading Literacy Study
PTA	Parent-Teacher Association
Q-Q Plot	Quantile-Quantile Plot
SDG 4	Sustainable Development Goal 4
SUBEB	State Universal Basic Education Board
TIMSS	Trend in International Mathematics and Science Study
TOLM	Two-Level Ordinal Logistic Model
UBE	Universal Basic Education
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations International Children's Emergency Fund

1. Introduction

To the dismay of governments, multinational organizations, and educators, the global increase in access to primary education over the course of a few decades has ultimately failed to improve the academic achievement of students in low- and middle-income countries. The United Nations Educational, Scientific and Cultural Organization (UNESCO) refers to this conundrum as a "global learning crisis" (UNESCO, 2013). Though primary school enrollment has reached near universal rates at 90% as of 2018 marking a 19% increase from 1970, the stagnation of achievement, a concept referring to student performance on a given assessment, has signaled a severe disconnect between the perceived value of formal schooling and the poor quality of student learning (UNESCO, 2018; Gove & Cvelich, 2011; Snilstveit et al., 2017).

This learning crisis has been particularly acute in Nigeria: primary achievement scores have remained abysmally low for more than two decades despite the persistent efforts of the federal government and international development organizations to improve the education system (Cameron, et al 2016). Various national learning assessments conducted between 1996 and 2016, such as the Nigeria Population Comission's Nigeria Education Data Survey in 2015 and the UK Department of International Development's Education Sector Support Programme in Nigeria Survey in 2016, attest to the unfortunate persistence of poor performance in literacy, numeracy, science, and other life-skills with signs of a slight decline over time (Falayajo et al, 1997; National Population Commission, 2015). These learning assessments, while insightful, have not been frequently used to empirically investigate what factors best predict academic achievement, warranting a closer read of the literature for a deeper understanding of this perennial problem.[1]

A burgeoning body of literature contends that academic achievement can be predicted by an array of interrelated factors emanating from *students* (e.g. gender, biological predisposition, personality, prior learning attainment, motivation, health, aspirational goals), *households* (e.g. parents' socio-economic status, family composition, maternal employment, technological devices, internet connectivity), *schools* (e.g. educational inputs, student-teacher ratio), *teachers* (e.g. gender, pedagogical strategy, length of teacher experiences, incentives), *institutional landscapes*

[1] In line with conventional terminology used in multilevel literature, the term "predict" in this study will refer to the probability of a student performing within a higher academic achievement category. It will not be used to insinuate a time-lag in the investigation of the inputs-achievement nexus (O'Connell, 2010).

(e.g. curriculum reform, corruption, political structure), and *international organizations* (e.g. OECD, UNESCO). These diverse domains of achievement predictors echo Bronfenbrenner's (1979) ecological systems model of human development, which suggest that child development, and by extension the academic success of students, can be subject to the influence of countless factors. Among those mentioned, scholars argue that educational inputs –available resources at the school-level– can serve as a critical catalyst for learning when faced with limited resources (Callaway & Musone, 1968). Lee & Zuze (2011) assert this notion in their statement: "the availability of resources to support learning is the essence of the educational quality debate in developing countries. Existing human and physical resources in a school –buildings, teachers, books, and desks– are typically allocated on the basis of national priorities" (p. 370).

In the context of Nigeria, the investigation of the input-achievement nexus appears to be of paramount importance as scarce resources are stretched too thin. In fact, the proportion of government budget allocated to the education sector over the past two decades has remained well below the 15 to 20 percent minimum for low-income nations as recommended by UNESCO, often oscillating within the 5-8 percent (Ameh & Aluko, 2019). Compounding onto this small budget, the federal government recently slashed funds towards Universal Basic Education (UBE), a program launched in 1990 to support the provision of free, universal, and compulsory education for children aged 6-15 years, by 54% (Salau et al., 2020). The decrease in budgets towards the UBE amidst the global coronavirus pandemic reduces the already limited resources available in primary schools (Abu et al., 2020). In addition to the education budget, the discretionary and often ad hoc approach to the distribution of scarce educational funds disregards the potential implications of such government policies on student performance.

A historical review of empirical literature further gestures to the need for a deeper dissection of this nexus in Nigeria. Over the course of a century, researchers grappling with issues of educational quality sought to examine the association between inputs and achievement (Norton, 1926). Pioneering studies in the 1930s and 1940s seemed to suggest a positive association between school resources and student outcomes until this consensus was challenged in 1966 upon the release of the Equal Educational Opportunity Study. Commonly referred to as the Coleman Report, the findings of the study indicated that educational inputs did not explain a substantial amount of the variance in student achievement between individual and schools when compared with the estimates of parental socioeconomic status, race, and social networks (Coleman et al., 1966). That

is, educational inputs were weak predictors of academic achievement, if at all. The Coleman Report opened Pandora's box across academic and policy circles chronicled by the plethora of debates, conferences, and research that emerged (Glewwe et al., 2011; Sabarwal et al., 2014).

However, much of the literature accumulated has centered on educational landscapes situated in the "Global North" (i.e. high-income countries) with limited research conducted in the "Global South" (i.e. middle- and low-income countries). Disregarding this literature gap would severely skew the narrative on the input-achievement nexus, as it has become apparent that the pattern of data tends to more consistently yield mixed findings in the "Global North," but positive findings in the "Global South" (Heyneman & Loxley, 1983; Walberg, 1991). This data pattern cannot be mapped onto Nigeria due to the absence of nationally representative research examining this association while internalizing the nested structure of schools (i.e. students nested in schools). The paucity of literature relating to Nigeria specifically signals the imperative need to investigate how educational inputs predict the academic achievement among primary students suffering from the dual effects of a learning crisis and thinning resources. In addition to insufficient scholarly research, much of the limited literature has explored how educational inputs moderate the association between student biological sex and achievement at the secondary level, but not at the primary level (Erinosho, 1999; Jegede, 1994).

As a result of the combined conceptual and geographical dearth of literature, this proposed study aims to investigate whether six educational inputs split into two main clusters, physical facilities (bad classroom condition, broken desk condition, presence of toilet/s) and educational materials (textbook availability, exercise book availability, and presence of corner library/ies) predict the academic achievement of fourth grade students in English, Mathematics, and Science across primary schools in Nigeria. The decision to place certain educational inputs under these two theoretical clusters not only adheres to customary practice, but also positions this study to concentrate on two important layers under this broad association (Jencks, 1972; Onolemhenmhen, 2014). In any case, literature related to these two input clusters mirror the aforementioned difference in data pattern between the "Global North" and "Global South." For a robust analysis of this association, this study explores cross-level interactions to evaluate whether these educational inputs moderate the relationship between sex and achievement.

In tandem with conceptual and geographical gaps in the literature, there appears to be a methodological gap. Studies have failed to account for the nested school structures and quantitative

epistemologies, which has perpetuated misleading and atomistic findings. Research interested in unpacking the learning crisis in Nigeria traditionally employed anecdotal and qualitative methods, with the remaining few quantitative studies interested in the effects of parental characteristics, parental involvement, gender, regional location, and socio-economic background (Onwuameze, 2013; Nevo & Egenti. 2019; Adetunji et al., 2019). To address this methodological gap, this proposed study makes an empirical contribution to the academic discourse by pursuing multilevel modelling –a quantitative method that can internalize nested school structures. Through a nationally representative data set collected in 2019 under the World Bank's Better Education Service Delivery for All (BESDA) development project, this study particularly employs a two-level ordinal logistic regression model.

In light of the non-experimental nature of this research, this study acknowledges that it will not ascertain a causal or a linear relationship due to limitations in the data set, but rather gauge for the presence of a correlation between the two constructs. While an experimental design would be an ideal way to account for spurious variables (i.e. variables that predict both educational inputs and academic achievement) such as socioeconomic status, a non-experimental multilevel model still has immense merit in producing results that more holistically depict educational performance in Nigeria. As opposed to a diagnostic tool to solve the "learning crisis," this study begins to scratch the surface of the problem to, ultimately, garner a greater appreciation of the association between educational inputs and academic achievement among fourth grade primary students as the West-African nation charts its course for a more prosperous future for its children in the coming decades.

This paper is organized into seven sections. Following this introduction, Section 2 will provide a contextualized overview of educational landscape in Nigeria. Section 3 will critically review relevant literature investigating the input-achievement nexus. Section 4 will introduce the research design, methodology, and hypotheses as derived from the literature. Section 5 will report the results of the multilevel logit models. Section 6 discusses the findings, validity, and limitations of the results. This study will conclude in Section 7.

2. Background

For an enriched engagement with the academic discourse, it is imperative that this study presents a brief background of Nigeria. This section will first illustrate how Nigeria has been afflicted with a set of interrelated social, economic, and political problems, which have compounded onto the education system. While this study will not empirically assess the impact of these problems on education, the purpose of mentioning them in this section are twofold: 1) To recognize the complex interplay of structures involved in any investigation of achievement predictors within post-colonial societies;, and 2) To reassert the pressing need for a multilevel analysis on the input-achievement nexus in Nigeria. With this national context in mind, this section will, secondly, unpack the educational landscape of primary schooling in the country. Specific emphasis will be given to the demonstration of stagnant academic achievement and deficient educational inputs. within the West-African nation-state.

Education in a Post-Colonial Nigeria

The Federal Republic of Nigeria –a sovereign state bordering Niger, Cameroon, and Guinea– has been plagued with social, economic, and political instability, which, according to scholars, has led to the production of a dysfunctional educational system as reflected in the stagnation of primary achievement.

Since the declaration of independence from British colonial rule nearly 60 years ago, social instability in the form of violent conflict has adversely affected education in Nigeria. The multi-ethnic society, composed of more than 250 ethnic groups and 500 languages, suffers from religious and socio-economic stratifications that have been mapped onto a geopolitical divide between the impoverished Muslim North and affluent Christian South. This polarization has led to the frequent eruption of violent conflict in the region, most notably the Nigerian Civil War (1967-1970), the Ogoni riot (1995), the Niger Delta crisis (2003-Present), and the Boko Haram insurgency (2009-Present) (Adewale, 2011). These ethno-religious conflicts can be traced to the consolidation of the Northern Nigerian Protectorate and Southern Nigerian Protectorate in 1914 by the British imperialist Frederick Lugard (Princewill et al., 2020). The impact of colonialism, and the ensuing social instability, has disrupted learning. For instance, in response to Boko Haram militant attacks on schools, Onyemelukwe-Waziri (2017) states “even if classes resume after an attack, it is hard

to revert to normal as the quality of education can suffer when students and teachers are afraid to go back" (p. 18). With studies showing that the psychological trauma of school attacks has decreased achievement, it has become apparent that social instability has contributed to the learning crisis in Nigeria (Alonge & Okunade, 2014). Such instability has also been driven by colonial legacies on education. For example, the narrative behind Boko Haram has been to dismantle English and Westernized education in Nigeria, which they believe has peripheralized local languages from learning as well as has rendered indigenous and religious educational systems the non-normative mode of teaching and learning (Bertoni et al., 2019). While it may be unclear whether peace can be actualized in the near future, it is, nevertheless, important to acknowledge that social instability, often a legacy of colonialism, has negatively affected education in Nigeria.

In addition to social instability, economic instability in Nigeria has adversely affected the federal education budget. Though considered the second-largest economy in the African continent largely due to an over-reliance on hydrocarbons, the Nigerian economy has been left in a precarious position in the wake of a drop in oil and gas prices triggered by the global pandemic (Nordea, 2020). This instability has resulted in the recent slash of the primary education budget by 54% (Salau et al., 2020). The uncertainty of when the virus infections may stop domestically and globally not only clouds the economic outlook of the nation, but also risks reversing previous efforts made to enhance learning in the country. But, even prior to this pandemic, a stable economic growth rate of 5% or more in the past decade has seldom translated into adequate education funding. The proportion of GDP allocated to the education sector has remained low, shifting ever so slightly within the range of 5-8%. Additionally, corruption further limits resources, with officials stealing up to 66% of the federal education budget in 2019 (Adegboyega, 2019). In a similar fashion to conflict, it may take years before the economy fully diversifies and eliminates corruption, thus, a focus on educational inputs can be a concrete, albeit small, means to begin to address the learning crisis.

Adding to social and economic instability, political instability has adversely affected the Nigerian education system. Setting aside instability caused by thirty years of military rule, 7 attempted coup d'états, and frequent violent conflict, the central government has been highly politicized and ethnically charged. Tensions in the National Assembly and the Ministry of Education have hindered attempts to effectively and inclusively implement a coherent set of educational policies. The failure of the program such as the Universal Primary Education at the

regional-scale in the 1950s and 1960s, the Universal Primary Education at the national-scale in 1976 further revised in 1981 and 1990, the National Primary Education Commission's cyclical disagreements over decrees between 1988 and 1993, and the Compulsory, Free Universal Basic Education in 2004 all attest to this perception (UBEC, 2020). Given that political reforms may take years, this study chooses to focus on educational inputs due to its existing involvement in educational policy planning and its proximity to formal learning spaces, which can discourage politicized methods of policy implementation. Ultimately, in a country besieged by social, economic and political instability, the list of factors contributing to the learning crisis is long. But, the need to approach this problem from the dimension of educational inputs remains ever more salient in light of increasing fiscal scarcity. For a contextualized investigation of the input-achievement nexus, this section introduces the educational scene in Nigeria.

Unpacking the Educational Landscape in Nigeria

Expressed in the 4th edition of the National Policy on Education, Nigeria views education as an instrument "par excellence" for the enhancement of national development; however, the willingness of the federal government to invest in the human capital of its children has failed to translate into better student achievement over the years (National Policy on Education, 2004, p. 4). The Africa Learning Barometer, an interactive comparative dataset on learning assessments, found that Nigeria was one of three countries where more than half of in-school students were not learning basic skills upon graduating from primary school (Center for Universal Education at Brookings, 2012). Before substantiating this claim in greater detail, it is important to explain the basic structure of the education system in Nigeria.

The education system is currently split into pre-primary education, primary education, junior secondary education, senior secondary education and tertiary education, also referred to as the 1-6-3-3-4 scheme denoting the years required to complete each phase (Asaaju, 2015). Primary education is a six-year long endeavor from Grades 1 to 5, where the subjects taught at the primary level include English, mathematics, science, home economics, religious studies (i.e. Christian or Islamic), and one of three prominent indigenous languages (i.e. Hausa-Fulani, Yoruba, or Igbo).

While primary school is free and compulsory as mandated by the UBE since 1999, Nigeria has been home to 13.2 million children out-of-school –the largest in the world. To increase enrolment rates without improving the stagnant state of achievement, however, would counter the

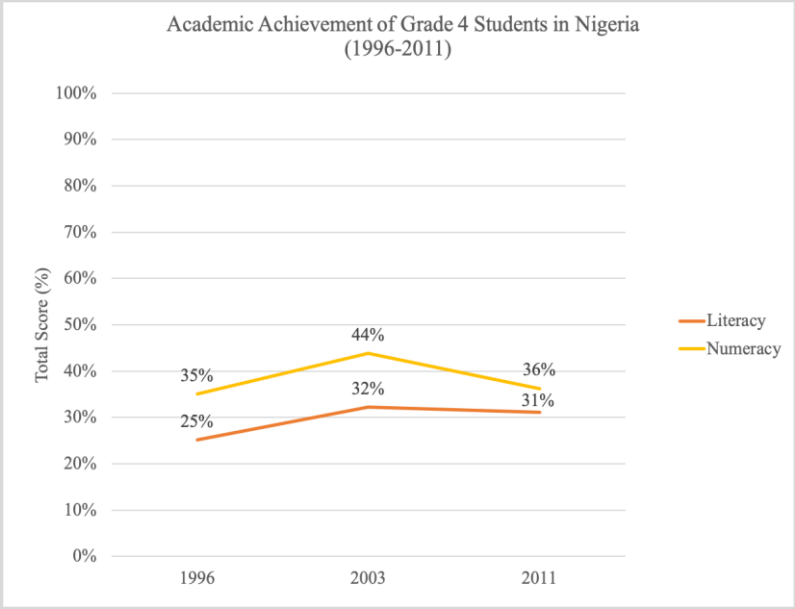
prime purpose of schooling in the first place, which, according to the UBE lies with the "acquisition of literacy, numeracy, life skills and values for lifelong education and useful living" (UBEC, 2020, p. 1). Some scholars criticize assessments as reductionist in format, failing to capture the various facets of learning (Weinberger et al., 2007). Achievement scores may not quantify the value of primary education, but it does serve as one benchmark for learning that can hold policy makers accountable. This section now turns to demonstrating stagnant trend of primary achievement in Nigeria.

The Stagnation of Academic Achievement in Nigerian Primary Schools

A range of international and national learning assessments concur that the academic achievement scores of primary students in Nigeria have been stagnant. In chronological order, this section will demonstrate this through two assessments: Monitoring Learning Achievement (MLA) and Education Sector Support Program in Nigeria (ESSPN). The MLA, widely considered the first nation-wide learning assessment geared at assessing student achievement in Grade 4 and Grade 6, was carried out by the Federal Government of Nigeria in collaboration with the United Nations International Children's Emergency Fund (UNICEF) and UNESCO across three waves (1996, 2003, and 2011) (Falayajo et al., 1997).

Figure 1

Academic Achievement Scores of Grade 4 Students in Nigeria (MLA 1996-2011)

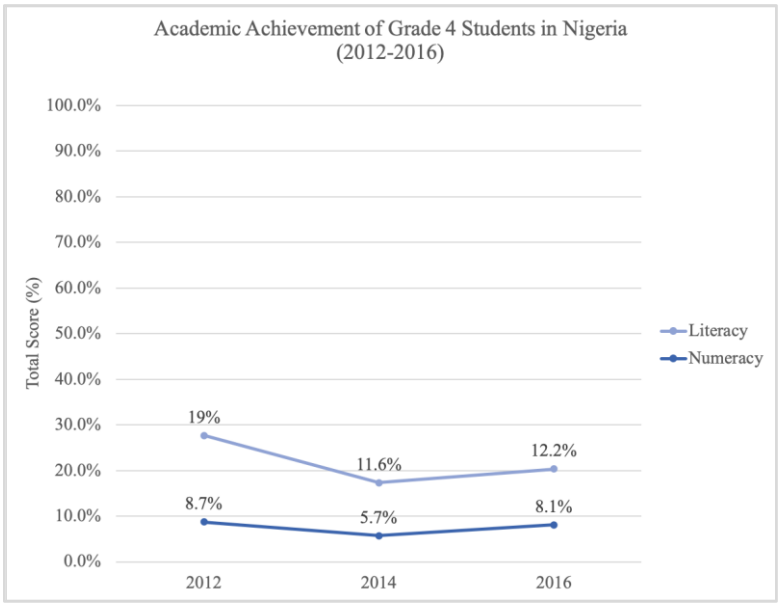


Source. Data extracted from *Reflections on why basic education learning outcomes are declining in Nigeria*, by S. Ogbonna, 2016, Education Sector Support Programme in Nigeria.

The overall findings of the MLA, illustrated in Figure 1, displays the prevalence of low and relatively stagnant achievement in numeracy and literacy score for fourth grade students over time, with signs of a slight decline between 2003 and 2011. The findings of the MLA echo that of the UK Department of International Development's (DFID) ESSPIN. Among other indicators, the assessment evaluated the academic achievement scores of Grade 2 and Grade 4 students in public schools in six states, Enugu, Jigawa, Kaduna, Kano, Kwara and Lago, across three iterations (2012, 2014, 2016). Within the Grade 4 student sample, literacy was measured as the proportion of students who could correctly read a sufficient number of words from a curriculum-level passage as well as could answer at least four out of five reading comprehension items. Numeracy was measured as the proportion of students who could correctly answer addition, subtraction, multiplication, and division questions extracted from the curriculum.

Figure 2

Academic Achievement Scores of Grade 4 Students in Nigeria (ESSPIN 2012-2016)



Source. Data extracted from *ESSPIN Composite Survey 2*, by S. Cameron, 2015, Oxford Policy Management.

As illustrated, Figure 2 suggests that the academic achievement of Grade 4 students in Nigeria has been quite low between 2012 and 2016. While numeracy appears to be somewhat stagnant, there are alarming signs of a decline in the average score assessments in literacy. Derived from these learning outcomes, DFID has posited that "nearly half of all children who have completed primary school cannot read a complete sentence and more than two-thirds of children in the north remain illiterate by the end of primary school" (DFID, 2016). The Africa Learning Barometer, an interactive comparative dataset that relied on the ESSPIN, found that Nigeria was at the bottom rungs of primary achievement in the continent (Center for Universal Education at Brookings, 2012). Overall, the MLA and ESSPIN learning assessments, while not directly comparable due to variations in instruments and sample populations, both underscore the stagnating and sometimes declining levels of fourth grade achievement in literacy and numeracy. Student performance in science has been notably absent from similar assessments in Nigeria.

These two assessments indicate the need for a robust multilevel analysis on predictors of academic achievement. A specific emphasis on investigating the predictive power of educational inputs pertaining to physical facilities and educational materials has the added advantage of dealing with conventional school resources that are often debated in policy circles. Granted, some scholars such as Hickey & Hossain (2019) argue that answers to improving achievement may not necessarily lie within the realm of the educational landscape as hinted earlier in this section (Panglayan, 2017), and, thus, may not concern educational inputs. Given the growing scarcity of funds to basic education in Nigeria, however, one cannot contest the pressing need to examine the input-achievement nexus as a means to complement broader approaches keen on reorienting the education system.

The Deficient State of Educational Inputs in Nigeria

Adopted in 2015, the Education 2030 Framework for Action reaffirmed the international community's commitment towards the actualization of Sustainable Development Goal 4 (SDG 4), a goal established to foster a more inclusive and equitable quality education for all (UNESCO, 2016). Among the targeted objectives, the framework particularly highlighted the pivotal role of educational inputs in improving achievement evident in the statement: "we commit to quality education and to improving learning outcomes, which requires strengthening inputs, processes and evaluation of outcomes and mechanisms to measure progress" (UNESCO, 2015, p. 8). In Nigeria,

scholars, teachers, educators, and civil society organizations have found the state of educational inputs to be highly deficient across primary schools. Agabi (2012) stipulates that "resources provided by [the] government for the execution of education projects in Nigeria are inadequate and irregular" (p. 96).

One can observe the poor state of educational inputs within the six educational inputs selected in this study. In terms of classroom condition, the Federal Ministry of Education revealed that 224,412 primary school classrooms were considered to be in poor condition as of 2019, representing roughly a third of all classrooms (Amoo, 2019).[2] Akin to classroom condition, desks were frequently reported to be broken (Salau, 2017). In addition to classroom and desk conditions, WaterAid in 2018 found that 52% of Nigerian schools did not possess toilets. In terms of textbook availability, there appears to be a dearth of textbooks available for primary students, with Okere & Mnim (2016) revealing that only 20% of all Grade 4 students in Rivers State have access to some type of textbook. In the absence of textbooks, students use exercise books as an alternative source of learning in the classroom to copy from fellow peers or from blackboard instructions, but even there a shortage exists (Fredriksen, 2015). Classroom corner libraries were also reported to be limited in availability (Baro & Eze, 2016). The recent decrease in the federal budget to basic education by 54% prolongs the deficient state of educational inputs in Nigeria, which has serious implications on student learning. A detailed demonstration of how each of these six inputs relate academic achievement will be discussed in the literature review.

[2] Bad classrooms were classified as those possessing no electricity, internet, computers, clear drinking water, single-sex toilets, and hand washing facilities (UNESCO, 2013).

3. Literature Review

This section will discuss pertinent theoretical and empirical literature exploring associations between "educational inputs" and "academic achievement." Split into three subsections, this literature review will, first, conceptualize and operationalize the two aforementioned concepts. The second subsection will, consequently, outline literature investigating how the literature frames relations between the six inputs and student achievement with special attention paid to studies concerning Nigeria. For greater nuance, the third subsection will trace present academic discourse on cross-level interactions effects, specifically regarding the view that educational inputs may moderate the relationship between student biological sex and academic achievement, respectively.

Conceptualizing and Operationalizing "Educational Inputs" & "Academic Achievement"

Literature broadly refers to "educational inputs" as the basic means allocated for the provision of education. Such inputs may refer to a plethora of resources including but not limited to teacher wages, free student meals, indoor toilets, pedagogical materials, school libraries and classroom supplies (Konstantopoulos & Borman, 2011). While various educational inputs may be investigated to further understand what predicts academic achievement among fourth grade primary students in Nigeria, this study will operationalize the concept to specifically denote six inputs—percentage of classrooms in good condition, percentage of desks in functional condition, presence of toilet/s, percentage of students with textbooks, percentage of students with exercise books, and the presence of corner library/ies. To clearly situate this study within the diverse spectrum of published educational research and public policy briefs on educational inputs, the selected six inputs will be subsumed under two thematic clusters. A brief overview of the clusters and the accompanying inputs and perceptions can be observed below:

Cluster 1: Physical Facilities

Educational Input 1: Classrooms in Good Condition

Educational Input 2: Desks in Functional Condition

Educational Input 3: Presence of Toilet/s

Cluster 2: Educational Materials

Educational Input 4: Textbook Availability

Educational Input 5: Exercise Book Availability

Educational Input 6: Presence of Corner Library/ies

In tandem with educational inputs, literature suggests that a consensus on the definition of achievement, often interchangeably used with "student performance," "learning outcomes," and "educational assessments," has not been reached. Guskey & Lung (2012) conveys this discursive contestation in his statement that "not only do policy makers, legislators, school leaders, teachers, parents, and researchers often define student achievement differently, tremendous variation exists in the definitions of individuals within each of these groups" (p. 3). Academics navigate this hurdle using reverse operationalism –a strategy that defines a term using the instrument(s) employed in research (Wigfield & Cambria, 2010). This study relies on reverse operationalism to conjointly define and operationalize "academic achievement" as the total performance of fourth grade students on the English, Mathematics and Science assessments aggregated at the school-level. The decision to use reverse operationalism does not negate the validity of alternative ways of conceptualizing academic achievement nor does it claim to comprehensively capture all possible domains of learning. Rather, it enables this study to circumvent the absence of a consensus on the definition.

Before delving into a literature review on how educational inputs relate academic achievement, this study must unpack the variations in the term "academic achievement," which has the added benefit of positioning the aforementioned definition within an established theoretical foundation. To begin, one must first recognize that "academic achievement" has been considered a multifaceted construct due to its capacity to address an assortment of learning domains, all of which may be measured in distinct ways and for distinct purposes. In education, "academic achievement," relates to the accomplishment of articulated learning goals (Cummins, 2015). And, while learning may certainly transpire beyond the confines of pronounced learning goals, educational research tends to commonly classify learning goals into three salient domains: *cognitive* (Bloom et al, 1956), *affective* (Krathwohl, Bloom, & Masia, 1964), *psychomotor* (Simpson, 1966). Notwithstanding the critical importance of *affective goals* (achievements related to student attitudes, beliefs, emotions, interests, and characters) and *psychomotor goals* (achievements associated with the demonstration of skills and knowledge), *cognitive goals* (achievements linked with concepts and skills that educators aim to instill within their students through instructional activities) will be the principal focus of this study given the relevance of this domain to the selected indicator.

Cognitive goals, moreover, tend to serve as core elements of academic curriculum within formal education institutions and can be applied onto a wide range of subdomains from critical thinking to subject knowledge. The latter subdomain, student subject knowledge in specifically English, Mathematics, and Science, is most relevant to the indicator of this study. After situating the indicator of this study within this theoretical framework, scholarly definitions of "academic achievement" within cognitive goals pertaining to student subject knowledge can now be discerned. Steinmayr et al. (2017), for instance, define "academic achievement" as "performance outcomes that indicate the extent to which a person has accomplished specific goals that were the focus of activities in instructional environments, specifically in school, college, and university" (p. 1). Guskey & Jung (2012) contend that the default usage of performance outcomes may be attributed to the absence of a shared definition and the assessment movement of the mid-1980s. With this narrower theoretical scope, one may draw on various cognitive instruments to operationalize achievement. This study relies on the three learning assessments as opposed to international standardized tests in order to evaluate achievement using instruments that were tailored by academic consultants at the World Bank to fit the local curriculum primary school students in Nigeria. Additionally, it eschews reductionist comparative metrics that unfairly assess low-income nation-states through a Western-centric educational lens.

Educational Inputs and Academic Achievement in the "Global South"

Literature empirically investigating how educational inputs relate to student performance has yielded mixed findings. On one hand, studies such as Glass & Smith (1979), Hedges & Stock (1983), and Hanushek (1996) suggest the presence of some significant relationship, albeit inconsistently, between educational inputs and student performance; on the other hand, more recent findings by Coleman et al. (1996), Glewwe et al. (2009), and Sabarwal et al. (2014) argue that educational inputs, in general, display no effect on student performance. A meta-analysis study conducted by Hedges et al. (1994), ultimately, contends that the data appears "more consistent with a pattern that includes at least some positive relation between dollars spent on education and output, than with a pattern of no effects or negative effects." (p. 5).

In an effort to unscramble this discussion, a close examination of the geographical contexts of these studies, through the theoretical lens of those located in the "Global North" versus the "Global South," can offer incredible insight. The use of these two terms are more powerful than

their economic counterparts as they acknowledge the historical narrative that contributed to the creation of this wealth disparity. A comparison of the geographic origins of research on educational inputs and academic achievement highlights that mixed findings or the absence of significant findings appears to be mainly derived from studies situated in the "Global North," whereas contemporary research conducted in the "Global South," albeit highly limited, suggests that findings appear to display a positive relationship. This contrast should not imply a blind adoption of this statement in absolute terms due to the fact that different inputs and different metrics may produce divergent findings, but rather, the contrast invites scholars to question and interrogate the appropriateness of the global pivot away from school inputs in low-income nations. This disparity further urges us to explore the input-achievement nexus within underrepresented low-income nations to foster a more rich, nuanced, and inclusive discussion, as well as productively address matters related to the efficient and equitable distribution of resources to educational institutions in such countries.

Input Cluster 1: Physical Facilities & Academic Achievement

Consistent with the overall pattern concerning educational inputs and achievement, literature evaluating the relationship between physical facilities and academic achievement tends to yield mixed results on studies conducted in the "Global North" as opposed to those in the "Global South" (Earthman, 2002; Ogunsaju, 1980; Nwagwu, 1978). Mwamwenda & Mwamwenda (1987) reflects this view in his statement: "though some studies in the West have cast doubt on the importance of school facilities (Jencks et al., 1972), a substantial number of studies in the Third world and some Western countries have shown correlations between achievement and the availability of facilities" (p. 228).

Physical facilities –referring to the visible elements of an educational institution including but not limited to furniture, lighting, toilets, libraries, storage facilities, and playgrounds– have been substantially explored within educational research (Earthman & Lemaster, 1996). Until the 1970s, school design professionals assumed that "as long as certain minimum standards for size, acoustics, lighting and temperature were met, a productive environment existed and teaching and learning would proceed normally" (Onolemhenmhen, 2014, p. 146). A compilation of recent studies suggests otherwise. For instance, in Earthman's (1996) meta-analysis review, the author argued that the built environment accounts for a significant portion of the variance in student

achievement, with studies finding weak to strong effects. Despite the fact that physical facilities tend to explain a small portion of the variance in achievement, the limited portion can be controlled through the cooperative efforts of design engineers, educators and school administrators.

In the context of Nigeria, literature supports the notion that good physical facilities can be a predictor of higher academic achievement at least at the secondary-level, with a notable absence of (1) studies empirically examining this relationship at a primary-level, (2) studies assessing classroom condition, desk conditions, and toilets in Nigeria, and (3) studies pursuing a multilevel ordinal logistic regression. Koroye (2016), for instance, finds that the infrastructural facilities significantly predict higher academic achievements of secondary students in Nigeria's Bayelsa State. This result may be attributed to the presence of decaying physical infrastructure across the majority of schools, in urban and rural settings, as reported by the Federal Ministry of Education of Nigeria, but also attributed to the dwindling government expenditure allocated to the education sector (Omotor, 2009).

Educational Input 1: Classrooms in Good Condition

Generally, scholars argue that good classroom conditions can be a predictor of higher student academic achievement (Edwards, 2006; Earthman 1985), but literature investigating this association has been scarce in Nigeria. A rare quasi-experimental study by Onolemhenmhen (2014) found that good classroom conditions positively predict the academic achievement of 450 secondary students in English and Mathematics across three public schools in Nigeria's Edo district. This finding aligns with Ekpo et al. (2009), who posit that classroom climate positively predicts the academic achievement of 1,200 secondary students in Social Science across 24 public schools located in Nigeria's Cross River State. These studies, while insightful, do not critically engage with the primary student population nor with a broader set of educational inputs and academic subjects, reflecting the need for further research. Ultimately, these two publications compel this study to hypothesize that good classroom conditions positively predict academic achievement.

Educational Input 2: Desks in Functional Condition

Given the sedentary nature of formal education, it is imperative that scholars, policy makers, and engineers gather a better understanding of how broken desks in a school predict

student achievement; however, a review of the literature investigating statistical associations between desk conditions and student academic achievement has been sparse (Zacharkow, 1987). As opposed to the functional condition of desks, researchers have heavily explored how the architectural design of desks predict achievement from sit-to-stand desks to desk weight (Purwaningrum et al., 2015). A systematic review by Castellucci et al. (2017) found that two out of the three studies analyzed displayed a positive correlation between classroom furniture design adjustments and student performance (Smith-Zuzovsky & Exner, 2004; Schilling et al., 2003). The third study, however, indicated that a change in the design and dimensions of classroom desks did not significantly contribute to a change in academic achievement (Ryan, Rigby, and Campbell, 2010). While the predictors discussed in these studies slightly differ in this case, they, nevertheless, gesture to the importance of investigating the infrastructural quality of desks in relation to student achievement. With reference to these publications, this study is inclined to hypothesize that the functional desk condition positively predicts academic achievement in Nigeria.

Educational Input 3: Presence of Toilet/s

Though Nigeria struggles with the basic provision of toilets, literature empirically exploring the relationship between the presence of toilet/s and academic achievement has astonishingly been meager. Reflected in Maslow's (1943) five-tier model on the hierarchy of needs, the prevalence of toilet facilities within school grounds addresses a basic physiological human need that must be fulfilled before a student can be expected to pursue higher-level needs such as self-actualization. Yet, within the context of Nigeria, recent reports indicate that roughly 50% of primary schools do not possess a single toilet, leaving students no choice but to openly defecate (Idoko, 2020). A study by Jaiyeoba and Atanda (2011), however, has been a notable exception to the absence of literature, whereby the authors found that toilets positively predicted the Mathematics achievement of 1,014 secondary students in South-Western and North-Central Nigeria. With this publication in mind, this study hypothesizes that toilet/s positively predict academic achievement in Nigeria.

Input Cluster 2: Educational Materials & Academic Achievement

Literature largely agrees that educational materials, defined as the collection of visual, auditory, and kinesthetic resources intended for student learning and teacher instruction, tends to

operate as a positive but moderate predictor of academic achievement in the "Global South" (Jamison et al., 1981; Lockheed et al., 1986; Fuller, 1987). In this study, the term educational materials will be operationalized to encompass three salient inputs: textbook availability, exercise book availability, and the presence of corner library/ies. Scholars assert that there appears to be a positive association between educational materials and academic achievement. A systematic review by Snilstveit et al. (2017) supports this view evident in the statement "teaching and learning materials are both logically and empirically associated with school achievement" (p. 370). However, while literature investigating this relationship in the "Global South" appears to be substantial, this association has not been extensively explored in Nigeria. Among the rare studies available, the literature tends to skew towards a focus on textbook availability with less published on the two textually oriented resources exercise books and corner library/ies.

Concerning textbook and exercise book availability, a rare, albeit outdated, study by Lockheed and Komenan (1986) investigates this association in the context of Nigeria. The authors revealed that the use of textbooks and exercise books displayed a significant positive correlation with student achievement as measured using the Mathematics scores of eighth grade secondary students. However, the authors, simultaneously, found that these same educational inputs demonstrated a significant negative correlation with student achievement in Swaziland. Plausible reasons for the prevalence of mixed findings in this study may relate to the complex relationship between educational materials and academic achievements, particularly that the increased academic benefits of providing educational materials, ultimately, depends on the creation, production, distribution, and utilization of textbooks and exercise books across the learning ecosystem (Fuller & Clarke, 1994). Another layer of complication lies with the use of colonial languages to design these materials, which can unfairly disadvantage and systematically undervalue student performance of those who speak a different language (Mingat et al., 2003).

Educational Input 4: Textbook Availability

Writing on the importance of textbooks to learning, Heneveld and Craig (1996) contend that textbooks are the "single most important instructional material" as it is the common medium used to deliver curriculum and design assessment (p. 23). The integral role of textbooks has been considered especially important in the "Global South" given local variations in teacher training, pedagogical strategies, and school management. Empirical literature has extensively supported the

notion that textbook availability positively predicts academic achievement at the primary level with the effect being particularly salient among under-resourced schools (Heyneman & Jamison, 1980; Mwamwenda & Mwamwenda, 1987). For instance, Heyneman et al. (1978) compiled several studies that analyzed the effects of textbook availability on academic achievement within 10 low-income nations. The authors revealed that fifteen out of the eighteen statistics unearthed from these studies yielded positive correlations, with two studies on Iran yielding negative correlations and one study on Ecuador revealing a non-significant correlation. However, in the context of Nigeria, literature has been scarce with the exception of the Lockheed & Komenan (1986) study published more than three decades ago. In line with the overarching trend in the literature, this study hypothesizes that textbook availability positively predicts academic achievement in Nigeria.

Educational Input 5: Exercise Books

In the context of Nigeria, the few studies on Nigeria that have empirically explored the relationship between exercise books and academic achievement indicate a positive association. Exercise books, a textual learning aid often subsidiary in nature to a textbook, may assume different formats, of which the common format among Nigerian primary students relates to plain lined paper (Momoh, 1980). Bearing in mind possible variations in the format of exercise books, a study by Ebuk and Bamijoko (2016) illuminates this discussion. Using a randomized control trial, the authors found that mathematics workbooks filled with problem sets positively predicted the mathematics performance of 400 secondary school students in the town of Bwari. As a result, this study errs on the side of hypothesizing that exercise book availability positively predicts academic achievement.

Educational Input 6: Presence of Corner Library/ies

Reflecting a general scarcity of input research in Nigeria, literature evaluating the relationship between the presence of at least one classroom corner library at the school-level and academic achievement has been sparse. While scholars have seldom studied corner libraries –a term that refers to the collection of books, learning materials, and teaching resources positioned inside a central location in a classroom– has been seldom studied, literature on the predictive power of school libraries have been more prevalent (Fractor et al., 1989). Building on the findings of

Adeche-Apeji (1990) and Omolayole (2001) who highlighted the decaying nature of school libraries evidenced, Adeyemi (2010) revealed a significant positive relationship between school library use and student performance level using stratified random sampling of 560 students across 120 secondary schools in Nigeria's Ekiti State. The concentration on school libraries as opposed to classroom corner libraries may be attributed to the absence of a significant number of observations, necessary for the pursuit of multilevel models. As a result of Adeyemi's findings, this study is more inclined to hypothesize that the presence of corner library/ies positively predicts academic achievement.

Cross-Level Interaction Effects: Educational Inputs, Biological Sex, & Achievement

For a more nuanced exploration of the input-achievement nexus, this study aims to gauge whether student biological sex predicts academic achievement as well as whether educational inputs moderate the association between student biological sex and academic achievement. In sum, the literature revealed three insights: 1) no scholarly consensus on the association between student biological sex and academic achievement in Nigeria, 2) insufficient educational research on cross-level interaction effects in the "Global North," and 3) available studies on cross-level interaction effects displayed mixed findings.

The thrust of the substantial body of research in Nigeria yielded mixed findings regarding the relationship between student sex and academic achievement, reflecting an absence of scholarly consensus on the matter (Agbaje & Alake, 2014; Unity & Igbudu, 2015). Jegede (1994) contends that biological sex does not predict the English achievement of the 160 secondary students in Nigeria's Kwara State of Nigeria, while Erinosho (1999) argues that biological sex predicts the performance of the 325 secondary students sampled on the Science assessment in the Nigeria's Ogun State. This difference may be due to the subject assessed, the sample measured, and the state selected; however, situating these two studies within a wider scope of literature signals to the importance of biological sex in predicting some of the variance in achievement. Thus, this study hypothesizes that student biological sex does predict academic achievement.

With the absence of literature exploring the specific cross-level interactions in Nigeria, scholars focused on the interactions of student biological sex and academic motivation on achievement (Vecchione et al., 2014) and student sex and teaching style on achievement (Thompson & O'Brien, 1991) as opposed to the interactions of educational input and student

biological sex on achievement. For instance, in the context of the United States, Brown (1991) found that classroom resources did not significantly moderate the association between student biological sex and academic achievement; however, this finding may not necessarily apply to non-classroom resources and to non-American contexts. Considering these limitations, this study is more inclined to hypothesize that educational inputs does moderate the relationship between student sex and academic achievement.

Conclusion

This section reviewed relevant literature exploring associations between educational inputs and academic achievement, with a focus on studies conducted in Nigeria or the "Global South" more broadly. In addition to introducing the overarching narrative, this section discussed cluster-specific and input-specific associations, demonstrating the existence of a significant conceptual and methodological gap in the literature pertaining to the context of Nigeria. This gap is a microcosm of a dearth in the literature on the "Global South." The unfortunate implication of the absence of a contextualized empirical study is that the wider academic discussion would not only be distorted and subsumed under sub-Saharan African meta-analyses, but also Nigeria as a fast-growing under-resourced nation in and of itself necessitates an individual exploration of what determines the variance in the student performance at the primary-level. Available studies operating within the same theoretical and geographic vicinity of this study suggest that educational inputs tend to positively predict academic achievement, thereby, forming the foundation of the ensuing hypotheses. This does not necessarily conflict with existing research arguing that input-based school policies are limited in the actualization of systematic and consistent academic improvement, it merely suggests that there is still merit in evaluating the predictive power of educational inputs within low-income nations such as Nigeria.

4. Research Design and Methodology

In the preceding section, a review of the literature illustrated the absence of scholarly research quantitatively assessing the association between educational inputs and the academic achievement of primary school children within the context of Nigeria. In tandem with an absence of nuanced research, one can observe that, whereas studies investigating the effect of educational inputs on academic achievement on the "Global North" reveal mixed evidence, studies conducted on nations in the "Global South" predominantly support the presence of a positive correlation between these two constructs. Thus, with the intention to further shed light on this multidimensional association, this study proposes the following research questions:

1) Do educational inputs, specifically good classroom condition, functional desk condition, presence of toilet/s, textbook availability, exercise book availability, and the presence of corner library/ies, significantly predict higher academic achievement among fourth grade students in Nigeria?
2) Does the biological sex of students significantly predict higher academic achievement among fourth grade students in Nigeria?
3) Do educational inputs significantly moderate the relationship between student sex and academic achievement among fourth grade students in Nigeria?

To address these questions, this study conducted a two-level ordinal logistic regression analysis using nationally representative secondary data collected, cleaned, and merged by the World Bank between the years 2019 and 2020.

Data Set & Data Access

Under the development project Better Education Service Delivery for All, the World Bank, in collaboration with the Federal Ministry of Education in Nigeria, designed and dispensed assessments to test the proficiency of students and teachers in the subject areas of English, mathematics, and science at the primary (Grade 4) and secondary levels (Grades 6-12) (World Bank, 2020). Alongside these assessments, school leaders were interviewed to measure two conceptual domains: mental models (i.e. school leader perceptions on student learning, school improvements, and institutional concerns) and educational inputs (i.e. available school resources).

As a research assistant to Dr. David Johnson, a chartered psychologist and academic at the University of Oxford's Department of Education, I had analyzed this data set to examine links

between teacher competency on student achievement. With the approval of Dr. Johnson, I was allowed to repurpose this extensive data set for my dissertation. Wary of replicating the current research pursuits of Dr. Johnson exploring associations between teacher competency and student achievement and of mistreating the research endeavor of analysts at the World Bank studying mental models, I decided to investigate the oft convoluted relationship between educational inputs and overall academic achievement among fourth grade students in Nigeria.

Multilevel Modeling

This study used multilevel modeling to investigate the input-achievement nexus. In statistics, a multilevel model allows researchers to test whether relationships between lower level variables (e.g. student sex and academic achievement) varies as a function of higher-order variables (e.g. educational inputs). This approach was selected for its capacity to internalize the nested data structure, whereby individual observations appear clustered within higher-level units. In this study, this structure pertains to the natural configuration of Grade 4 students (level-1 observations) nested within classrooms (level-2 observations) nested within schools (level-3 observations).

According to methodological literature, the prevalence of clustering insinuates that students share some common variance that tends to be associated with learning in the same educational institution (Searle et al., 1992). Thus, any analysis of student achievement predictors using this dataset cannot be treated as independent, which, by extension, implies that the pursuit of any non-multilevel method—such as the traditional multiple linear regression— would produce invalid standard errors and parameter estimates (Rabe-Hesketh & Skrondal, 2008). Yet, which type of multilevel model would be the most empirically valid and theoretically suitable for this research scope?

Using version 15 of the statistical software Stata, this study explored various types and versions of multilevel models before coming to the conclusion that a two-level ordinal logit model (TOLM) would be the most suitable approach (See Appendix C for a detailed explanation of this process). In the search for a valid and parsimonious model, I frequently checked my decisions with my dissertation supervisor as well as received further technical advice from a QuantMethods Statistical Surgery session with Dr. Ellie Suh.

Prior to a TOLM, this study pursued a three-level hierarchical linear model (HLM) –a multilevel model for linear dependent variables however, the data failed to converge in part due to the absence of a significant quantity of observations at the classroom-level (Level-2). The failure to converge, an indicator of a poorly constructed model, persisted despite the down scaling of the mode from a three-level to a two-level HLM due to non-normal residuals (See Appendix D). Alterations in model composition still yielded an unviable model. For instance, the inclusion of a third input cluster pertaining to parent teacher association (PTA) participation to assess the predictive powers of PTA budget and PTA meeting frequency not only produced non-normal residuals, but also posed a risk to the predictive assessments of inputs within the physical facilities and educational materials clusters.

In light of HLM assumption violations and model complications, this study proceeded to explore multilevel ordinal logistic models (MOLM) as an alternative method recommended by statisticians (Liu, 2015). To do so, this study converted the continuous dependent variable into an ordinal dependent variable split into 10 achievement thresholds, but the limited number of observations available for each of the ten thresholds produced a weak model that failed to converge. To curb this problem, this study recategorized academic achievement from ten-thresholds to three-thresholds, which produced a much stronger model (See Appendix H). While various multilevel model types and versions were explored, this study selected the two-level ordinal logistic regression as the most suitable method capable of producing a simple and valid model that aligns with the theoretical background.

Two-Level Ordinal Logistic Regression

A TOLR is a multilevel model designed to fit an ordered dependent variable (i.e. a categorical outcome that follows an ordinal scale) and account for nested data structure (Demidenko, 2004). Similar studies have employed TOLR to assess achievement predictors including Roscingo et al. (2006), Fielding et al. (2003), and Byrd & Chavous (2009), but such publications have seldom applied this method within contexts in the "Global South." This study, therefore, offers a compelling empirical contribution to the literature.

For an informed interpretation of TOLR outputs, this section introduces the terms: "log odds," "odds ratio," "fixed effects," "random effects." TOLR allows researchers to predict the log odds of belonging to a higher category, whereby the term "log odds" refers to the cumulative

probability of a student performing in a higher achievement category (Tamura & Giampaoli, 2010). The exponential form of a log odds is equivalent to the "odds ratio" (OR), which denotes the strength of the association between a given input and achievement (Szumilas, 2010). Log odds and odds ratio are terms that can be used to interpret predictor estimates. These estimates can be considered to exert a fixed effect or a random effect. A fixed effect predictor, known as a fixed parameter, has a slope and intercept that is assumed to be constant, whereas a random effect predictor, known as a random parameter, has a slope and intercept that varies across groups.

In terms of the process of model construction, this study followed best practice by conducting the following five models of TOLR:

Model 1: Null unconditional model

Model 2: Random-intercept model with four fixed effects

Model 3: Random-intercept model with six fixed effects

Model 4: Random-covariate model with moderator effects

Model 5: Random-covariate model with cross-level interaction effects

These five models incrementally introduce parameters to search for a parsimonious model –a simple and valid model that yields substantial predictive power (Tenenbaum & Filho, 2016). Model 1 allows this study to test whether student achievement significantly varies by schools. That is, to test if a multilevel model is warranted in the first place. Models 2 and 3 are statistical models where the predictors are only fixed parameters, whereby it tests for the input-achievement nexus without student age, student biological sex, and moderators. Models 4 and 5 are statistical models where the predictors include fixed and/or random parameters. Mode 4 tests for whether educational inputs moderate the association between student biological sex and academic achievement. Model 5 tests for cross-level indirect effects (For mathematical notations of all models, see Appendix G).

The TOLR employed the maximum likelihood estimation (MLE) as opposed to alternative estimations (e.g. Bayesian) due to its ability to consistently produce unbiased estimates when dealing with large data (Raudenbush & Bryk, 2002). MLA relies on an optimization feature that iteratively searches for the largest log likelihood in order to calculate the probability of attaining a particular achievement category (Verbeke & Molenberghs, 2000).

Sampling Procedure & Sample Size

In terms of the sampling procedure, the World Bank pursued systematic random selection for schools and students to simultaneously increase generalizability and decrease bias. Regarding

schools, data analysts at the multinational organization randomly selected 160 primary schools from a vast list of schools in Nigeria. The World Bank held a training session in Abuja with 148 enumerators, where they discussed data collection procedures and split enumerators into 36 teams (i.e. one team per state). Concerning students, enumerators were requested to randomly select three sets of 10 Grade 4 students to complete only one assessment type. For schools with less than 30 students, random selection was no longer feasible; thus, enumerators were instructed to proceed with administering all three assessments to students while noting assessment overlap using identification numbers. To deliver all assessments and questionnaires, the World Bank provided enumerators a tablet with geolocation features. The first sample was collected on October 21st of 2019 and the last sample was collected on December 5th of 2019, spanning less than two months.

Concerning sample size, the World Bank intended to sample 4,440 students and 160 schools. The achieved size, albeit slightly lower, was quite close to the original objective with a total of 4,016 students and 148 schools. Ideally, all sampled Grade 4 students would be included in the final data analysis, however, to standardize observations and compare logit models, missing values were dropped representing roughly 4% of total observations and resulting in a sample size of 3,856 students in 131 schools. According to Schafer (1999), missing values may be inconsequential if the missing rate is 5% or less, thus, the 4% missing rate is not of significant concern. In addition to missing values, the sample size decreased to a total of 3,236 in 97 schools due to the desk condition parameter given that schools without desks were excluded from the sample (See Appendix K for the shifts in sample size). Despite this decrease, the proportion of school types and location did not significantly change. While the survey collected secondary students and the teacher responses, they are beyond the research scope of this study.

Following data-collection, World Bank analysts merged and weighted the data sets to conjointly construct cluster units and to increase sample size, both of which are necessary for the pursuit of a multilevel logit model (Hedeker & Gibbons, 2006). The data merge involved mapping responses between, on the one hand, school leaders and teachers, and, on the other hand, students. That is, all students assessed in the same school and classroom would have identical school leader and head teacher surveys linked to their identification number. In tandem with the data merge, the data weighted school, school leader, and student observations based on the probability of selection to ensure that the public-private and urban-rural strata were nationally representative. This merge was necessary given that the data oversampled private schools for alternative research.

With regards to financing the survey, the World Bank transferred funds to the Federal Ministry of Education.[3] The ministry then distributed these funds to the State Universal Basic Education Board (SUBEB), the government entity overseeing project implementation, which handled enumerator payments. The enumerators, who were government employees working at SUBEB, received reimbursements and per diems for their work. The compensation of enumerators for their time aligns with ethical constructs of "giving back." And, while it is unknown whether enumerators reimbursements or affiliation with SUBEB may have biased responses, their presence in the data collection likely improved access to schools.

The World Bank did not receive approval from an institutional review board to conduct this study, rather, they were content with using the signed agreements of the Permanent Secretary Arc Sonny Echono and the Education Minister Adamu Adam to pursue a government sanctioned exercise. This approval did not mean that ethical guidelines were neglected. World Bank data analysts stressed that enumerators were instructed to respectfully ask for consent and to clearly inform all participants of their data privacy rights, explicitly that all collected information would be de-identified. Enumerators informed participants that they may leave at any point in time.

This dissertation did not need approval from the Central University Research Ethics Committee, nonetheless, it did aim to abide by best practice in secondary data analysis. Strong ethical arguments further support the use secondary data in this dissertation, in which this method can foster greater transparency, encourage replicability of empirical research, maximize the investment value allocated towards data collection by the World Bank and the Ministry of Education, reduce the burden on Nigerian students and SUBEC enumerators to participate in future research, and illuminate the literature on educational inputs in Nigeria and the "Global South" (Tripathy, 2013).

Questionnaire Design for the English, Mathematics, and Science Assessments

The English assessment consists of 31 questionnaire items, the Mathematics assessment contains 34 questionnaire items, and the Science assessment has 53 questionnaire items (See Appendix L for full extracts of the three assessments). As opposed to a deterministic design that compartmentalizes and confines student achievement to reaching certain scores, Dr. Johnson

[3] The World Bank cannot publicly disclose the precise amounts allocated to all parties officially given that an official report on the project has yet to be published.

designed a probabilistic assessment for each of the three subjects through a combination of TIMSS, PIRLS, the Grade 4 Nigerian curriculum (Wilson, 1989). Nigerian educational consultants reviewed these assessments to ensure that the questions correspond to the curriculum.

Reliability and Validity

Various elements in this project were used to increase the reliability[4] and validity[5] of the findings. In addition to random sampling methods and Dr. Johnson's internal consistency tests, android tablets and code checks enhanced confidence in the reliability and validity of this study. In the data collection phase, the World Bank loaned tablets to every enumerator and, consequently, instructed them to use the geolocation feature to track the time, date, and location of all respondents. Not only did tablets allow for rapid digital transfer of data from collection to cleaning, but also, it enabled the study to minimize the likelihood of enumerators filling out paperwork themselves as was rampant within development research in sub-Saharan Africa (Hughes & Lin, 2018). In the data analysis phase, economists independently checked my 720 lines of code to ensure assumptions were appropriately tested and models were logically constructed based on statistical best practice. Further attention was given towards the identification of syntax errors. These two methods increased response accuracy, measurement quality, and, ultimately, garners greater trust in the reliability and validity of the findings discussed in this study.

Variable Selection Process

With the principal aim to create a theoretically and empirically valid simple model, this study selected the following clusters and variables for this study (See Appendix M to view the original variable questionnaire items within the teacher interview and school survey).

Clusters

The decision to place classroom condition, desk condition, and toilets within the physical facilities cluster and to place textbook availability, exercise book availability, and corner libraries within the educational materials cluster was, in part, influenced by studies exploring educational

[4] Reliability refers to the consistency of a measure. A particular piece of research can demonstrate high reliability if it can produce consistent findings under similar conditions.

[5] Validity refers to the accuracy of a measure in representing the real world. A given research study, for instance, can demonstrate high validity if the method used to produce findings conforms to real physical or social characteristics.

inputs and determined by the substantial observations available for each input. Ideally, this study would have explored a broader spectrum of clusters and their corresponding inputs including budgets (e.g. water budgets, construction budgets, teacher budgets), supplies (blackboards, chalk, electricity), toilets (e.g. separated, clean, accessible), but due to constraints in sample size and model capacity this study opted for a more parsimonious model design. And, while there was some hesitancy in classifying corner libraries as educational material input as opposed to the physical facilities input, this decision aligns with much of the discourse in which corner libraries are viewed as a learning resource for students.

Through the use of Cronbach's Alpha —a test of internal consistency used to assess how closely related a series of items are in measuring an underlying construct, this study found that the theoretical association among educational inputs in a given cluster did not empirically translate to significant internal reliability (See Appendix B). Thus, this study will not explore interaction effects between inputs within each cluster but will instead individually explore predictor effects.

Level-2 Variables

This study selected educational inputs and student age as level-2 fixed variables to examine how they predict academic achievement while controlling for nested data. In ideal circumstances, this study would have classified educational inputs as level-2 random variables, student age as a level-1 random variable, and add level-2 variables such as school ownership type, location, and budget; however, all these model compositions failed to converge reflecting a rather weak model. Thus, to minimize threats to validity, this study selected the six educational inputs and student age.

Level-1 Variables

This study selected overall academic achievement as a level-1 outcome and biological sex as level-1 random variable in alignment with empirical literature. The prospect of incorporating alternative variables associated with student-level achievement, most notable parental socio-economic status and previous achievement scores, was not feasible due to their absence from the data collection process. With the realm of data collected, the inclusion of these two level-1 variables in the model heeds the recommendations of multilevel modeling experts.

Variables

Academic Achievement: This dependent variable measures the academic performance of fourth grade students at the school-level through an aggregation of student scores on all three subjects, English, mathematics and science, as a percentage of overall achievement. After transforming the outcome from a continuous variable into an ordinal variable, overall academic achievement has been uniformly split into three thresholds. The higher the threshold, the higher the student performed on a given assessment.

Achievement Threshold 1 (%):	0 - 33.33
Achievement Threshold 2 (%):	33.34 - 63.33
Achievement Threshold 3 (%):	63.34 – 100

Good Classroom Condition: This level-2 fixed parameter refers to the percentage of classrooms enumerators classified as "good" in a given school. The proportion of bad classrooms to total classrooms was converted into a percentage and has been treated as a continuous variable.[6]

Functional Desk Condition: This level-2 fixed parameter refers to the percentage of functional desks classified as "not broken" by enumerators at a given school and has been treated as a continuous variable. In light of the sampling method (i.e. enumerators measured desks in one randomly selected classroom), the quantity of functional desks was multiplied by the quantity of classrooms available to produce a more representative school-level estimate. This proportion was then turned into a percentage.[7]

Presence of Toilet/s: This level-2 fixed parameter, treated as a dichotomous variable, refers to whether one can identify at least one school toilet in its premises, whereby 1 represents the presence of one or more toilets and 0 represents the absence of a toilet.

[6] The variable, good classroom condition, was originally coded as the percentage of classrooms in "bad" condition. This study reversed the direction of the variable to reflect classrooms in "good" condition to compare estimates with other inputs and align with the theoretical discourse on the input-achievement nexus. One must bear in mind that STATA outputs in the Appendix display the variable in its original form (i.e. classrooms in bad condition).

[7] Similar to good classroom condition, functional desk condition was originally coded as the percentage of desks in "broken" condition, which was reversed to denote classrooms in "non-broken" condition. All outputs in the Appendix display the variable in its original form (i.e. desks in broken condition).

Textbook Availability: This level-2 fixed parameter, treated as a continuous variable, refers to the percentage of students with textbooks in a given school. After randomly selecting a class, enumerators measure the number of students who held a textbook. This study multiplied by this total by the number of classrooms in a given school for a more representative estimate and converted it into a percentage to derive stronger inferences.

Exercise Book Availability: This level-2 fixed parameter, managed as a continuous variable, refers to the percentage of students with exercise books in a given school; thus, it has been treated as a continuous variable. Upon the random selection of a class, enumerators counted the number of raised hands indicating they had brought an exercise book to class. Similar to textbooks, this study multiplied this total by the number of classrooms in a given school and converted it into a percentage. Information on the type of exercise book was not recorded.

Presence of Corner Library/ies: This level-2 fixed parameter has been treated as a dichotomous variable, whereby 1 denotes the presence of at least one classroom corner library and 0 denotes the absence of a classroom corner library at the school-level. Information on materials in the corner library was not documented.

Student Age: This level-1 fixed parameter, treated as a continuous variable, measures student age in terms of years. In this dataset, student age ranges from seven years old to nineteen years old.
Student Sex: This level-1 random parameter, computed as a dichotomous variable, categorizes students into female or male students, whereby 1 refers to females and 0 refers to males.

5. Hypotheses

This study proposes eight hypotheses (H1-H8), whereby H1 will test the overarching association between all educational inputs (Level-2 predictors) and student achievement (Level-1 outcome), H2-H7 will test for the cross-level direct effects between individual educational inputs (Level-2 predictors) and academic achievement (Level-1 outcome), H8 will test for the lower-level direct effects between student sex (Level-1 predictor) and academic achievement (Level-1 outcome), and, lastly, H9 will test whether student sex moderates the relationship between the six educational inputs and academic achievement. Reflecting on how these hypotheses will relate to the three multilevel logit models (M1-M5), it is pertinent to note that M1 will be used to evaluate whether academic achievement significantly varies across schools, M2-M3 will test H1-H7, M4 will test H8, and M5 will test H9.

Cross-level Direct Effects for All Inputs

Hypothesis 1 (H1): Educational Inputs

Plain English Form

H_0: Educational inputs, as measured by good classroom condition, functional desk condition, presence of toilet/s, textbook availability, exercise book availability, and the presence of corner library/ies, does not predict the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

H_a: Educational inputs, as measured by good classroom condition, functional desk condition, presence of toilet/s, textbook availability, exercise book availability, and the presence of corner library/ies, predict the log odds of among Grade 4 primary school students in Nigeria

Algebraic Form

H_0: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y \leq K)] = (a_k - u_{0j}) - (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j$ (1)
$\gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij})$,
$\gamma_n = 0$ for at least one predictor

H_a: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y \leq K)] = (a_k - u_{0j}) - (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j +$ (2)
$\gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij})$,
$\gamma_n \neq 0$ for at least one predictor

Cross-level Direct Effects of Cluster 1: Physical Facilities

Hypothesis 2 (H2): Good Classroom Condition

Plain English Form

H_0: Controlling for predictors specified in the logit model, the percentage of classrooms in good condition does not predict the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

H_a: Controlling for predictors specified in the logit model, the percentage of classrooms in good condition positively predicts the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

Algebraic Form

H_0: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y \leq K)] = (a_k - u_{0j}) - (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j +$ (3)
$\gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij})$, $\gamma_{01} = 0$

H_a: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y \leq K)] = (a_k - u_{0j}) - (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j +$ (4)
$\gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij})$, $\gamma_{01} > 0$

Hypothesis 3 (H3): Functional Desk Condition

Plain English Form

H_0: Controlling for predictors specified in the logit model, the percentage of desks in a functional condition does not predict the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

H_a: Controlling for predictors specified in the logit model, the percentage of desks in a functional condition positively predicts the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

Algebraic Form

H_0: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}),\ \gamma_{02} = 0$ (5)

H_a: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}),\ \gamma_{02} > 0$ (6)

Hypothesis 4 (H4): Presence of Toilet/s

Plain English Form

H_0: Controlling for predictors specified in the logit model, the presence of toilets in schools does not predict the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

H_a: Controlling for predictors specified in the logit model, the presence of toilets in schools positively predicts the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

Algebraic Form

H_0: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}),\ \gamma_{03} = 0$ (7)

H_a: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}),\ \gamma_{03} > 0$ (8)

Cross-level Direct Effects of Cluster 2: Educational Materials

Hypothesis 5 (H5): Textbook Availability

Plain English Form

H_0: Controlling for predictors specified in the logit model, the percentage of students with textbooks available does not predict the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

H_a: Controlling for predictors specified in the logit model, the percentage of students with textbooks positively predicts the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

Algebraic Form

H_0: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}),\ \gamma_{04} = 0$ (9)

H_a: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}),\ \gamma_{04} > 0$ (10)

Hypothesis 6 (H6): Exercise Book Availability

Plain English Form

H_0: Controlling for predictors specified in the logit model, the percentage of students with exercise books does not predict the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

H_a: Controlling for predictors specified in the logit model, the percentage of students with exercise books positively predicts the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

Algebraic Form

H_0: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j$ (11)
$+ \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}), \gamma_{05} = 0$

H_a: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j$ (12)
$+ \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}), \gamma_{05} > 0$

Hypothesis 7 (H7): Presence of Corner Library/ies

Plain English Form

H_0: Controlling for predictors specified in the logit model, the presence of at least one corner library does not predict log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

H_a: Controlling for predictors specified in the logit model, the presence of at least one corner library positively predicts the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

Algebraic Form

H_0: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j$ (13)
$+ \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}), \gamma_{06} = 0$

H_a: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j$ (14)
$+ \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}), \gamma_{06} > 0$

Lower-Level Direct Effects

Hypothesis 8 (H8): Student Sex

Plain English Form

H_0: Controlling for predictors specified in the logit model, student sex does not predict the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

H_a: Controlling for predictors specified in the logit model, student sex positively predicts the log odds of performing within a higher category of academic achievement among Grade 4 primary school students in Nigeria

Algebraic Form

H_0: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}),\ u_1 = 0$ (15)

H_a: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}),\ u_1 > 0$ (16)

Cross-Level Interaction Effects

Hypothesis 9 (H9): Cross-Level Interaction Effects

Plain English Form

H_0: The strength of the relationship between student sex and the log odds of performing within a higher category of academic achievement does not change as a function of educational inputs and student age

H_a: The strength of the relationship between student sex and the log odds of performing within a higher category of academic achievement positively changes as a function of educational inputs and student age

Algebraic Form

H_0: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j +$ (17)

$\gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + \gamma_{08}\text{ClassroomCondition}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{09}\text{DeskCondition}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{10}\text{Toilets}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{11}\text{Textbooks}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{12}\text{ExcerciseBooks}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{13}\text{CornerLibraries}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{14}\text{StudentAge}_j \text{ x } \text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}), \gamma_{08} = \gamma_{09} = \gamma_{10} = \gamma_{11} = \gamma_{12} = \gamma_{13} = \gamma_{14} = \gamma_{15} = 0$

H_a: For $\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + \gamma_{08}\text{ClassroomCondition}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{09}\text{DeskCondition}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{10}\text{Toilets}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{11}\text{Textbooks}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{12}\text{ExcerciseBooks}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{13}\text{CornerLibraries}_j \text{ x } \text{StudentSex}_{ij} + \gamma_{14}\text{StudentAge}_j \text{ x } \text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij}), \gamma_{08\text{-}15} > 0$ for at least one cross level interaction (18)

6. Results

In accordance with the proposed research design, the results section will present the data findings of this study, report the findings of the two-level ordinal logistic regression models in order to test hypotheses H1-H7 in a data structure where 3,236 students (level-1) were nested within 97 schools (level-2). Of specific interest was the relationship between student academic achievement (level-1 ordinal outcome variable) and the following six educational inputs: classroom condition, desk condition, toilets, textbooks per student, exercise books per student, and corner libraries (level-2 predictor variables). For a descriptive analysis of primary schools by ownership type and location, primary school leaders by position, fourth grade students by age and sex, student academic achievement by assessment type and threshold, and educational inputs by cluster, see Appendix A.

Data analysis centered on five models: an unconditional null model (M1), a random-intercept model inclusive of four educational inputs (M2), a random-intercept model inclusive of all educational inputs and student age (M3), a random-covariate model with student sex as the random effect parameter (M4), and, a random-covariate model with cross-level interactions between the fixed school-level parameters (i.e. educational inputs and student age) and random student-level parameter (i.e. student sex) (M5). The parameter estimates of all five models can be illustrated in Table 1.

Two-Level Ordinal Logistic Regression Analysis

Table 1

Two-Level Ordinal Logistic Regressions of Educational Inputs and Academic Achievement Among Fourth Grade Students in Nigeria

Variables	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effect Estimates					
Achievement Threshold 1 (a_1)	-1.132	-0.873	-1.030	-2.570	-2.552
	(0.02)	(0.294)	(0.311)	(0.517)	(0.518)
Achievement Threshold 2 (a_2)	1.594	1.889	1.714	0.213	0.233
	(0.02)	(0.296)	(0.312)	(0.513)	(0.514)
Classroom Condition (γ_1)		0.008***	0.009***	0.009***	0.008***
		(0.003)	(0.003)	(0.003)	(0.003)
Desk Condition (γ_2)		0.009**	0.011***	0.012***	0.014***
		(0.004)	(0.004)	(0.004)	(0.004)
Presence of Toilets (γ_3)		0.167	0.165	0.109	0.131

		(0.222)	(0.224)	(0.223)	(0.244)
Textbooks Availability (γ_4)		0.012***	0.014***	0.013***	0.013***
		(0.002)	(0.002)	(0.002)	(0.002)
Exercise Books Availability (γ_5)			-0.005*	-0.005*	-0.007**
			(0.003)	(0.003)	(0.004)
Corner Libraries (γ_6)			-0.517**	0.558***	-0.511**
			(0.214)	(0.214)	(0.234)
Age (γ_7)				0.096***	-0.081***
				(0.027)	(0.030)
Cross-Level Interaction Effects	-1.132	-0.873	-1.030	-2.570	-2.552
Sex & Classroom Condition (γ_8)	(0.02)	(0.294)	(0.311)	(0.517)	(0.518)
	1.594	1.889	1.714	0.213	0.233
Sex & Desks Condition (γ_9)	(0.02)	(0.296)	(0.312)	(0.513)	(0.514)
		0.008***	0.009***	0.009***	0.008***
Sex & Toilets (γ_{10})		(0.003)	(0.003)	(0.003)	(0.003)
		0.009**	0.011***	0.012***	0.014***
Sex & Textbooks (γ_{11})		(0.004)	(0.004)	(0.004)	(0.004)
		0.167	0.165	0.109	0.131
Sex & Exercise Books (γ_{12})		(0.222)	(0.224)	(0.223)	(0.244)
		0.012***	0.014***	0.013***	0.013***
Sex & Corner Libraries (γ_{13})		(0.002)	(0.002)	(0.002)	(0.002)
			-0.005*	-0.005*	-0.007**
Sex & Age (γ_{14})			(0.003)	(0.003)	(0.004)
			-0.517**	0.558***	-0.511**
Random Effect Estimates			(0.214)	(0.214)	(0.234)
School-Level Intercept (τ_{00})				0.096***	-0.081***
				(0.027)	(0.030)
Student-Level Intercept (τ_{01})	-1.132	-0.873	-1.030	-2.570	-2.552
	(0.02)	(0.294)	(0.311)	(0.517)	(0.518)
Sex (τ_{10})	1.594	1.889	1.714	0.213	0.233
	(0.02)	(0.296)	(0.312)	(0.513)	(0.514)
Model Fit Summary					
Likelihood Ratio Test (χ^2)	1334.67***	858.49***	861.29***	833.84***	829.64***
Intraclass Correlation Coefficient (ICC)	0.5256	0.433	0.433	0.437	0.437
Akaike's Information Criterion (AIC)	5610.93	5560.48	5557.84	5545.46	5553.59
Bayesian Information Criterion (BIC)	5629.18	5603.05	5618.66	5606.28	5656.99
Degrees of Freedom (df)	3	7	6	10	17
Observations	3,236	3,236	3,236	3,236	3,236
Number of groups	97	97	97	97	97

Note. Standard errors in parentheses
*** p <0.01, ** p <0.05, * p <0.1

Model 1 (Unconditional Null Model)

Table 1. illustrates that there was a significant between-group variance: $\tau_{00} = 3.646$, which is the intercept variance across all schools. The log likelihood ratio test comparing the

unconditional model to the single-level ordinal logistic regression model suggests that the unconditional model that estimated the between group variance was preferred ($\chi^2_{(1)} = 1334.67$, $p < 0.001$). The ICC in the unconditional model is 0.526, which indicated that 52.6% of the total variance in the chances of performing within a higher academic achievement percentage bracket was accounted for by schools in level-2. Therefore, the use of a multilevel ordinal logistic regression was justified (See Appendix J for all hand calculations of ICC).

Model 2 (Random-Intercepts Model with Four Fixed Parameters)

Model 2 was embedded in this series of logistic regressions due its superiority in model fit relative to other versions of Model 3 (See Appendix I for a parameter-by-parameter construction of model 3 from M3.1 to M3.6 as well as for a list of log likelihood ratio tests).

The logit coefficient for classroom condition ($\gamma_1 = .008$, $p < .01$, OR = .99) indicates that a one-unit increase in the percentage of classrooms in good condition marginally increases the logit of students performing within a higher category of academic achievement by a factor of 0.008. The corresponding odds ratio suggests that students in schools with better classroom conditions scoring are 0.99 times more likely to perform within a higher achievement category than students in schools with worse classroom conditions, holding predictors in M2 constant. Therefore, using a 99% confidence interval, this study can conclude that at least one educational input can predict higher academic achievement; in this instance, classrooms in good condition appear to positively predict higher academic achievement. Thus, this study can, by extension, reject the null hypothesis of H1 and H2 in light of the findings of M2.

Similarly, the logit coefficient for desk condition ($\gamma_2 = .009$, $p < .05$, OR= .99) suggests that a one-unit increase in the percentage of desks in functional condition slightly increases the logit of students performing within a higher category of academic achievement by a factor of 0.009. Controlling for all other predictors in M2, the accompanying odds ratio implies that students in schools with a greater percentage of functional desks are 0.99 times more likely to perform within a higher achievement category compared to students in schools with a lower percentage of functional desks. Therefore, using a 95% confidence interval, the odds of Grade 4 students yielding scores at a higher academic achievement category seems to be empirically associated with fewer broken desks in schools. As a result of the findings of M2, this study can reject the null hypothesis of H3.

The logit coefficient for the presence of toilet/s (γ_3 = .167, p = .541, OR= 1.18) indicates the possession of at least one toilet in schools increases the logit of students performing within a higher category of academic achievement. The related odds ratio reveals that students with toilet access are 1.18 times more likely to perform within a higher achievement category compared to students with no toilet access, controlling for all other predictors. This finding is, however, not statistically significant. Thus, we cannot conclude that the presence of toilets positively predicts better academic achievement; hence, this study fails to reject the null hypothesis of H4 based on M2.

The logit coefficient for textbook availability (γ_4 = .012, p < .01, OR = 1.01) indicates that a one-unit increase in the percentage of students with textbooks in schools corresponds to a faint increase in the logit of performing within a higher category of academic achievement by a factor of 0.012. When controlling for all other predictors specified in M2, the odds ratio suggests that students in schools with a larger percentage of textbooks available are 1.01 times more likely to perform within a higher achievement category than students in schools with a smaller percentage of textbooks available. On the basis of M2, this study concludes that textbook availability positively predicts higher academic achievement, which provides enough grounds to reject the null hypothesis of H5 using a 99% confidence interval.

The ICC in model 2 is 0.433, indicating that 43.3% of the total variance in the chances of performing within a higher academic achievement percentage bracket can be explained by between-school differences; conversely, 56.7% of the total variance can be accounted for by within-school differences.

Model 3 (Random-Intercepts Model with Six Fixed Parameters)

The logit coefficient for the exercise book availability (γ_5 = -.005, p < .093, OR =0.99) indicates that a one-unit increase in the percentage of students with exercise books in schools correlates with a 0.005 decrease in the logit of students performing at a higher category of academic achievement. Controlling for all other predictors in M3, the related odds ratio reveals that students in schools with a larger percentage of exercise books available are 0.99 times more likely to perform within a higher achievement category compared to students in schools with a smaller percentage of exercise books available. However, the 90% confidence interval suggests that there is insufficient evidence to argue that exercise book availability adversely predicts higher

academic achievement. Thus, this study fails to reject the null hypothesis of H6 from the lens of M3.

The logit coefficient for the presence of corner library/ies (γ_6 = -.517, p < .05, OR = .59) indicates the prevalence of at least one classroom corner library in school decreases the logit of students performing within a higher category of academic achievement by a factor of 0.517. Controlling for all other predictors in M2, the odds ratio estimate suggests that students in schools with one or more corner libraries are 0.59 times more likely to perform within a higher achievement category compared to students in schools that do not possess a corner library. Therefore, using a 95% confidence interval, this study finds that the presence of at least one classroom corner library adversely predicts higher academic achievement. With a similar logic to that of exercise book availability, this result deviates from the hypothesized positive association, which implies that this study fails to reject the null hypothesis of H7 based on M3.

Mirroring M2, the ICC in M3 remains at 0.433 indicating that 43.3% of the total variance in the chances of performing within a higher percentage category of academic achievement can be explained by between-school differences. Inversely, the ICC reveals that 56.7% of the total variance can be accounted for by within-school differences.

Model 4 (Random-Covariate Model with Seven Fixed Parameters and One Random Parameter)

The logit coefficient for the seventh fixed parameter, student age (γ_7 = -.096, p < .01, OR = .91), signifies that a one-year increase in the age of fourth grade students decreases the logit of students performing within a higher category of academic achievement by a factor of 0.096. Controlling for all other predictors in M4, the odds ratio estimate implies that older students are 0.91 times more likely to perform within a higher achievement category compared to their younger counterparts. In light of these findings, this study can empirically argue with a 99% confidence interval that age adversely predicts higher academic achievement.

The logit coefficient for the only random parameter, student sex (τ_{10} = 2.38e-41, OR = 1.51e-23), reveals that female students increases, albeit faintly, the logit of performing above a particular level of academic achievement by a factor of 2.38e-41. Controlling for all other predictors in M4, the odds ratio estimate indicates that female students are 1.51e-23 times more likely to perform within a higher achievement category compared to male students. This finding is not significant enough for even the statistical software to return a p-value; thus, we cannot

conclude that student sex can predict higher academic achievement. Thereby, this study fails to reject the null hypothesis of H8 based on M4.

The ICC in M4 is 0.437, reflecting a minor increase of 0.004 from the ICC in M2 and M3. The interpretation of this estimate implies that 43.7% of the total variance in the chances of performing within a higher academic achievement percentage bracket can be explained by between-school differences, while 56.3% of the total variance can be accounted for by within-school differences.

Model 5 (Random-Covariate Model with Cross-Level Interactions)

The estimates of the fixed effect parameters and the random effect parameter in M5 were fairly consistent with the previous results of M2 to M4, thus, this section will not outline the findings of the seven parameters but rather report the findings of the cross-level interactions.

The logit coefficient for the cross-level interaction between the random parameter (student sex at level-1) and the fixed parameters (six educational inputs and student age at level-2) yielded miniscule and non-significant findings. For instance, the interaction between student sex and classroom condition (γ_8 = -.002, p = .389, OR = 1) suggests that the percentage of classrooms in bad condition within schools does not significantly moderate the relationship between student sex and the logit of performing above a particular level of academic achievement. The prevalence of non-significant small estimates can be further illustrated in cross-level interactions between student sex and desk condition (γ_9 = .003, p = .262, OR = 1), student sex and presence of toilet/s (γ_{10} = -.026, p = .889, OR = .97), student sex and textbook availability (γ_{11} = .002, p = .453, OR = 1), student sex and exercise book availability (γ_{12} = .003, p = .293, OR = 1), student sex and presence of corner library/ies (γ_{13} = -.078, p = .688, OR = .93), and student sex and age (γ_{14} = -.024, p = 338, OR = .98). These results are likely an outcome of the non-significant results of student biological sex. Thus, this study does not have enough evidence to argue that educational inputs and student age significantly moderate the relationship between sex and achievement; hence, this study fails to reject the null hypothesis of H9.

Reproducing the findings of M4, the ICC in M5 does not depart from 0.437. The interpretation of this estimate suggests that 43.7% of the total variance in the chances of performing within a higher academic achievement percentage bracket can be explained by between-school differences, while 56.3% of the total variance can be accounted for by within-school differences.

Variance and Covariance Comparisons

Regarding the variance and covariance components of the random effects, the between-school variance, also referred to as the intercept variance (τ_{00}), decreased from 3.65 in M1 to 2.52 in M2 as a result of the introduction of the four educational inputs, which indicates that shifting from the null model to the random-intercept model decreased between-school variance by 31%. Consequently, the introduction of exercise book availability and corner library/ies as two fixed parameters to M3, counterintuitively, increased between-school variance from 2.52 to 3.148, reflecting a rise of 20% as a result of this model shift. The initial drop and the consequent rise imply that the inclusion of all six inputs in M3 does not better explain the contextual phenomenon underlying student academic achievement when compared to merely four inputs in M2. That is, more educational inputs do not equate to more between-school variance explained. However, the random-covariate model without (M4) and with cross-level interactions (M3) counters this narrative as they both display a higher between-group variance of 2.56 relative to M3, albeit a lower between-group variance when compared to M1 and M2. This second rise may be attributed to the inclusion of additional statistically significant variables in the model, namely student age. In tandem with the intercept variance, the slope variance (τ_{10}) for student sex, prevalent in only M4 and M5, yields near zero results in both instances ($\tau_{10\ (1)}$ = 1.66e-33, $\tau_{10\ (2)}$ = 2.38e-41). This finding implies that the effect of student sex on the odds of being at a higher achievement category does not significantly vary from one school to another, which appears contrary to the expectations of literature propagating the view that student performance can often relate to biological sex. Similar to slope variance, the covariance —the extent to which two random variables (i.e. the intercept and the slope) vary conjointly— remains near zero throughout the construction of the five models ($\sigma_e^{2}{}_{(1)}$ = 1.94e-33, $\sigma_e^{2}{}_{(2)}$ = 3.99e-33, $\sigma_e^{2}{}_{(3)}$ = 3.47e-34, $\sigma_e^{2}{}_{(4)}$ = 1.69e-33, $\sigma_e^{2}{}_{(5)}$ = 2.36e-33), reflecting a weak and possibly an independent relationship.

Model Selection

Likelihood Ratio Tests

A common way to assess and compare the fit of models lies with the use of the log likelihood ratio test. This goodness of fit test relies on parameter estimates and degrees of freedom to ascertain the difference between the log likelihoods of two nested models: a full model and a restricted model. A significant p-value would mean that the full model with more parameters has a better model fit. Applied on all multilevel models in this study, the results indicate that M4 appears to be significantly superior in model fit relative to alternative models including the parameter-by-parameter constructions of model 3 (See Appendix H and Appendix I for log likelihood ratio tests for M1-M5 and M3.1-M3.6, respectively).

Akaike's Information Criterion & Bayesian Information Criterion

Extensively employed in model selection, this study examines the AIC and BIC values to corroborate the model selection decisions of the LR tests, while ensuring parsimony –the selection of the simplest model without a substantial loss in the amount of data fit. The two tests are penalized-likelihood criteria computed as a function of the maximized likelihood and parameter count (Akaike, 1974; Schwarz, 1978). However, they mathematically differ in the sense that the Akaike approach measures the relative goodness of fit of any estimated model through a comparison of the fitted likelihood function and the unknown true likelihood function derived from the data, whereas the Bayesian approach measures the relative goodness of fit of only parametric models through the estimation of the posterior probability of a model appearing true under a certain Bayesian framework (Dziak et al., 2012; Atkinson, 1981). In both cases, however, a lower AIC and BIC value would denote a better model fit (Vrieze, 2012).

In this study, the AIC estimates indicate that M4 has the lowest value (M1 = 5611 , M2 = 5560, M3 = 5558, M4 = 5545 , M5 = 5554), but the BIC estimates suggest that M2 has the lowest value (M1 = 5629, M2 = 5603, M3 = 5619, M4 = 5606, M5 = 5657). On the surface, one may argue for the prevalence of a conflict in model selection. Which of the two models, M2 or M4, should be selected based on better model fit? Upon digging deeper into the value and estimation differences, this study can confidently dispel this apparent conflict. With regards to value differences, the BIC between M2 and M4 differ by a small amount of 3. With regards to estimation differences, Kuha (2004) contends that BIC penalizes additional parameters more strongly compared to AIC, thereby, favoring M2 due to the presence of four fewer parameters compared to M4. When compounding the value and estimation differences, this conflict in model selection may

not be such a salient concern. From a critical view, some may argue that M2 might be a more parsimonious selection, however, with the large sample size, it seems that choosing M2 over M4 might jeopardize a substantial loss in data fit. This study, thus, selects M4 as the best model fit in alignment with the findings of the likelihood ratio tests.

Conclusion

This section conducted a two-level ordinal logistic regression analysis to investigate the relationship between educational inputs and academic achievement. Three key insights were derived as a result. First, the ICC of the null model validated the pursuit of a multilevel model. Secondly, the findings supported the rejection of hypotheses H1 (at least one educational input), H2 (good classroom condition), H3 (functional desk condition), H5 (textbook availability), but did not merit the rejection of hypotheses H4 (toilets), H6 (exercise book availability), H7 (corner libraries), H8 (student sex), and H9 (cross-level interactions) across all models. Lastly, a comparison of the models indicated that M4, a random-covariance model with seven fixed parameters and one random parameter, has the best model fit. While the Bayesian's Information Criterion initially pointed to M2, a deeper engagement with statistical literature, accompanied with the results of the likelihood ratio tests and Akaike's Information Criterion, suggest that the value and estimation difference support the selection of M4 as the best model as opposed to M2. With a 95% confidence interval, M2 suggests that classroom condition, desk condition, and textbook availability positively predict higher academic achievement, while corner library/ies and student age negatively predict higher category. The presence of toilet/s, exercise book availability, and student biological sex, while non-significant, are important parameters to include in the model.

7. Discussion & Limitations

Discussion

Using nationally representative data from the World Bank, this study attempted to examine three pertinent questions within the context of 2,976 Grade 4 students across 97 randomly selected schools in Nigeria: 1) do educational inputs predict the log odds of performing at a higher category of academic achievement? 2) does the biological sex of students predict the log odds of performing at a higher category of academic achievement? 3) do educational inputs significantly moderate the association between student biological sex and academic achievement? First and foremost, this study recognizes the need for caution when discussing the implications of these findings given the absence of a causal direction as, for instance, high-performing schools may receive more adequate educational inputs and vice versa. Bearing this in mind, this section proceeds to discuss these three research questions in light of the two-level ordinal logistic models and relevant educational literature.

In response to the three research questions, the five multilevel models investigating the association between educational inputs and academic achievement yielded mixed findings with a noticeable skew towards an overall positive association. From the lens of M4 (the model with the best fit), this study specifically found that students exposed to schools with better classroom conditions, better desk conditions, and more textbooks were more likely to perform academically higher, whereas students with at least one classroom corner library were more likely to perform academically poorer. Whereas toilet/s and exercise book availability were non-significant, these two parameters increased model fit. In response to the second research question, student biological sex did not predict higher academic achievement; yet, the inclusion of this random parameter improved model fit. In response to the third research question, cross-level interactions revealed that educational inputs did not moderate the association between student biological sex and higher academic achievement, but, these interactions did not increase model fit.

Situating these results within academic discourse, it becomes clear that this study aligns with much of the literature arguing that the data pattern regarding educational inputs and academic achievement tends to be more consistent with that of positive relationship in the context of nations in the "Global South" than that of an adverse or non-significant relationship commonly found in the "Global North" (Hedges et al., 1994; Glass & Smith, 1979; Hedges & Stock, 1983; Hanushek,

1996). This view conflicts with literature, often stemming from researchers, policy makers, and multinational organizations in the "Global North," deeming such inputs to be inconsequential to student achievement (Coleman et al., 1996; Glewwe et al., 2011; and Sabarwal et al., 2014). Rather than dismissing educational inputs from the discourse on the learning crisis, one can surely adopt a dual stance on this contentious matter. The mixed findings may indicate that certain educational inputs are valuable in predicting better performance in low-income nations, but also it can mean that the influence of these inputs on student achievement may be conditional on alternative components of the educational system. This duality intertwines with a burgeoning body of literature suggesting that the predictive power of educational inputs can be highly influenced by not only the metrics selected, but also by contextual factors such as instruction styles, reform processes, and political corruption, among others. In addition to this overarching association, the non-significance of student biological sex aligns with the limited number of studies on Nigeria (Jegede, 1994), but contradicts those claiming a significant association (Erinosho, 1999; Agbaje & Alake, 2014). This result further clashes with the wider global discourse in educational research that often views student sex as a baseline indicator of achievement. The absence of a consensus may, once again, concern the overrepresentation of secondary studies, the mismatch in the learning assessment subject, and temporal differences in publications. Given the scarcity of relevant literature on cross-level interaction effects in Nigeria, the non-significant finding introduces baseline reference for future scholars to expand on.

There are various reasons as to why these research findings emerged, of which the following are a few examples. For instance, the positive predictive power of classroom condition, desk condition, and textbook availability may not directly drive achievement, but instead, may reflect the socio-economic status of students, parents, and communities. Due to the north-south divide in Nigeria, students in the wealthier south may be more prone to attending adequately resourced schools that receive financial support and administrative supervision from PTAs and local states. Moreover, the positive predictive power of these three inputs may denote the absence of conflict. Given the concentration of conflict in the north, schools in the south may also experience less Boko Haram attacks, and, thus, many not suffer from destroyed classrooms, broken desks, and insufficient textbooks as well as interrupted learning more generally. Concerning the slight adverse predictive power of corner library/ies, it may be the case that dimensions concerning the language, accessibility, and academic appropriateness of books located in such libraries may

hamper student learning. These dimensions are of paramount importance given the multi-lingual context of Nigeria.

In addition to these positive predictors, the non-significance of exercise book availability may be attributed to an emphasis on measuring input quantity as opposed to input quality. With the extensive prevalence of exercise books among primary students in Nigeria, it may be that using the indicator percentage of exercise books available was not as valuable in measuring this educational input as initially expected. Dimensions that capture student interaction with exercise books such as the frequency of use, the type of content written, and the alignment of teacher-student language may offer more compelling ways to measure this predictor. In tandem with exercise books, the non-significance of toilet/s may imply that the binary way of measuring toilets was not as helpful in understanding this educational input as originally thought, whereby, for instance, schools without separate toilets may not be regularly used by female students. An investigation into toilet types may be a more insightful way to measure this predictor.

In tandem with these six educational inputs, plausible reasons for the non-significance of student sex may relate to the exclusion of out-of-school children from this sample as they disproportionately are composed of female students in northern Nigeria. Given that the cross-level interactions explored moderation effects between student biological sex and achievement, the non-significance of the interactions may be due to the non-significance of student biological sex. Ultimately, all these plausible reasons illustrate that this study does not completely capture primary learning, but it does signal avenues for future research which will be mentioned at the end of this section.

With regards to the input-achievement nexus from the perspective of the two educational input clusters, the results yielded mixed findings in each cluster; mirroring the overarching pattern earlier established. Within the physical facilities cluster, the mixed findings of this study echoes the empirical conclusions of Earthman (2002), Ogunsaju, (1980), Nwagwu (1978), but the non-significant finding of toilet/s implies that this study slightly contradicts the positive correlations found in the publications of Onolemhenmhen (2014), Koroye (2016), and Edwards (2006). Among those conducted in Nigeria, a reason for this contradiction may be the focus on secondary students as opposed to primary students. Another reason may be attributed to temporal differences between previous publications and this current study as conditions and materials may have improved in certain areas. In parallel with the physical facilities cluster, the prevalence of mixed findings in the

educational materials cluster conflicted with the majority of the literature, whereby scholars argue that educational materials tend to positively predict academic achievement in the “Global South” (Jamison et al., 1981, Fuller, 1987, and Snilstveit et al., 2017). A reason for this clash may be attributed to the wealth of studies focused on textbook availability. Despite these mixed findings, internal reliability tests revealed that there was no statistical basis for the clustering of these inputs.

This study contributes to the literature beyond a reaffirmation or diversion from former educational research in three main ways: geographically, methodologically, and conceptually. First, the investigation of this association within Nigeria’s understudied educational landscape allows this study to shed light on a noticeable literature gap. Secondly, the use of multilevel ordinal logistic regression was a novel method that internalized realistic dimensions of society such as the nested structures in schooling. Thirdly, this study conceptually repositions the growing narrative often originating from scholars and educational policy circles in the “Global North" from one that deems educational inputs to be fruitless to one that dually distinguish between predictive associations and policy implications. Therefore, in tandem with an alignment with a substantial number of empirical and meta-analytical literature, this study geographically, methodologically, and conceptually contributes to the creation of a more inclusive, multidimensional, and nuanced conversation.

Accompanying these academic contributions, this study further informs the educational policy landscape in Nigeria by suggesting the following policy implications. Derived from multilevel analysis of this study, the Nigerian government must take into account the nested structures of primary schooling when designing federal policies. In addition, policy makers must be aware that while educational inputs may not directly or sufficiently enhance student achievement, they may still be critical in retaining students and enhancing the learning environment. Even beyond the realm of learning, there may be a moral argument in improving school resources especially toilet presence and functional desks as they can have positive health and societal outcomes. Moreover, this study hints at the need to address the learning crisis beyond school-based input-policies but through a systems lens in order to trace the deployment, management, quality, and availability of resources and the academic achievement of students. A notable example is the Research on Improving Systems of Education (RISE) program, which has been conducting research to highlight the severity and scale of the learning crisis using a systems thinking approach in Nigeria (RISE, 2020). Furthermore, future learning assessments implemented

with the collaboration of international development partners must include socio-economic indicators to eliminate the prevalence of spurious variables predicting both educational inputs and academic achievement.

Validity & Limitations

To adequately ascertain the credibility and applicability of the results gathered, this section will discuss the internal validity[8], external validity[9], and general limitations of this study.

Random selection, control variables, low attrition, and a substantial sample size increased confidence in the internal validity of this study. In using random selection at the school and student levels, this study limited the risk of systematic selection bias that otherwise may have distorted the predictive estimates of individual inputs. Moreover, the use of control variables enabled this study to account for confounding variables that may have interfered with the estimates. In addition to random selection and control variables, the low attrition rate of 0.03%, the percentage of students who did not consent to the study, in the sample further strengthens the internal validity of the findings. Lastly, the large sample size of 97 schools representing level-2 observations, which exceeded the minimum requirement of 10 observations for a random-intercept model and 30 observations for a random-covariate model as recommended by the two multilevel model specialists Austin (2010) and Moineddin et al. (2007), respectively.

Whereas random selection, control variables, low attrition, and a substantial sample size increased confidence in the internal validity of this study, the pursuit of correlational research, the partial violation of the proportional odds assumption, and the absence of classroom clustering were potential threats. The first threat concerns the limited interpretative power of pursuing non-experimental research. The non-experimental nature of this research does not fundamentally weaken internal validity given that no causal claims were made throughout the interpretation of the results; nevertheless, non-experimental research prevents the attribution of a direction to the predictors. The second threat relates to the violation of the proportional odds assumptions by toilet/s and textbook availability, implying that the estimates of these two inputs cannot be generalized across all cumulative achievement splits in the data. This assumption has been highly

[8] Internal validity refers to the extent to which a particular study can confidently claim that the tested causal relationship has not been influenced by alternative confounding variables.

[9] External validity refers to applicability or the extent to which the findings of a particular study can be applied to real situations, populations, or events.

contested within literature with scholars deeming it as anti-conservative for continuous variables and larger sample sizes (Brant, 1990; Allison, 1999). Accordingly, one may argue that this violation may not jeopardize the validity of the model given the continuous nature of textbooks and the large sample size found for exercise books. In addition to these two threats, the absence of classroom clustering in pursuing a two-level model instead of a three-level model may have distorted estimates. But, to avoid compromising the large sample size, the elimination of classroom clustering was a necessary modification.

Random selection, weighted data, and inclusion-exclusion criteria increased the external validity of this study. Random selection had the dual effect of increasing internal and external validity, thereby, fostering more credible and applicable results. Specifically, randomly sampling a stratum of schools and students allowed the study to generalize findings onto the population of Grade 4 students in Nigeria. Alongside random selection, the use of weighted data enabled estimates to be more representative of the population as private schools were initially heavily sampled for alternative research pursuits. In addition to random selection and weighted data, inclusion-exclusion criteria enable this study to clearly demarcate a cross-section of Grade 4 students, thereby, increasing generalizability of this study.

Whereas random selection, weighted data, and inclusion-exclusion criteria increased external validity, instrument composition and logit model restrictions were threats. The first threat pertains to the composition of learning instruments using content from the Nigerian formal curriculum and international assessments impedes this study from generalizing findings onto other countries. This novel composition, however, has immense merit in countering the deterministic instruments often used to evaluate achievement and, ultimately, does not hamper the primary objective of this study to investigate the input-achievement nexus in this understudied context. Compounding onto instrument composition, model restrictions could be construed as a threat to external validity. For instance, the need to convert academic achievement from a continuous to an ordinal variable to conduct a logit model prevented the study from comparing results with other metrics of achievement. Without this conversion, however, this study would not have been able to internalize the nested data structure given assumption violations in the hierarchical linear model. Another restriction pertains to the logit model itself as many scholars opt out of using this approach due to the need for large level-2 observations and the limitations on the number of variables, both

of which prevented this study from including school budgets, school types, and school location as random level-2 parameters.

Ultimately, when considering the structural and systematic factors used to enhance internal and external validity, the corresponding validity threats did not appear to have significantly undermined the credibility and applicability of these findings. In tandem with the methodological limitations mentioned in the form of internal and external validity threats, it is imperative to note the pressing non-methodological limitations embedded in this study. Specific emphasis will be placed on discussing limitations related to the conceptualization and operationalization of academic achievement, implications of an evolving educational landscape, and the perceived reductionist association between inputs and achievement.

Firstly, the conceptualization and operationalization of academic achievement can be construed as a limitation. Concerning the conceptualization of achievement, the use of colonial language and content in these learning assessments may fail to measure or underestimate the cognitive form of this latent construct when disregarding indigenous languages and knowledge forms. This matter is of particular importance given the diverse multi-ethnic composition of communities in Nigeria. The adoption of colonial languages and content in formal schooling over decades may, alternatively, warrant the need to include them in any educational assessment as it echoes the state-driven reality of the learning ecosystem. Regarding the operationalization of achievement, the ordinal collation of all three subjects may not clearly lend itself to direct comparisons with subject-specific, national or international achievement instruments commonly found in the literature. This collation does, on the other hand, challenge normative, orientalist and deterministic modes of testing student performance. It may, for instance, garner a more representative image of the academic schedules of most primary students who are prone to learning more than one subject at any given year. Therefore, while one may consider the operationalization of achievement to have limited capacity to be applied comparatively, one may also understand that this variable can challenge present assessment modes and more accurately capture the academic reality.

Secondly, the constantly evolving nature of educational inputs and student performance implies that the interpretation of these findings and the consequent implementation of educational policies can be a highly time-sensitive issue. The stagnant state of achievement in Nigeria may suggest that these findings will be relevant for a longer time than expected. Thirdly, the

investigation of the input-achievement nexus without the inclusion of influential achievement predictors (e.g. student motivation, parental socio-economic status) or systems thinking dynamics (e.g. school-government interactions, PTA-student interactions) can produce reductionist depictions of the learning process. An expansion of research design, however, can often compromise statistical reliability. The process of balancing depth and breadth in quantitative educational research is a common challenge. Overall, these three limitations are far from comprehensive, but in identifying them, this study can clearly demarcate future avenues for research.

There are various future research directions. Within the confines of this data set, it may be worthwhile to examine the input-achievement nexus within the secondary student sample in order to compare findings with their primary counterparts. Researchers may also explore a broader set of educational inputs such as toilet types, school budgets, lunch meals etc. Beyond the confines of this data set, a qualitative study that uses ethnographic research to explore how schools leaders, teachers, students, and government officials in a given state manage, maintain, and employ resources for the purpose of learning can be an insightful way to shed light on the human dimension of the nexus. A mixed methods study that triangulates the findings of ethnography and multilevel modeling can foster a more multidimensional understanding that counters the reductionistic critique often associated with quantitative methods. Using a systems thinking framework, this mixed methods study can further identify inefficiencies in resources distribution, allocation, and use as well as systematic problems in the administration, instruction, and assessment of curriculum. Additionally, a panel study could also help address the time-sensitive nature of results interpretations and policy implementations.

8. Conclusion

Through the use of nationally representative data from the World Bank, this study sought to determine whether six educational inputs predict higher academic achievement in English, Mathematics, and Science among 3,262 Grade 4 students across 97 randomly selected schools in Nigeria. The six educational inputs were theoretically and proportionally classified into two input clusters: physical facilities and educational materials. In reference to the literature, the three inputs –classroom condition, desk condition, presence of toilet/s– were categorized within the physical facilities cluster. In parallel, the three inputs –textbook availability, exercise book availability, and presence of corner library/ies– were embedded within the educational materials cluster. Internal reliability tests found that this clustering had no empirical basis.

The analysis revealed mixed evidence concerning the association between educational inputs and academic achievement, with the results suggesting that the pattern of data tends to yield more positive than negative results. Specifically, good classroom conditions, non-broken desk conditions, and textbook availability were significant positive predictors of achievement, but exercise book availability and corner library/ies were significant negative predictors of achievement. Regardless of the direction, all significant parameters held moderate predictive power. Toilet/s yielded non-significant results. Student-sex also produced non-significant results. Educational inputs also did not moderate the association between student sex and achievement.

The skew towards a significant association suggests that, contrary to growing academic and policy discourse in international circles often stemming from the "Global North," educational inputs can predict achievement, an association more likely to occur in the "Global South." The varied direction and small effect of these educational inputs, however, are most certainly limiting factors that must be accounted for in this discussion. The prevalence of mixed findings simultaneously echoes the spectrum of views in the literature with scholars predominantly researching the "Global North" arguing for non-significant association, while others investigating this nexus in the "Global South" claim a slightly more positive association. Many scholars agree on a combination of both. This geographic divide may not be perfectly mapped in the literature, but a pattern over recent decades has been discerned. Considering this caveat, this study argues that, instead of negating the importance of educational inputs entirely, one may adopt a systems thinking approach to this perceived "learning crisis."

Aiming to foster a more inclusive academic discussion on what predicts student performance in Nigeria, this study geographically, methodologically, and conceptually contributes to the literature. Not only did this study address a literature gap within an understudied national context, but also it employed a statistical method accounting for the nested structure of formal schooling and, simultaneously, challenged the normative global discourse on education.

One must be cautious when applying these findings onto educational policy given the absence of a causal direction between inputs and achievement. Nevertheless, a few key insights relate to the importance of considering the nested structure of formal school in any evidence-based approach to policy making. Moreover, these results may indicate a greater need to focus on the quality of the educational learning environment. As opposed to exercise books which are more widely available, more funds should be administered towards ensuring each student has access to textbooks in light of the essential content students are often tested within them. The absence of significant findings or the slight negative findings of certain predictors does not necessarily suggest their irrelevance to learning. Rather, it may mean that they do not directly enhance student performance but serve as critical factors in creating a more favorable learning ecosystem that can attract students to enrolling and remaining in school.

To address conceptual limitations that criticize the reductionist portrayal of learning in this study, future avenues for research may pursue a mixed methods study triangulate insights from ethnography and experimental multilevel modelling to trace how various actors in the system navigate the distribution, management, maintenance, and use of resources for enhanced learning.

References

Abu, A., Jimoh, S., & Olaniyi, S. (2020, June 4). Outrage over cut in health, education budgets. *The Guardian (Nigeria)*. https://guardian.ng/news/outrage

Adeche-Apeji, E. (1990). The development of school library services. *International library review*, *22*(1), 41-51. https://doi.org/10.1016/0020-7837(90)90038-H

Adegboyega, A. (2019, May 15). More than half of Nigeria's education budget lost to corruption- Transparency International. *Premium Times Nigeria.* https://www.premiumtimesng.com/news/top-news/330104

Adetunji, A. E., Adeniran, K. A., Olomu, S. C., Odike, A. I., Ewah-Odiase, R. O., Omoike, I. U., & Akpede, G. O. (2019). Socio-demographic factors associated with overweight and obesity among primary school children in semi-urban areas of mid-western Nigeria. *PloS On*e, *14*(4), 1-12. https://doi.org/10.1371/journal.pone.0214570

Adewale, A. (2011). *The political, economic and social dynamics of Nigeria: A synopsis* (Asia Policy Brief, pp. 1-9, Issue brief No. 39). Africa Institute of South Africa. https://media.africaportal.org/documents/No-39.-The-Political-Economic-and-Social-.pdf

Adeyemi. T.O. (2010). The School Library and Students' Learning Outcomes in Secondary Schools in Ekiti State, Nigeria. *Asian Journal of Business Management*. *2*(1), 1-8. https://www.researchgate.net/publication/45266783

Agabi, C. O. (2010). Prudential approach to resource management in nigerian education: A theoretical perspective. *International Journal of Scientific Research in Education 3*(2), 94-106. http://ijsre.com/assets/vol.,-3_2_-agabi.pdf

Agbaje, R. O., & Alake, E. M. (2014). Students' variables as predictor of secondary school students' academic achievement in science subjects. *International Journal of Scientific and Research*, *4*(9), 1-5. https://citeseerx.ist.psu.edu

Agresti, A., & Natarajan, R. (2001). Modeling clustered ordered categorical data: A survey. *International Statistical Review*, *69*(3), 345-371. https://doi.org/10.1111/j.1751-5823.2001.tb00463.x

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716-723. http://doi.org/10.1109/TAC.1974.1100705

Allison, P. D. (1999). *Multiple regression: A primer*. Pine Forge Press.

Alonge, R., & Okunade, H. (2014). Ethnic Conflict and Sustainable Development in Nigeria: The Role of Peace Education. International Journal of Education and Literacy Studies, *2*(3), 22-25. http://dx.doi.org/10.7575/aiac.ijels.v.2n.3p.22

Ameh, J., & Aluko, O. (2019, January 04). 2019 budget: Education gets N620.5bn, against UNESCO's advice. *Punch.* https://punchng.com/2019

Amoo, A. (2019, December 26). 315,579 classrooms in bad condition across Nigeria. *Educeleb.* https://educeleb.com/classrooms-bad-condition-nigeria

Anero, N. (2014). Relevance and challenges of primary education to the overall development of the child and the Nigerian society. *Global Journal of Educational Research*, *13*(2), 55-62. https://doi.org/10.4314/gjedr.v13i2.1

Baro, E., & Eze, M. (2016). Enhancing quality learning: The impact of school library services to students in Nigeria. *22*(1), 8-19. https://www.researchgate.net/profile/Emmanuel_Baro

Asaaju, O. A. (2015). The inconsistency of Nigeria's education system and its implication for curriculum implementation. *Journal of US-China Public Administration*, *12*(3), 167-179. https://doi.org/10.17265/1548-6591/2015.03.001

Atkinson, A. C. (1981). Likelihood ratios, posterior odds and information criteria. *Journal of Econometrics, 16*(1), 15-20. https://doi.org/10.1016/0304-4076(81)90072-5

Austin, P. C. (2010). Estimating multilevel logistic regression models when the number of clusters is low: a comparison of different statistical software procedures. *The international journal of biostatistics, 6*(1). https://doi.org/10.2202/1557-4679.1195

Bertoni, E., Di Maio, M., Molini, V., & Nistico, R. (2019). Education is forbidden: The effect of the Boko Haram conflict on education in North-East Nigeria. *Journal of Development Economics, 141*. https://doi.org/10.1016/j.jdeveco.2018.06.007

Bloom, B. (1956). *Taxonomy of educational objectives : The classification of educational goals. Handbook 1, Cognitive domain*. London: Longmans.

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics, 46*(4), 1171-1178. https://doi.org/10.2307/2532457

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge - Mass. & London: Harvard University Press.

Brown, B. W. (1991). How gender and socioeconomic status affect reading and mathematics achievement. *Economics of education review, 10*(4), 343-357. https://doi.org/10.1016/0272-7757(91)90024-J

Byrd, C. M., & Chavous, T. M. (2009). Racial identity and academic achievement in the neighborhood context: A multilevel analysis. *Journal of Youth and Adolescence, 38*(4), 544-559. https://doi.org/10.1007/s10964-008-9381-9

Callaway, A., & Musone, A. (1968). *Financing of education in Nigeria*. Paris: UNESCO, International Institute for Educational Planning.

Cameron, S. (2015). *ESSPIN Composite Survey 2*. Oxford Policy Management. https://www.opml.co.uk/files/Publications/a0148

Castellucci, H., Arezes, P., Molenbroek, J., De Bruin, R., & Viviani, C. (2017). The influence of school furniture on students' performance and physical responses: Results of a systematic review. Ergonomics, 60(1), 93-110. https://doi.org/10.1080/00140139.2016.1170889

Center for Universal Education at Brookings. (2012). *Africa Learning Barometer: Technical Appendix*. https://www.brookings.edu/wp-content/

Clogg, C. C., & Shihadeh, E. S. (1994). *Statistical Models for Ordinal Variables* (Advanced Quantitative Techniques in the Social Sciences). Thousand Oaks, Calif: Sage. https://doi.org/10.1002/(SICI)1097-0258(19970228)

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., & Weinfeld, F. (1966). *Equality of educational opportunity study*. United States Department of Health, Education, and Welfare.

Cummins, J. (2015). Intercultural education and academic achievement: A framework for school-based policies in multilingual schools. *Intercultural Education: Honouring the Contributions of Jim Cummins to Language and Intercultural Education Internationally*, *26*(6), 455-468. https://doi.org/10.1080/14675986.2015.1103539

Davis, C. R., & Salkind, N. J. (2010). Homogeneity of variance. *Encyclopedia of Research Design. Thousand Oaks (US): SAGE Reference*, 577-80.

Demidenko, E. (2004). *Mixed models: Theory and Applications* (Wiley Series in Probability and Statistics). John Wiley & Sons. https://doi.org/10.1002/0471728438

Doornik, J. A., & Hansen, H. (2008). An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, *70*, 927-939. https://doi.org/10.1111/j.1468-0084.2008.00537.x

Dziak, J. J., Nahum-Shani, I., & Collins, L. M. (2012). Multilevel factorial experiments for developing behavioral interventions: Power, sample size, and resource considerations. *Psychological methods*, *17*(2), 153. https://doi.org/10.1037/a0026972

Earthman, G. I. (1985). Evaluating the Completed Project. *CEFP Journal*, *23*(3), 5-6.

Earthman, G. I. (2002). School facility conditions and student academic achievement. https://escholarship.org/uc/item/5sw56439

Earthman, G. I., & Lemaster, L. K. (1996). Review on the research on the relationship between school buildings, student achievement, and student behaviour. Paper presented in the annual meeting of the Council of Educational Facilities Planners. *International, Tarpon Springs, Florida, USA.* https://files.eric.ed.gov/fulltext/ED416666.pdf

Ebuk, L., & Bamijoko, O. (2016). The Effective Management of Mathematics Worktext: Sure Remedy to Students' Performance in Mathematics. *International Journal of Advanced Multidisciplinary Research, 3*(5). http://s-o-i.org/1.15/ijarm-2016-3-4-9

Edwards, N. C. (2006). *School facilities and student achievement: Student perspectives on the connection between the urban learning environment and student motivation and performance* [Doctoral dissertation, Ohio State University]. https://etd.ohiolink.edu/

Egbas, J. (2018, December 13). "13.2 million Nigerian children now out of school, highest in the world," UNICEF says. *Pulse*. https://www.pulse.ng/news/local/132

Ekpo, K., Akpan, O., Essien, E., & Imo-Obot, M. (2009). Classroom Climate and Students' Academic Achievement in Social Studies in Cross River, Nigeria. *African Research Review*, 3(4), 412-428. https://doi.org/10.4314/afrrev.v3i4.47576

Erinosho, S. Y. (1999). Gender Differences in Interests and Performance of High School Science Students: A Nigerian Example. *Studies in Educational Evaluation*, *25*(2), 163-71. https://doi.org/10.1016/S0191-491X(99)00019-X

Falayajo, W., Makoju, G. A. E., Okebukola, P., Onugha, D. C., & Olubodun, J. O. (1997). Monitoring of learning achievement project, assessment of learning achievement of primary four pupils in Nigeria. Federal Government of Nigeria/UNICEF.

Fielding, A., Yang, M., & Goldstein, H. (2003). Multilevel ordinal models for examination grades. *Statistical modelling, 3*(2), 127-153. https://doi.org/10.1191/1471082X03st052oa

Fredriksen, B., & Brar, S. (2015). *Getting textbooks to every child in Sub-Saharan Africa: Strategies for addressing the high cost and low availability problem*. The World Bank. https://doi.org/10.1596/978-1-4648-0540-0

Fuller, B. (1987). What School Factors Raise Achievement in the Third World? *Review of Educational Research, 57*(3), 255-292. https://doi.org/10.3102/00346543057003255

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.

Gilavand, A., Gilavand, M., & Gilavand, S. (2016). Investigating the impact of lighting educational spaces on learning and academic achievement of elementary students. *International Journal of Pediatrics, 4*(5), 1819-1828. https://doi.org/10.3102/00346543064001119

Glass, G. V., & Smith, M. L. (1979). Meta-analysis of research on the relationship of class-size and achievement: The class size and instruction project. *Educational Evaluation and Policy Analysis, 1*(1), 2-16. https://doi.org/10.3102/01623737001001002

Glewwe, P., Kremer, M., & Moulin, S. (2009). Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics*, *1*(1), 112-35. https://doi.org/10.1257/app.1.1.112

Glewwe, P. W., Hanushek, E. A., Humpage, S. D., & Ravina, R. (2011). *School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010* (No. w17554). National Bureau of Economic Research. https://www.nber.org/papers/w17554.pdf

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. *Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, 82-88.* http://hdl.handle.net/1805/344

Gove, A., & Cvelich, P. (2011). *Early reading: Igniting education for all.* Early Grade Learning Community of Practice, 1-66. https://eric.ed.gov/?id=ED520290

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of educational research*, *66*(3), 361-396. https://doi.org/10.3102/00346543066003361

Hanushek, E. (1996). A More Complete Picture of School Resource Policies. Review of Educational Research, *66*(3), 397-409. https://doi.org/10.3102/00346543066003397

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis* (Vol. 451). John Wiley & Sons.

Hedges, L., Laine, R., & Greenwald, R. (1994). Money Does Matter Somewhere: A Reply to Hanushek. Educational Researcher, *23*(4), 9-10. https://doi.org/10.3102/0013189X023004009

Hedges, L., & Stock, W. (1983). The Effects of Class Size: An Examination of Rival Hypotheses. American Educational Research Journal, *20*(1), 63-85. https://doi.org/10.3102/00028312020001063

Heneveld, W., & Craig, H. (1996). *Schools count: World Bank project designs and the quality of primary education in Sub-Saharan Africa*. The World Bank.

Henze, N., & Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Communications in statistics-Theory and Methods*, *19*(10), 3595-3617. https://doi.org/10.1080/03610929008830400

Heyneman, S. P., & Jamison, D. T. (1980). Student learning in Uganda: Textbook availability and other factors. *Comparative education review*, *24*(2), 206-220. https://doi.org/10.1086/446116

Heyneman, S. P., & Loxley, W. A. (1983). The effect of primary-school quality on academic achievement across twenty-nine high-and low-income countries. American Journal of Sociology, *88*(6), 1162-1194. https://doi.org/10.1086/227799

Hossain, N., & Hickey, S. (2019). The Problem of Education Quality in Developing Countries. In The Politics of Education in Developing Countries. Oxford University Press.

Hughes, S. M., & Lin, Y. C. (2018). Survey Data Collection in Sub-Saharan Africa (SSA) Challenges, Strategies, and Opportunities. *Advances in Comparative Survey Methods: Multinational, Multiregional, and Multicultural Contexts (3MC)*, 533-554. https://doi.org/10.1002/9781118884997.ch25

Idoko, C. (2020, January 08). 50% of public primary schools in Nigeria have no toilets -UBEC: Education. Nigerian Tribune. https://tribuneonlineng.com/50-of-public-primary-schools-in-nigeria-have-no-toilets-ubec/

Jaiyeoba, A. O., & Atanda, A. I. (2011, June). School quality factors and secondary school students' achievement in mathematics in South-western and North-central Nigeria. *The African Symposium*, *11*(1), 91-100. https://pdfs.semanticscholar.org/ec9d/

Jamison, D. T., Searle, B., Galda, K., & Heyneman, S. P. (1981). Improving elementary mathematics education in Nicaragua: An experimental study of the impact of textbooks and radio on achievement. *Journal of Educational Psychology*, *73*(4), 556. https://doi.org/10.1037/0022-0663.73.4.556

Jegede, O. (1994). African Cultural Perspectives and the Teaching of Science. In J. Solomon & G. Aikenhead (Eds.), *STS Education: International Perspectives on Reform (Ways of Knowing in Science Series)* (pp. 120-130). Teachers College Press.

Jencks, C. (1972). Inequality: A reassessment of the effect of family and schooling in America. Basic Books.

Jung, L., & Guskey, T. (2012). *Grading exceptional and struggling learners*. Corwin Press.

Konstantopoulos, S., & Borman, G. (2011). Family Background and School Effects on Student Achievement: A Multilevel Analysis of the Coleman Data. *Teachers College Record (1970), 113*(1), 97-132. https://www.researchgate.net/publication/286614398

Koroye, T. (2016). The influence of school physical environment on secondary school students academic performance in Bayelsa State. *Asian Journal of Educational Research*, *4*(2), 1-15. https://pdfs.semanticscholar.org/e1ea/

Krathwohl, D., Bloom, B., & Masia, B. (1964). *Taxonomy of educational objectives : The classification of educational goals. Handbook 2, Affective domain*. Longmans.

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological methods & research*, *33*(2), 188-229. https://doi.org/10.1177/0049124103262065

Lee, V., & Zuze, T. (2011). School Resources and Academic Performance in Sub-Saharan Africa. Comparative Education Review, *55*(3), 369-397. https://doi.org/10.1086/660157

Liew, J., Cameron, & Lockman. (2016). Early Education and Development Special Issue. Early Education and Development, *27*(7), 1101-1102. http://dx.doi.org/10.1080/10409289.2016.1216643

Liu, X. (2015). *Applied ordinal logistic regression using Stata: From single-level to multilevel modeling*. Sage Publications.

Lockheed, M., & Komenan, A. (1989). Teaching quality and student achievement in Africa: The case of Nigeria and Swaziland. *Teaching and Teacher Education, 5*(2), 93-113. https://doi.org/10.1016/0742-051X(89)90009-7

Lockheed, M., Vail, S., & Fuller, B. (1986). How textbooks affect achievement in developing countries: Evidence from Thailand. *Educational Evaluation and Policy Analysis, 8*(4), 379-392. 10.3102/01623737008004379

Malá, D., & Čerešník, M. (2015). The Evaluation of the Course Description Quality by Students of the Psychology Teaching Training Programme. *Procedia, Social and Behavioral Sciences, 174*, 3523-3530. https://doi.org/10.1016/j.sbspro.2015.01.1067

Mansfield, E. R., & Helms, B. P. (1982). Detecting multicollinearity. *The American Statistician*, *36*(3a), 158-160. https://doi.org/10.1080/00031305.1982.10482818

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519-530. https://doi.org/10.1093/biomet/57.3.519

Maslow, A. (1943). A theory of human motivation. *Psychological Review*, *50*(4), 370-396. https://doi.org/10.1037/h0054346

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86-92. https://doi.org/10.1027/1614-2241.1.3.86

Mbah, A. K., & Paothong, A. (2015). Shapiro–Francia test compared to other normality test using expected p-value. *Journal of Statistical Computation and Simulation*, *85*(15), 3002-3016. https://doi.org/10.1080/00949655.2014.947986

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review, 28*(2), 295-314. https://doi.org/10.1007/s10648-014-9287-x

Mingat, A., Tan, J., & Sosale, S. (2003). *Tools for education policy analysis.* The World Bank. https://www.researchgate.net/publication/5086319_Tools_for_education_policy_analysis

Moineddin, R., Matheson, F. I., & Glazier, R. H. (2007). A simulation study of sample size for multilevel logistic regression models. *BMC medical research methodology*, *7*(1), 1-10. https://doi.org/10.1186/1471-2288-7-34

Molenberghs, G., & Verbeke, G. (2000). *Linear mixed models for longitudinal data*. Springer Publishing.

Momoh, S. (1980). A study of the relationship between instructional resources and academic achievement of students in Ilorin local government Kwara State [Unpublished M.Ed Thesis]

Murnane, R., & Ganimian, A. (2014). Improving educational outcomes in developing countries: Lessons from rigorous evaluations. *NBER Working Paper Series.* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2471177

Mwamwenda, T. S., & Mwamwenda, B. B. (1987). School facilities and pupils' academic achievement. *Comparative Education*, *23*(2), 225-235. https://doi.org/10.1080/0305006870230208

National Policy on Education (4th Edition), 1 (Nigerian Educational Research and Development Council 2004). http://wbgfiles.worldbank.org/documents/hdn/ed/saber/supporting_doc

National Population Commission. (2015). *Nigeria Education Data Survey*. https://pdf.usaid.gov/pdf_docs/pa00kzpv.pdf

Nevo, C. M., & Egenti, S. (2019). A disaggregated analysis of wealth status and educational attainment in Nigeria using the multinomial logit approach. *Economies, 7*(2), 1-9. https://doi.org/10.3390/economies7020039

Nordea. (2020). The economic context of Nigeria. https://www.nordeatrade.com/en/explore-new-market/nigeria/economical-context

Norton, J. (1926). Education and the Federal Government [Doctoral Dissertation, Columbia University].

Nwagwu, N. (1978). *Primary school administration*. Macmillan Nigeria.

Onolemhenmhen, P. E. (2014). Effectiveness of good classroom conditions on the academic achievement of secondary school students in Edo Central Senatorial District. *Asian Journal of Education and e-Learning*, *2*(2), 146-149. https://python.zzx.us/index.php/AJEEL/article/view/720

OConnell, A. A. (2006). *Logistic regression models for ordinal response variables*. Sage Publications.

O'Connell, A. A. (2010). Multilevel Models for Discrete Response Data. *Wiley Series in Probability and Statistics Multilevel Statistical Models,* 111-146. https://doi.org/10.1002/9780470973394.ch4

Ogbonna, S. (2016). *Reflections on why basic education learning outcomes are declining in Nigeria*. Education Sector Support Programme in Nigeria.

Ogunsaju, O. A. (1980). An analysis of procedures and methods of educational planning utilized in Ogun State, Nigeria [Doctoral Dissertation, Loyola University Chicago]. https://ecommons.luc.edu/luc_diss/1921

Okere, L., & Mnim, M. (2016). Textbook crisis and the challenges for sustainable educational quality in nigeria. *International Journal of Innovative Education Research*, *4*(2), 40-45. https://seahipaj.org/journals-ci/june-2016/IJIER/full/IJIER-J-5-2016.pdf

Omolayole, O. O. (2001). Revitalizing libraries in Nigeria. *Lokoja: A Memorandum Submitted at the Joint Consultative Council on Education (JCCE) held in Ilorin, Kwara State, Nigeria*, 1-10.

Omotor, D. (2009). Determinants of federal government health expenditures in Nigeria. *International Journal of Economic Perspectives*, *3*(1), 5-18. https://www.researchgate.net/publication/246044767

Onwuameze, N. C. (2013). Educational opportunity and inequality in Nigeria: Assessing social background, gender, and regional effects [Doctoral Dissertation, The University of Iowa]. https://doi.org/10.17077/etd.bs85au87

Onyemelukwe-Waziri, H. (2017). Impact of two wars on the educational system in Nigeria [Master's Thesis, Clark University].

Peterson, B., & Harrell Jr, F. E. (1990). Partial proportional odds models for ordinal response variables. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *39*(2), 205-217. https://doi.org/10.2307/2347760

Piper, B. (2017). Improving learning in sub-Saharan Africa using rigorous research designs. *Journal of Research on Educational Effectiveness*, *10*(3), 535-540. https://doi.org/10.1080/19345747.2017.1335093

Princewill, O., Lucas, O., Daniel, I. (2020). Colonialism and amalgamation of southern and northern protectorates: Analysis of emerging issues in Nigeria. *Journal of Political Science and Leadership Research*, *6*(1), 49-64. https://www.researchgate.net/publication/336901133

Purwaningrum, L., Funatsu, K., Xiong, J., Rosyidi, C., Muraki, S., & Dalal, K. (2015). Effect of Furniture Weight on Carrying, Lifting, and Turning of Chairs and Desks among Elementary School Children. PloS One, *10*(6), 1-19. https://doi.org/10.1371/journal.pone.0128843

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata*. STATA Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage Publications.

RISE. (2020). The learning crisis. https://riseprogramme.org/systems-thinking/learning-crisis

Roscigno, V. J., Tomaskovic-Devey, D., & Crowley, M. (2006). Education and the inequalities of place. *Social Forces*, *84*(4), 2121-2145. https://doi.org/10.1353/sof.2006.0108

Royston, J. P. (1983). Some techniques for assessing multivarate normality based on the Shapiro-Wilk W. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *32*(2), 121-133. https://doi.org/10.2307/2347291

Ryan, S. E., P. J. Rigby, and K. Campbell. (2010). Randomised controlled trial comparing two school furniture configurations in the printing performance of young children with cerebral palsy. *Australian Occupational Therapy Journal*, *57*(4), 239–245. https://doi.org/10.1111/j.1440-1630.2009.00831.x

Sabarwal, S., Evans, D. K., & Marshak, A. (2014). The permanent input hypothesis: The case of textbooks and (no) student learning in Sierra Leone. The World Bank. https://documents.worldbank.org/806291468299200683

Salau, G. (2017, August 27). Where desks, qualified teachers are unaffordable luxuries. *The Guardian (Nigeria)*. https://www.pressreader.com/nigeria/the-guardian-nigeria

Salau, H. M., Alkassim, A., & Iloani, U. (2020, June 05). Nigeria: Revised budget - Anger over cuts in health, education votes. *All Africa*. https://allafrica.com/stories/202006040446.html

Salo, B., Laaksonen, T., & Santtila, P. (2016). Construct validity and internal reliability of the Finnish risk and needs assessment form. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, *17*(1), 86-107. https://doi.org/10.1080/14043858.2016.1161940

Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical methods in Medical Research*, *8*(1), 3-15. https://doi.org/10.1177/096228029900800102

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, *12*(2), 347-367. https://doi.org/10.1177/1094428107308906

Schilling, L., Washington, K., Billingsley, F. & Deitz, J. (2003). Classroom seating for children with attention deficit hyperactivity disorder: Therapy balls versus chairs." *American Journal of Occupational Therapy*, *57*(5), 534–541. https://doi.org/10.5014/ajot.57.5.534

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461-464. https://projecteuclid.org/euclid.aos/1176344136

Searle, S.R., Casella, G. and McCulloch, C.E. (1992), *Variance components.* John Wiley & Sons. http://dx.doi.org/10.1002/9780470316856

Shapiro, S. S., & Francia, R. S. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, *67*(337), 215-216. https://doi.org/10.1080/01621459.1972.10481232

Simpson, E. J. (1966). The classification of educational objectives, psychomotor domain, 1-45. https://eric.ed.gov/?id=ED010368

Smith-Zuzovsky, N., & Exner, C. (2004). The effect of seated positioning quality on typical 6- and 7-year-old children's object manipulation skills. *American Journal of Occupational Therapy*, *58*(4), 380–388. https://doi.org/10.5014/ajot.58.4.380

Snijders, T. A., & Bosker, R. J. (1999). *An introduction to basic and advanced multilevel modeling.* Sage Publications.

Steinmayr, R., Meissner, A., Weidinger, A. F., & Wirthwein, L. (2015). Academic Achievement. Oxford Bibliographies. https://doi.org/10.1093/obo/9780199756810-0108

Snilstveit, B., Gallagher, E., Phillips, D., Vojtkova, M., Eyers, J., Skaldiou, D., Davies, P. (2017). PROTOCOL: Interventions for improving learning outcomes and access to education in

low- and middle-income countries: A systematic review. *Campbell Systematic Reviews*, *13*(1), 1-82. https://doi.org/10.1002/CL2.176

State Universal Basic Education Board. (2017). Lagos, SUBEB. https://subeb.lagosstate.gov.ng/

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *19*(3), 227-229. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938757/

Tamura, K. A., & Giampaoli, V. (2010). Prediction in multilevel logistic regression. *Communications in Statistics-Simulation and Computation*, *39*(6), 1083-1096. https://doi.org/10.1080/03610911003790106

Tenenbaum, G., & Filho, E. (2016). Measurement considerations in performance psychology. *Performance Psychology*, 31-44. https://doi.org/10.1016/b978-0-12-803377-7.00003-x

Thompson, M. J., & O'brien, T. P. (1991). Learning styles and achievement in postsecondary classrooms. Paper presented at the Annual Conference of the American Educational Research Association, 3-18.

Tripathy, J. P. (2013). Secondary data analysis: Ethical issues and challenges. *Iranian Journal of Public Health*, *42*(12), 1478-1479. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4441947/

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2013). The global learning crisis: Why every child deserves a quality education. https://unesdoc.unesco.org/ark:/48223/pf0000223826

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2016). Every child should have a textbook. https://unesdoc.unesco.org/ark:/48223/pf0000243321

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2018). School enrollment, primary (% net). https://data.worldbank.org/indicator/SE.PRM.NENR

Unity, O., & Igbudu, U. (2015). Influence of gender on students' academic achievement in government subject in public secondary schools in Oredo Local Government area of Edo State, Nigeria. *Journal of Educational and Social Research*, *5*(2), 101-106. https://doi.org/10.5901/jesr.2015.v5n2p101

Universal Basic Education Commission: Home. (2020). https://www.ubec.gov.ng/about/who-we-are/

Vecchione, M., Alessandri, G., & Marsicano, G. (2014). Academic motivation predicts educational attainment: Does gender make a difference?. *Learning and Individual Differences*, *32*, 124-131. https://doi.org/10.1016/j.lindif.2014.01.003

Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, *17*(2), 228. https://doi.org/10.1037/a0027127

Fractor, J., Woodruff, M., Martinez, M., & Teale, W. (1993). Let's not miss opportunities to promote voluntary reading: Classroom libraries in the elementary school. *The Reading Teacher*, *46*(6), 476-484. http://www.jstor.org/stable/20201114

Walberg, H. J. (1991). Improving school science in advanced and developing countries. Review of educational research, *61*(1), 25-69. https://doi.org/10.3102/00346543061001025

Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and Instruction*, *17*(4), 416-426. https://doi.org/10.1016/j.learninstruc.2007.03.007

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, *30*(1), 1-35. https://doi.org/10.1016/j.dr.2009.12.001

Wilson, M. (1989). A comparison of deterministic and probabilistic approaches to measuring learning structures. *Australian Journal of Education*, *33*(2), 127-140. https://doi.org/10.1177/168781408903300203

World Bank. (2020). Better Education Service Delivery for All. https://projects.worldbank.org/en/projects-operations/project-detail/P160430

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2015). Education 2030: Incheon Declaration and Framework for Action. https://unesdoc.unesco.org/ark:/48223/pf0000245656

Yu, H., Jiang, S., & Land, K. C. (2015). Multicollinearity in hierarchical linear models. *Social Science Research, 53,* 118-136. https://doi.org/10.1016/j.ssresearch.2015.04.008

Zacharkow, D. (1987). *Posture: Sitting, standing, chair design & exercise*. Charles C Thomas Publisher, LTD.

Appendix A: Descriptive Analysis

This section reports the descriptive statistics of primary schools by ownership type and location (Table 2), primary school leaders by position (Table 3), Grade 4 students by age and sex (Table 4), student academic achievement by assessment type and threshold (Table 5), and educational inputs by cluster (Table 6). The use of descriptive statistics enables this study to empirically envisage the essential features of this nationally representative data sample, thereby, supporting the interpretation of the five multilevel models.

Table 2

Descriptive Statistics of Primary Schools by Ownership Type and Location in Nigeria

Variables	No. Schools	Percent (%)	No. Students	Percent (%)
School Ownership Type				
Public	92	62.16	2,531	64.65
Private Not-For-Profit	17	11.49	316	8.07
Private For-Profit	34	22.97	934	23.86
Community	3	2.03	72	1.84
Other	2	1.35	62	1.58
School Location				
Urban	33	22.3	941	24.04
Semi-Urban	80	54.05	944	51.85
Rural	35	23.65	2030	24.11
Total	148	100	3,916	100

Table 2 displays the descriptive statistics of the Nigerian primary schools sampled in this study and the distribution of these schools by ownership type (i.e. public, private not-for-profit, private for-profit, community, and other) and location (i.e. urban, semi-urban, and rural). As illustrated above, a total of 148 Nigerian primary schools and 3,916 nested students were sampled. Among this sample, the 92 publicly-owned schools constituted the predominant school ownership type (62.16%), the 34 private not-for-profit schools were the second most common ownership type (22.97%), the 17 private not-for-profit schools were the third most common ownership type (11.49%), and, the 3 community schools (2.03%) and 2 other school types (1.35%) represented miniscule portions of all schools. In tandem with variation in school ownership type, the 148 schools considerably varied in geographic location. While the majority of primary schools, or 80 schools (51.85%), were situated in semi-urban regions, the 33 urban (24.04%) and 35 rural (24.11%) schools represented roughly identical portions. Given the nested data structure, it is

pertinent to lastly note that the proportion of schools by type and location closely mirrors the proportion of students surveyed.

Table 3

Primary School Leaders in Nigeria by Position

Variables	Obs.	Percent (%)
School Leader Position		
Headteacher/ Principal	127	85.81
Owner/Director	8	5.41
Assistant Headteacher	11	7.43
Teacher	2	1.35
Total	148	100

Table 3 displays the descriptive statistics of primary school leaders in Nigeria according to position (i.e. headteachers/principals, owners/directors, assistant headteachers, and teachers). Among the 148 head teachers interviewed, 127 individuals were head teachers or school principals. This position constituted [EO1] the majority of the sample (85.81%). When enumerators failed to find a headteacher in place, they interviewed the individual in charge of school premises within that particular day. Concerning alternative school leaders, enumerators mostly resorted to interviewing 11 assistant headteachers (7.43%) and, to a slightly lesser extent, interviewing 8 school owners or directors (5.41%) and 2 regular teachers (1.35%). A dissection of the school leaders interviewed by position enables us to better comprehend the human dimension involved in the initial data collection process, the management of educational inputs, and the supervision of schooling often considered critical to student academic achievement.

Table 4

Descriptive Statistics of the Age and Sex of Grade 4 Primary School Students in Nigeria

Variables	Obs.	Percentage (%)	Mean	Std. Dev.	Min	Max.
Student Age & Sex						
Age (Years)	3,977	100	10.7	1.7	7	19
Male	1,971	49.2				
Female	2,036	50.8				

Table 4 presents the descriptive statistics concerning the age and sex of Grade 4 students. A brief descriptive analysis of the sample suggests that average age was approximately 11 years (*M* =

10.7, *SD* = 1.7*)*.[10] One can further observe that student sex was uniformly split with 49.2% male students and 50.8% female students represented. These two individual characteristics will play a crucial role in the construction of the multilevel models used to investigate the predictors of academic achievement.

Table 5

Descriptive Table of Overall Score of Grade 4 Students by Assessment Type and Threshold

Variables	Threshold 1 (0-33.33%)	Threshold 2 (33.34-63.33%)	Threshold 3 (63.34-100%)	Total
Assessment Type				
English	428	516	345	1,289
Mathematics	605	552	121	1,278
Science	280	587	471	1,338
Total	1,313	1,655	937	3,905

Table 5 presents the descriptive statistics of academic achievement scores of Grade 4 students by assessment type (i.e. English, mathematics, and science) and by threshold (Threshold 1 (T1): 0-33.33%, Threshold 2 (T2): 33.34-63.33%, and Threshold 3 (T3): 63.34-100%). A total of 3,905 Grade 4 students are represented in this study. An overview of the findings suggests that, while approximately 33.6% of all students performed in the percentage bracket of T1, the largest proportion of students, roughly 42.4%, performed in the percentage bracket of T2. Relative to the first and second achievement thresholds, T3 held the smallest proportion of students with merely 24% represented.

The findings further illustrate that students did not participate in all the three assessment types, rather, each third of the total student sample was randomly assigned to complete only one assessment. Among those who completed the English assessment, the distribution mirrored that of the general pattern, whereby the greatest portion of student performance occurred within T2 (40%), a smaller portion within T1 (33.2%) and the smallest portion within T3 (26.7%). In parallel to the English assessment, student performance on the Science assessment on average concentrated within T2 (43.87%), less so within T1 (20.9%). However, unlike the English assessment, student performance in science assessment within T3 (35.2%) was higher than T1, suggesting students

[10] Note that student age varies considerably with individuas ranging from 7 to 19 years old placed within the fourth grade. A histogram of age illustrates that the standard age bracket for this grade level (8-12 years) comprises 86% of all students. The inclusion of students situated outside the confines of this age bracket enables this study to emulate common patterns of age across the primary education system in Nigeria.

were on average able to attain higher scores on science than on English. In addition to these two assessment types, the distribution of student performance across thresholds on the mathematics assessment deviates from that of English and science. Among those who completed the Mathematics assessment, student performance on average tilted towards T1 (47.3%) with slightly less represented in T2 (43.19%) and a notable decline in T3 (9.5%). Comparatively assessing these findings, we can conclude that students performed the lowest on the mathematics assessment, but performed the highest on the science assessment, with English interposed between the two assessment types.

A plethora of factors may predict these students outcomes such as post-colonial educational policies, parental socio-economic status, and assessment design, however, in an attempt to focus on a single theoretical strand surrounding academic achievement, this study will examine the relationship between educational inputs and academic achievement among fourth grade students in Nigeria.

Table 6

Descriptive Statistics of Primary School Educational Inputs at the School-Level

Educational Inputs	Obs.	Mean	Std. Dev.	Min.	Max.
Cluster 1: Physical Facilities					
EI 1: Classrooms in Good Condition Per School (%)	145	68.22	37.18	0	100
EI 2: Desks in Functional Condition Per School (%)	131	80.47	25.78	0	100
EI 3: Presence of School toilet/s	148	0.70	0.46	0	1
Cluster 2: Educational Materials					
EI 4: Students with Textbooks (%)	145	46.71	46.71	0	100
EI 5: Students with Exercise Books (%)	144	86.42	26.47	0	100
EI 6: Presence of Corner Library/ies	145	0.20	0.40	0	1

Table 6 demonstrates the descriptive statistics of six educational inputs at the school-level split into two clusters: physical facilities and educational materials. Within the physical facilities cluster, the findings of the first educational input revealed that approximately 68% of all school classrooms surveyed were classified as "good" (M = 68.22, SD = 37.18*)*, while the findings of the second educational input indicated that roughly 80% of all school desks were deemed to be "unbroken" (M = 80.47, SD = 25.78). When comparing classrooms and desks, it becomes evident that the average percentage of functional desks appears to be more than 10 percentage points higher than good classrooms. This disparity, however, may be due to a noticeable decrease in sample size

from the 145 schools with recorded classroom observations to 131 schools with recorded desk observations. A plausible reason for this drop in percentage may be due to the absence of broken desks in these 16 schools. In tandem with the first and second educational inputs, findings from the third input reveals that nearly 70% of schools sampled contain at least one toilet (M = .7, SD = .46).

Within the educational materials cluster, the percentage of students with textbooks on average (M = 46.71, SD = 46.71) appears to be nearly half of students with exercise books (M = 86.42, SD = 26.547); thereby, suggesting students are more likely to possess exercise books than textbooks at the school-level. In addition to textbook and exercise book availability as the fourth and fifth salient educational input in this study, data on the sixth educational input reveals that 20% of schools had at least one classroom corner library (M = .2, SD = .4). Inversely, we can conclude that 80% of schools do not possess a corner library. The descriptive statistics of these school-level educational inputs coupled with the aforementioned tables on school types, school leaders, student age and sex, and achievement equips us with the knowledge to better engage with the ensuing multilevel models.

Appendix B: Cronbach's Alpha Outputs

Cronbach's alpha measures the internal reliability between items in a scale, which, in this case, can allow us to gauge whether the educational inputs theoretically classified under each cluster do indeed measure the same underlying properties from a statistical lens. That is, to examine (1) whether all six independent variables measure the underlying properties of "educational inputs," (2) whether classroom condition, desk condition, and presence of toilet/s measure the underlying properties of "physical facilities" and (3) whether textbook availability, exercise book availability, and the presence of corner library/ies measure the underlying properties of "educational materials." According to Gliem and Gliem (2003), the minimum acceptable value for Cronbach's value equates to 0.7. In other words, any value equal to or above 0.7 would indicate that the items measure the same construct, a value below this threshold would exhibit low internal consistency. As shown in Figure 3, the Cronbach's alpha for all inputs does not meet this threshold, but rather is situated at approximately 0.33. This value decreases considerably when examining the construct of physical facilities in Figure 4 and increases slightly when examining the construct of educational materials in Figure 5 given that they have, respectively, reached Cronbach's alpha values of approximately 0.11 and 0.35. As a result, this study finds that the theoretical clustering of these variables under "educational inputs" as well as further compartmentalizing them under "physical facilities" and "educational materials" does not hold true when using the reliability tests of Cronbach's alpha. This outcome, however, should not negate the significance of using input clusters as a beneficial framework for policy discourse and design.

Figure 3

Cronbach's Alpha for All Educational Inputs

Item	Obs	Sign	item-test correlation	item-rest correlation	average interitem covariance	alpha
classroom_~l	3775	-	0.5848	0.1182	49.39753	0.3060
desks_pct_~i	3476	-	0.4574	0.1166	62.38742	0.3141
toilets_pri	3915	+	0.4128	0.4065	81.9648	0.3360
stu_txtbks~t	3775	+	0.7067	0.2467	18.53581	0.1495
stu_exbks_..	3745	+	0.5644	0.2653	36.22201	0.2037
corner_lib~i	3855	+	0.2348	0.2272	82.08259	0.3371
Test scale					55.12294	0.3266

Figure 4

Cronbach's Alpha for Physical Facilities

Item	Obs	Sign	item-test correlation	item-rest correlation	average interitem covariance	alpha
classroom_~l	3775	+	0.8444	0.0328	3.648494	0.0218
desks_pct_~i	3476	+	0.6277	0.0765	6.666191	0.0196
toilets_pri	3915	-	0.4555	0.4590	70.15958	0.1252
Test scale					26.26286	0.1072

Figure 5

Cronbach's Alpha for Educational Materials

Item	Obs	Sign	item-test correlation	item-rest correlation	average interitem covariance	alpha
stu_txtbks~t	3775	+	0.8852	0.2489	-.4008593	.
stu_exbks_..	3745	+	0.7138	0.3110	3.442265	0.0082
corner_lib~i	3855	+	0.0889	0.1390	355.0269	0.4578
Test scale					120.3128	0.3500

Appendix C: Multilevel Method Selection

This study explored various versions and types of multilevel methods before coming to the conclusion that a two-level ordinal logit model would be the most suitable approach to investigate the association between inputs and achievements (Agresti, 2010). At the start of this dissertation, the study pursued a three-level hierarchical linear model (HLM) –a multilevel model for linear dependent variables; however, the data failed to converge in part due to the absence of a significant quantity of observations at the classroom-level (Level-2). The failure to converge, an indicator of a poorly constructed model, meant that I would need to down scale from a three-level to a two-level HLM (i.e. students nested in schools). But, the two-level HLM was also not a viable option as it violated the assumption of normal residuals (i.e. residuals of the model were not normally distributed) (See Appendix D). This HLM assumption was still violated even when the composition of inputs in the model altered. For instance, this dissertation also considered including a third input cluster pertaining to Parent-Teacher Association (PTA) participation, which specifically related to assessing the predictive powers of PTA budget and PTA meeting frequency. The inclusion of this third cluster not only violated the HLM assumption, but also posed a risk to the predictive assessments of inputs within the physical facilities and educational materials clusters.

As per best practice, I turned to exploring multilevel ordinal logistic models (MOLM) as an alternative method when faced with assumption violations and parameter complications within HLM. To test the suitability of both methods, this study converted the continuous dependent variable into an ordinal dependent variable split into 10 achievement thresholds. However, when testing three-level and a two-level logit model, this model often failed to converge most likely due to the small amount of observations dedicated to each of the ten thresholds and the partial violation in the parallel lines assumption (i.e. the relationship between each pair of outcome groups was not identical). To reconcile this violation, I thought of pursuing a generalized ordinal logistic model (GOLM) as it has the capacity to relax the assumption of parallel lines for some variables, but this method does not account for an indispensable component of the data – the nested structure (See Appendix E). Not only is the nested structure an integral element of the data, but dependence itself should be viewed as a fascinating dimension that can enable us to unscramble variability at the student and school levels. Stata, the statistical software employed, unfortunately does have any

commands or methods that can pursue multilevel logit models while relaxing the parallel lines assumption.

Upon reading further statistical literature, it becomes apparent that the need to test for this assumption has been highly contested with a bourgeoning number of scholars contending that a failure does not nullify the model for two technical reasons. First, Stata does not provide a clear command that can test this assumption, rather, researchers would have to use graphical methods. Secondly, the normative means of testing this assumption tends to yield biased estimates. Clearly, future technical innovations must construct statistical tools that allow researchers to test for the parallel lines assumptions and to implement a generalized version of a multilevel mixed effects ordered logistic regression. With the parallel lines assumption reconciled, there remains the issue of a limited number of observations that is attributed to a ten-threshold ordinal variable. To curb this concern, I recategorized academic achievement from a ten-threshold ordinal variable to a three-threshold ordinal variable, which produced a stronger model that, ultimately, did converge (See Appendix H).

In parallel to the transitions in model selection, I initially intended to pursue an alternative research method. As opposed to a mono-method quantitative approach, I planned to pursue a multi-method quantitative approach in order to triangulate the input-achievement associations using two different input metrics: (1) standard indicators and (2) headteacher perceptions. The purpose behind this approach was to internalize criticism stemming from educational research, in which scholars purported that mono-method quantitative analyses were privy to producing reductionist portrayals of the educational system and prone to yielding incorrect statistical inferences. A multi-method quantitative approach was, ultimately, not a feasible option. The World Bank data set did not contain data items for all inputs. As a result, this study decided to use a mono-method approach (i.e. MOLM) that relies on standard input indicators to empirically examine the input-achievement nexus.

Appendix D: Hierarchical Linear Model Violation

In hierarchical linear modeling, the assumption of residual normality or multivariate normality is considered essential for valid hypothesis testing (i.e. valid p-values, t-test values, and F-test values). To test for this assumption, this study will rely on graphical and non-graphical methods: (1) Q-Q Plot, (2) P-P Plot, (3) Kernel Density Estimate, (4) Shapiro-Francia Test, (5) Skewness/Kurtosis Test, (6) Multivariate Normality Tests (i.e. Mardia, Henze-Zirkler, and Doornik-Hansen). The reliance on these six methods not only allows this study to test whether the sample was randomly selected from a normal distribution, but also it enables this study to corroborate estimations from four different statistical methods. Relative to alternative hierarchical linear models explored, this study will focus on model 3 –a fixed effect model with all six educational inputs – due to its superiority in model fit reflected in the likelihood ratio tests below.

Figure 6

HLM Nested Model Comparisons M1– M5: Likelihood Ratio Tests

```
. lrtest M1 M2

Likelihood-ratio test                                 LR chi2(3)  =     37.23
(Assumption: M1 nested in M2)                         Prob > chi2 =    0.0000

. lrtest M2 M3

Likelihood-ratio test                                 LR chi2(3)  =     19.86
(Assumption: M2 nested in M3)                         Prob > chi2 =    0.0002

. lrtest M3 M4

Likelihood-ratio test                                 LR chi2(2)  =      0.06
(Assumption: M3 nested in M4)                         Prob > chi2 =    0.9714

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space.  If this is not true, then
      the reported test is conservative.

. lrtest M3 M5

Likelihood-ratio test                                 LR chi2(14) =     18.63
(Assumption: M3 nested in M5)                         Prob > chi2 =    0.1795

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space.  If this is not true, then
      the reported test is conservative.
```

```
. lrtest M2 M4

Likelihood-ratio test                                 LR chi2(5)  =     19.92
(Assumption: M2 nested in M4)                         Prob > chi2 =    0.0013

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space.  If this is not true, then
      the reported test is conservative.

. lrtest M2 M5

Likelihood-ratio test                                 LR chi2(17) =     38.49
(Assumption: M2 nested in M5)                         Prob > chi2 =    0.0021

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space.  If this is not true, then
      the reported test is conservative.

. lrtest M4 M5

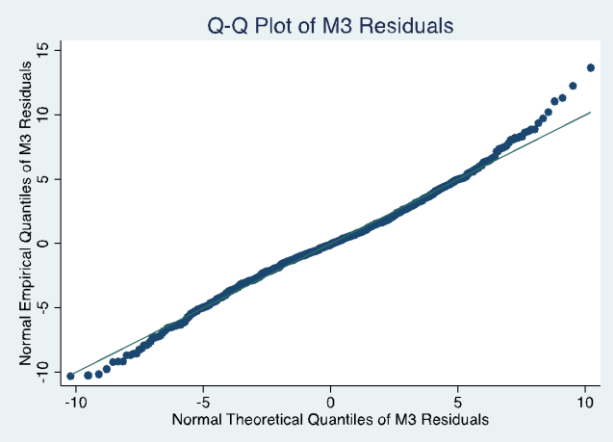
Likelihood-ratio test                                 LR chi2(12) =     18.57
(Assumption: M4 nested in M5)                         Prob > chi2 =    0.0993
```

Residual Normality Tests for HLM

Method 1: A Q-Q plot, also known as a quantile-quantile plot, maps the quantiles of the standardized residuals against the quantiles of a normal distribution (a 45-degree line). To ascertain for the presence of residual normality, values must closely conform to the predicted quantiles; otherwise, deviations from the normal distribution may hint at the skewness or significant outliers in the dataset. As illustrated in Figure 7, the Q-Q plots of the residuals in models 3 appears to be slightly non-normal in light of deviation in the tails.

Figure 7

Q-Q Plot of Model 3 Residuals

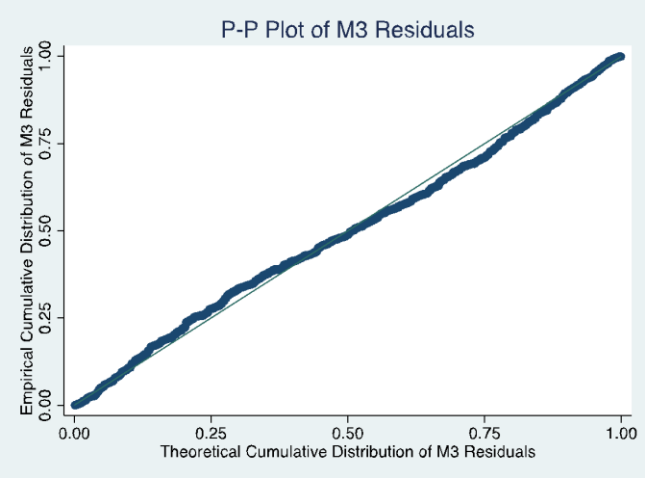


Method 2: To verify the findings of a Q-Q plot, the P-P plot, commonly regarded as a probability-probability plot, is an alternative probability plot used to examine the distribution of two data sets. As opposed to non-cumulative distribution manifested in the form of quantiles, the P-P plot estimates the theoretical cumulative distribution against the empirical cumulative distribution.

Contrary to the Q-Q plot, Figure 8 suggests that the P-P plots of the residuals closely align with the theoretical cumulative distribution line suggesting that the residuals are normally distributed.

Figure 8

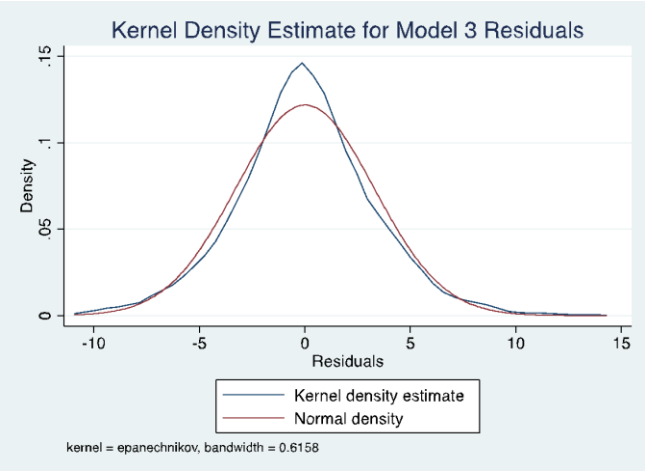
P-P Plot of Model 3 Residuals



Method 3: The KDE estimates the probability density of a random variable using a non-parametric method by plotting, in this case, model 3 residuals against a normal density estimate. The Kernel density estimate in Figure 9 echoes the findings of the Q-Q plot, suggesting that, while the diagram predominantly aligns with the normal kernel density curve, the discrepancy centers at the peak. This peak indicates that residual distribution of the residuals suffers from kurtosis, and, thus does not appear to be normally distributed.

Figure 9

Kernel Density Estimate for Model 3 Residuals



Method 4: The Shapiro-Francia test is an analysis of variance test for normality quite similar to the Shapiro-Wilk test, the difference constituting that the latter is frequently used for large sample sizes (Shapiro & Francia, 1972; Royston, 1983). When conducting this test in Stata, the null hypothesis is assumed to be normally distributed. To reject this null hypothesis (yielding a p-value equal to or less than 0.05) is to claim that the random effects are not normally distributed. One of the reasons, this study relies on the Shaprio-Francia test can be attributed to the comparative simulation findings of Mbah and Paothong (2015), whereby the authors found that, amidst nine common normality tests such as Kolmogorov–Smirnov and Pearson's Chi-Square tests, the Shapiro-Francia test was best at detecting deviation from normality at varying sample sizes. In light of the contradiction among graphical estimations of normality, this study turns to non-graphical methods for further investigation. Starting with the Shapiro-Francia test in Figure 10, the p-values below 0.05 means that this study can reject the null hypothesis to contend that the residuals of model 3 are not normally distributed.

Figure 10

Shapiro-Francia Test for Model 3 Residuals

```
. sfrancia m3_residual

                  Shapiro-Francia W' test for normal data

    Variable |       Obs        W'         V'         z       Prob>z
-------------+--------------------------------------------------------
 m3_residual |     1,114    0.99130      6.422      4.293     0.00001
```

Method 5: The Skewness-Kurtosis test aims to identify skewness (a measure of asymmetry with reference to the mean of a probability distribution) and extreme kurtosis (a measure of the degree to which values cluster in the tails or peaks of a probability distribution). In terms of skewness, any value departing from 0 would indicate skewness; however, considering that some degree of skewness tends to be commonly prevalent in datasets, the general rule of thumb is that distribution between -0.5 and 0.5 can be interpreted as approximately symmetric. A skewness greater than the parameters of -1 and 1 is highly skewed. Similar to the conclusions of the Shapiro-Francia test, the findings of the Skewness-Kurtosis test in Figure 11 reveals p-values below 0.05, thus, indicating that the residuals are not normally distributed.

Figure 11

Skewness-Kurtosis Test for Model 3 Residuals

```
. sktest m3_residual

                    Skewness/Kurtosis tests for Normality
                                                      ------ joint ------
    Variable |        Obs  Pr(Skewness)  Pr(Kurtosis) adj chi2(2)    Prob>chi2
-------------+-------------------------------------------------------------
 m3_residual |      1,114     0.1058        0.0000       19.22         0.0001
```

Method 6: To test multivariate normality, this study examines three tests: (1) Mardia's (1970) multivariate kurtosis and skewness test (2) Henze and Zirkler's (1990) consistent test, and (3) Doornik and Hansen's (2008) omnibus test. For a normal distribution, this study would need to avoid rejecting the null hypotheses. As illustrated in Figure 12, the multivariate normality tests of Mardia, Henze-Zikler, and Doornik-Hansen cement the non-normal distribution of the residuals of model 3 given that all tests yield p-values below 0.05 with the exception of Mardia's skewness test.

Figure 12

Multivariate Normality Tests for Model 3 Residuals

```
Test for multivariate normality

    Mardia mSkewness =    .014001     chi2(1) =     2.614   Prob>chi2 =  0.1060
    Mardia mKurtosis =  3.869942     chi2(1) =    35.128   Prob>chi2 =  0.0000
    Henze-Zirkler    =  4.378031     chi2(1) =    15.364   Prob>chi2 =  0.0001
    Doornik-Hansen                   chi2(2) =    29.893   Prob>chi2 =  0.0000
```

These six graphical and non-graphical normality tests suggest that the residuals of model 3 are not normally distributed. While HLM models with non-normal random effects can be pursued, an HLM model with non-normal residuals may generate biased results. As a result of this violation, a hierarchical linear model cannot be pursued. Considering these circumstances, Peterson and Harrell (1990) suggest less restrictive models such as the multilevel mixed effects ordered logistic regression can be an appropriate alternative.

Appendix E: Generalized Ordinal Logistic Regression Outputs

While not relevant to the multilevel model chosen for this study, a proportionality test was empirically conducted within the framework of a generalized ordered logistic regression in order to better understand the data. The outputs below reveal that classroom condition, toilet/s, and corner library/ies adhere to the principle of proportional odds, whereas desk condition, textbook availability, and exercise book availability do not meet this assumption. These results do not align with the findings of the graphical assumption test conducted within the multilevel ordinal logistic regression framework in Figure 13. Not that this study compares apples to pears, but the difference in these findings demonstrate the importance of customizing the proportional odds assumption testing with respect to the model, suggesting that the hierarchical nature of the dataset can influence proportionality. Thus, this output consolidates the methodological rationale behind pursuing a multilevel logit model as opposed to a generalized logit model for this study.

Figure 13

Proportionality Tests for Generalized Ordinal Logistic Regressions

```
. gologit2 tot_pct_pri_ord classroom_bad__pct_pri_official desks_pct_pri i.toilets_pri stu_txtbks_pri_per_student stu_exbks_pri_per_stu
> dent i.corner_lib_pri, pl(0b.corner_lib_pri 1.corner_lib_pri 0b.toilets_pri 1.toilets_pri classroom_bad__pct_pri_official)

Generalized Ordered Logit Estimates             Number of obs     =      3,380
                                                LR chi2(30)       =     744.05
                                                Prob > chi2       =     0.0000
Log likelihood = -7130.8097                     Pseudo R2         =     0.0496

 ( 1)  [1]1.corner_lib_pri - [2]1.corner_lib_pri = 0
 ( 2)  [1]1.toilets_pri - [2]1.toilets_pri = 0
 ( 3)  [1]classroom_bad__pct_pri_official - [2]classroom_bad__pct_pri_official = 0
 ( 4)  [2]1.corner_lib_pri - [3]1.corner_lib_pri = 0
 ( 5)  [2]1.toilets_pri - [3]1.toilets_pri = 0
 ( 6)  [2]classroom_bad__pct_pri_official - [3]classroom_bad__pct_pri_official = 0
 ( 7)  [3]1.corner_lib_pri - [4]1.corner_lib_pri = 0
 ( 8)  [3]1.toilets_pri - [4]1.toilets_pri = 0
 ( 9)  [3]classroom_bad__pct_pri_official - [4]classroom_bad__pct_pri_official = 0
 (10)  [4]1.corner_lib_pri - [5]1.corner_lib_pri = 0
 (11)  [4]1.toilets_pri - [5]1.toilets_pri = 0
 (12)  [4]classroom_bad__pct_pri_official - [5]classroom_bad__pct_pri_official = 0
 (13)  [5]1.corner_lib_pri - [6]1.corner_lib_pri = 0
 (14)  [5]1.toilets_pri - [6]1.toilets_pri = 0
 (15)  [5]classroom_bad__pct_pri_official - [6]classroom_bad__pct_pri_official = 0
 (16)  [6]1.corner_lib_pri - [7]1.corner_lib_pri = 0
 (17)  [6]1.toilets_pri - [7]1.toilets_pri = 0
 (18)  [6]classroom_bad__pct_pri_official - [7]classroom_bad__pct_pri_official = 0
 (19)  [7]1.corner_lib_pri - [8]1.corner_lib_pri = 0
 (20)  [7]1.toilets_pri - [8]1.toilets_pri = 0
 (21)  [7]classroom_bad__pct_pri_official - [8]classroom_bad__pct_pri_official = 0
 (22)  [8]1.corner_lib_pri - [9]1.corner_lib_pri = 0
 (23)  [8]1.toilets_pri - [9]1.toilets_pri = 0
 (24)  [8]classroom_bad__pct_pri_official - [9]classroom_bad__pct_pri_official = 0
```

tot_pct_pri_ord	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1						
classroom_bad__pct_pri_official	-.0072711	.0009515	-7.64	0.000	-.009136	-.0054061
desks_pct_pri	-.0145383	.0025415	-5.72	0.000	-.0195195	-.0095571
toilets_pri						
1	.3123439	.0779266	4.01	0.000	.1596106	.4650771
stu_txtbks_pri_per_student	.0175557	.0020977	8.37	0.000	.0134443	.0216672
stu_exbks_pri_per_student	.01044	.0022077	4.73	0.000	.006113	.0147671
corner_lib_pri						
1	-.0093872	.0789585	-0.12	0.905	-.1641431	.1453686
_cons	1.520616	.2045011	7.44	0.000	1.119802	1.921431
2						
classroom_bad__pct_pri_official	-.0072711	.0009515	-7.64	0.000	-.009136	-.0054061
desks_pct_pri	-.0103914	.0018385	-5.65	0.000	-.0139949	-.0067879
toilets_pri						
1	.3123439	.0779266	4.01	0.000	.1596106	.4650771
stu_txtbks_pri_per_student	.0129919	.0013477	9.64	0.000	.0103505	.0156333
stu_exbks_pri_per_student	.0060586	.0018149	3.34	0.001	.0025015	.0096156
corner_lib_pri						
1	-.0093872	.0789585	-0.12	0.905	-.1641431	.1453686
_cons	.8587541	.1757262	4.89	0.000	.514337	1.203171
3						
classroom_bad__pct_pri_official	-.0072711	.0009515	-7.64	0.000	-.009136	-.0054061
desks_pct_pri	-.0096315	.0015716	-6.13	0.000	-.0127118	-.0065513
toilets_pri						
1	.3123439	.0779266	4.01	0.000	.1596106	.4650771
stu_txtbks_pri_per_student	.012107	.00113	10.71	0.000	.0098923	.0143218
stu_exbks_pri_per_student	.0055786	.0016163	3.45	0.001	.0024107	.0087464
corner_lib_pri						
1	-.0093872	.0789585	-0.12	0.905	-.1641431	.1453686
_cons	.2737829	.1604776	1.71	0.088	-.0407474	.5883133
4						
classroom_bad__pct_pri_official	-.0072711	.0009515	-7.64	0.000	-.009136	-.0054061
desks_pct_pri	-.0090847	.0014526	-6.25	0.000	-.0119318	-.0062376
toilets_pri						
1	.3123439	.0779266	4.01	0.000	.1596106	.4650771
stu_txtbks_pri_per_student	.01336	.0010132	13.19	0.000	.0113741	.0153459
stu_exbks_pri_per_student	.0043348	.0015417	2.81	0.005	.0013131	.0073564
corner_lib_pri						
1	-.0093872	.0789585	-0.12	0.905	-.1641431	.1453686
_cons	-.3894398	.1543808	-2.52	0.012	-.6920206	-.086859

5						
classroom_bad__pct_pri_official	-.0072711	.0009515	-7.64	0.000	-.009136	-.0054061
desks_pct_pri	-.0074597	.0014385	-5.19	0.000	-.0102791	-.0046404
toilets_pri						
1	.3123439	.0779266	4.01	0.000	.1596106	.4650771
stu_txtbks_pri_per_student	.014126	.0009957	14.19	0.000	.0121745	.0160774
stu_exbks_pri_per_student	.0024986	.0015835	1.58	0.115	-.0006051	.0056023
corner_lib_pri						
1	-.0093872	.0789585	-0.12	0.905	-.1641431	.1453686
_cons	-.8556834	.1576302	-5.43	0.000	-1.164633	-.5467338
6						
classroom_bad__pct_pri_official	-.0072711	.0009515	-7.64	0.000	-.009136	-.0054061
desks_pct_pri	-.0052319	.0015019	-3.48	0.000	-.0081755	-.0022882
toilets_pri						
1	.3123439	.0779266	4.01	0.000	.1596106	.4650771
stu_txtbks_pri_per_student	.0150829	.0010789	13.98	0.000	.0129683	.0171975
stu_exbks_pri_per_student	.0013916	.0018113	0.77	0.442	-.0021584	.0049416
corner_lib_pri						
1	-.0093872	.0789585	-0.12	0.905	-.1641431	.1453686
_cons	-1.574082	.1761532	-8.94	0.000	-1.919336	-1.228828
7						
classroom_bad__pct_pri_official	-.0072711	.0009515	-7.64	0.000	-.009136	-.0054061
desks_pct_pri	-.004416	.0017131	-2.58	0.010	-.0077737	-.0010583
toilets_pri						
1	.3123439	.0779266	4.01	0.000	.1596106	.4650771
stu_txtbks_pri_per_student	.0189591	.0013856	13.68	0.000	.0162434	.0216748
stu_exbks_pri_per_student	-.0010082	.0023842	-0.42	0.672	-.0056811	.0036647
corner_lib_pri						
1	-.0093872	.0789585	-0.12	0.905	-.1641431	.1453686
_cons	-2.416539	.2211011	-10.93	0.000	-2.849889	-1.983189
8						
classroom_bad__pct_pri_official	-.0072711	.0009515	-7.64	0.000	-.009136	-.0054061
desks_pct_pri	-.0083484	.0025706	-3.25	0.001	-.0133867	-.0033102
toilets_pri						
1	.3123439	.0779266	4.01	0.000	.1596106	.4650771
stu_txtbks_pri_per_student	.0245604	.0022339	10.99	0.000	.0201821	.0289387
stu_exbks_pri_per_student	-.0033448	.0040161	-0.83	0.405	-.0112162	.0045266
corner_lib_pri						
1	-.0093872	.0789585	-0.12	0.905	-.1641431	.1453686
_cons	-3.553283	.3601305	-9.87	0.000	-4.259125	-2.84744
9						
classroom_bad__pct_pri_official	-.0072711	.0009515	-7.64	0.000	-.009136	-.0054061
desks_pct_pri	-.0104602	.0051964	-2.01	0.044	-.0206449	-.0002755
toilets_pri						
1	.3123439	.0779266	4.01	0.000	.1596106	.4650771
stu_txtbks_pri_per_student	.0383364	.0064638	5.93	0.000	.0256675	.0510053
stu_exbks_pri_per_student	-.0155669	.007907	-1.97	0.049	-.0310644	-.0000695
corner_lib_pri						
1	-.0093872	.0789585	-0.12	0.905	-.1641431	.1453686
_cons	-5.080743	.6839505	-7.43	0.000	-6.421262	-3.740225

Appendix F: Assumption Testing for Two-Level Ordinal Logistic Regressions

In multilevel analysis, scholars must adhere to statistical assumptions to ensure that model design and interpretation are reliable. If certain assumptions are violated, the model design may be adjusted accordingly, or alternative statistical methods may be pursued for greater accuracy. It is pertinent to mention, however, statisticians principally agree that OLR needs to meet fewer assumptions compared to alternative methods such as hierarchical linear modeling, proportional odds logistic regression models, or repeated-measures analysis of variance (Raudenbush & Bryk, 2002; O'Connell, 2006). Specifically, a logistic regression does not need to meet the assumptions of linearity, normally distributed residuals, or homoscedasticity common in linear models. The assumptions of a generalized hierarchical ordinal logistic regression can be viewed below:

Assumption 1: Ordinal Dependent Variable (Met)

Assumption 2: Continuous, Ordinal, or Categorical Independent Variables (Met)

Assumption 3: Absence of Multicollinearity (Met)

Assumption 4: Substantial Sample Size (Met)

Assumption 5: Proportional Odds (Assumption Not Fully Met, Model Adjusted)

Assumption 1: Ordinal Dependent Variable (Met)

This study meets the assumption of an ordinal dependent variable to pursue an ordinal logistic regression given that the dependent variable—overall academic achievement—is measured at an ordinal interval. As evident in Figure 14, the overall percentage of Grade 4 students' academic achievement has been split into three tiers based on 33.33% intervals. Considering that academic achievement consists of an aggregation of three assessment subjects, one can further observe the frequency of student score within each interval according to assessment type.

Figure 14

Overall Academic Achievement of Grade 4 Students in Nigeria by Assessment Type

```
. tab assessment_type tot_pri_ord_30_pct
```

Assessment Type	tot_pri_ord_30_pct 0	1	2	Total
P4 English	428	516	345	1,289
P4 Maths	605	552	121	1,278
P4 Science	280	587	471	1,338
Total	1,313	1,655	937	3,905

Assumption 2: Continuous, Ordinal, or Categorical Independent Variables (Met)

This study meets the second assumption of a multilevel ordinal logistic regression given that all six independent variables can be classified as either continuous or categorical. As can be examined in Figures 15-20 below, the model contains 4 continuous variables (i.e. percentage of classrooms in bad condition, percentage of desks in broken condition, percentage of students with textbooks, percentage of students with exercise books and the number of corner libraries per school) and 2 dichotomous ordinal variables (i.e. presence of toilets in a school and the presence of a corner library).

Figure 15

Descriptive Statistics of the Percentage of Classroom in Bad Condition

```
. sum classroom_bad__pct_pri_official
```

Variable	Obs	Mean	Std. Dev.	Min	Max
classroom_~l	3,775	31.69453	37.03514	0	100

Figure 16

Descriptive Statistics of the Percentage of Desks in Broken Condition

```
. sum desks_pct_pri
```

Variable	Obs	Mean	Std. Dev.	Min	Max
desks_pct_~i	3,476	19.76305	26.50117	0	100

Figure 17

Tabulation of the Presence of toilet/s in Schools

```
. tab toilets_pri
```

toilets_pri	Freq.	Percent	Cum.
0	1,179	30.11	30.11
1	2,736	69.89	100.00
Total	3,915	100.00	

Figure 18

Descriptive Statistics of the Percentage of Students with Textbooks

```
. sum stu_txtbks_pri_per_student
```

Variable	Obs	Mean	Std. Dev.	Min	Max
stu_txtbks~t	3,775	46.71456	41.08055	0	100

Figure 19

Descriptive Statistics of the Percentage of Students with Exercise Books

```
. sum stu_exbks_pri_per_student
```

Variable	Obs	Mean	Std. Dev.	Min	Max
stu_exbks_..	3,745	86.42354	26.46707	0	100

Figure 20

Tabulation of the Presence of a Corner Library in Schools

```
. tab corner_lib_pri
```

corner_lib_pri	Freq.	Percent	Cum.
0	3,063	79.46	79.46
1	792	20.54	100.00
Total	3,855	100.00	

Assumption 3: Absence of Multicollinearity (Met)

This study meets the assumption of no multicollinearity. Before delving into how this assumption has been met, it is pertinent to first explain the concept of multicollinearity. Scholars contend that multicollinearity refers to a phenomenon whereby selected independent variables appear to be highly correlated with each other (Mansfield & Helms, 1982). The prevalence of multicollinearity in an ordinal logistic regression can distort coefficient estimates and standard errors, which often threatens the validity of the selected model. Given these adverse effects, testing for multicollinearity has been profoundly considered an integral preliminary step to the consequent

construction of an ordinal logistic regression (Yu, Jiang, and Land, 2015). To test this assumption, the standard method that relies on an assessment of the Variance Inflation Factor (VIF) may not be the most suitable in light of the presence of two dichotomous ordinal variables in this study. While statistically the VIF test may be run, the suitability of this for method a combination of ordinal and continuous variables is still a widely contested topic among scholars (Davis & Salkind, 2010). Studies such as Salo, Laksoonen, and Santtila (2016) have, instead, relied on the polychoric correlation test as a more reliable and appropriate method. The polychoric correlation refers to a statistic used to estimate the correlation between independent variables including ones of a dichotomous nature. In this study, the polychoric correlation matrix in Figure 21 below finds that all of the correlations do not appear to be significantly strong as they range from 0.0086 to 0.042. Based on these values we can conclude that this study does not suffer from multicollinearity.

Figure 21

Multicollinearity Test of Educational Inputs in an Ordinal Logistic Regression

```
. polychoric classroom_bad__pct_pri_official desks_pct_pri toilets_pri stu_txtbks_pri_per_student stu_exbks_pri_per_student corne
> _lib_pri

Polychoric correlation matrix

                            classroom_bad__pct_pri_official                  desks_pct_pri
classroom_bad__pct_pri_official                               1
                  desks_pct_pri                      .04106369                              1
                    toilets_pri                     -.42968146                       -.3459085
     stu_txtbks_pri_per_student                     -.15157834                       .00861506
      stu_exbks_pri_per_student                      .12298018                      -.12696742
                 corner_lib_pri                     -.39077474                      -.31104279

                                                    toilets_pri      stu_txtbks_pri_per_student
                    toilets_pri                               1
     stu_txtbks_pri_per_student                      .26133226                              1
      stu_exbks_pri_per_student                      .09564729                       .30730885
                 corner_lib_pri                      .42144706                       .30149254

                                         stu_exbks_pri_per_student                 corner_lib_pri
      stu_exbks_pri_per_student                                1
                 corner_lib_pri                     -.09878973                              1
```

Assumption 4: Substantial Sample Size (Met)

In accordance with statistical literature, this study contains substantial observations at each level, thereby, meeting the fourth assumption of a multilevel ordered logit model. According to Scherbaum and Ferreter (2009), the maximum likelihood estimation methods employed in multilevel models for ordinal data are asymptotic, as in, they assume that the sample size must be sufficiently large in order to yield accurate likelihood ratio tests, parameter estimates, and variances estimates. The operationalization of this asymptotic theory has, however, been contested.

That is, no consensus has been reached on the precise sample size needed for reliable estimates, yet, scholars largely agree that the prime restriction on sample size should concern the highest level of the hierarchy (McNeish & Stapleton, 2014). An investigation by Mass and Hox (2005) on how different group level sizes affect the accuracy of estimates and standard errors in a multi-level model revealed that "only a small sample size at level two (meaning a sample of 50 or less) leads to biased estimates of the second-level standard errors (p. 86). In all of the other simulated conditions the estimates of the regression coefficients, the variance components, and the standard errors are unbiased and accurate." Thus, this study will focus on testing this assumption at the group-level.

For guidelines on sample sizes specialized on ordinal group-level data, we turn to the recent simulation studies of Austin (2010) and Moineddin et al. (2007), whereby the authors relied on a random-intercept model and a random-coefficient model, respectively, to discern what range of cluster size may yield reliable estimates. For a random-intercepts model, Austin found that a cluster size of 10 or larger produces reliable results. For a random-coefficient model, Moineddin et al. revealed that a sample size equal to or larger than 30 clusters at the group-level produces reliable results. Applying these insights on this study, a sample size of 97 schools at the group-level seems to be well above the recommended number of clusters, which serves to enhance the robustness of the logit model estimates. Therefore, this study can conclude that the assumption of substantial sample size at the group-level has been met.

Assumption 5: Proportional Odds (Assumption Not Fully Met, Model Adjusted)

This study finds that the assumption of proportional odds has not been fully met across all parameters of the multilevel mixed effects ordered logistic regression. An investigation into methodological literature suggests that this slight violation may not be of prime concern to the construction of a valid model, which has prompted this study to proceed with the intended logistic regression without completely adhering to the proportional odds assumption. Additionally, adjustments to the interpretation of specific parameters in the model will be made as a result of this partial violation. Prior to disclosing test results, this study briefly provides an overview of the proportional odds assumption.

The concept of proportional odds, also known as parallel lines, refers to an underlying assumption within standard ordinal logistic regressions, which posits that the relationship between

each pair of outcome groups must appear identical. Within the context of this study, this assumption would imply that each set of delineated intervals of academic achievement should be the same. This assumption allows for the standardization of coefficient interpretation, whereby one may describe the relationship between highest and the lowest categories in a similar manner as the relationship between the second highest and the second lowest. Through the assumption of proportional odds, researchers rely on one model, as opposed to several models, in order to efficiently interpret and compare coefficients. Within statistical literature, the need to test for this assumption has been highly contested. While some scholars argue that this assumption functions as an important prerequisite to the valid construction and interpretation of the model, a growing number of scholars have recently argued that the violation of this assumption may not nullify an ordered logistic regression for two technical reasons.

First, statistical software does not permit this assumption to be tested while simultaneously retaining the nested structure of the data. Stata's "meologit" command uses an algorithm that assumes proportional odds (slopes estimate between pairs are the same across responses levels) and, unfortunately, the software does not provide a clear command to test this assumption nor a feasible mechanism to adjust the multilevel mixed model in face of violation. To conjointly test and adjust for proportional odds, researchers would have to switch to other models such as the generalized logistic regression that can rely on the "gologit2" command with the "autofit" feature (See Appendix E below for Stata outputs testing the proportional odds assumption within a generalized logistic regression). To refrain from ignoring an indispensable element of the study—the nested data structure, this study decided the pursuit multilevel mixed-effects ordered logistic regression would be a more accurate representation of the data than a generalized model. The following limitation further cements this conclusion.

In tandem with the technical limitations of the software, the second reason a partial violation of proportional odds can still produce a valid model lies with the tendency of assumption tests to yield biased results. A plethora of scholars described the assumption tests as "anti-conservative, that it nearly always results in rejection of the proportional odds assumption, particularly when the number of explanatory variables is large (Brant, 1990), the sample size is large (Allison, 1999; Clogg & Shihadeh, 1994) or there is a continuous explanatory variable in the

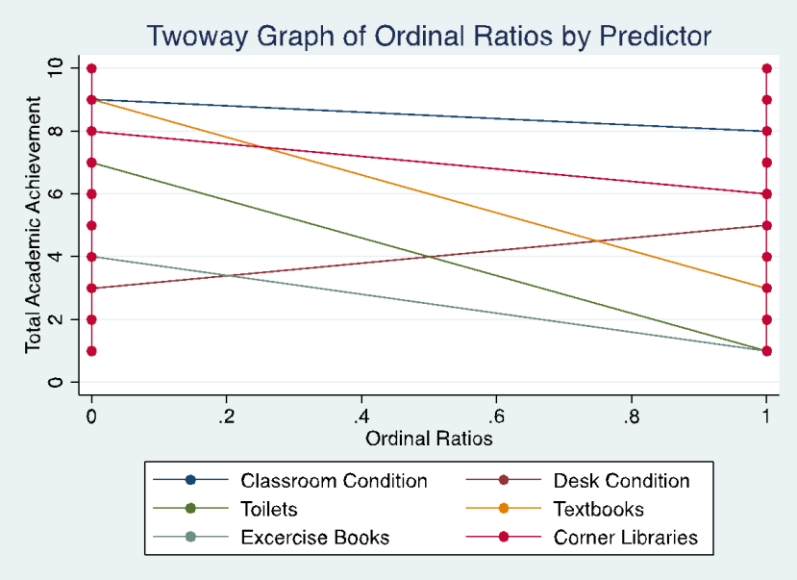
model (Allison, 1999)" (O'Connell, 2006, p. 29).[11] Applied to this study, one can note that the considerable sample size of the data coupled with the specific concentration of violations on large continuous variables compels us to disregard this partial violation.

The absence of a Stata package to test for proportionality within a multilevel ordinal logistic framework has commonly led researchers to adopt ad hoc graphical methods of assumption testing (Cohen et al., 2013). Through the process of creating and plotting logistic regression equations for each independent variable, one can test the proportional odds assumption by examining the extent to which predictor ordinal ratios vary at different thresholds of academic achievement. Consensus on what demarcates a significant departure in proportionality has not been reached, therefore, this study will consider departures above the net difference value of 3 as a violation of the assumption. As illustrated in Figure 22, the ordinal ratio of classroom condition, desk condition, exercise book availability, and corner libraries do not significantly vary at different academic achievement thresholds given net difference values of 0.5, 2, 3, and 2, respectively. On the other hand, the ordinal ratios of toilets and textbook availability depart significantly at different academic achievement thresholds, with net difference values of 7 and 6, respectively. That is, classroom condition, desk condition, exercise book availability, and corner libraries meet the assumption of proportionality, while toilets and textbook availability do not. Thus, this study finds it reasonable to conclude that the ordinal proportional odds model fairly summarizes the data patterns in relation to threshold selection for all but two predictors. The implication of this partial violation would mean that the covariate effects of toilets and textbook availability should not be generalized across all cumulative splits in the data.

[11] Harrell (2001) adds that violations tend to be miniscule in size and, thus, overemphasized. Data sets with large sample sizes should place more trust on the maximum likelihood estimation process as opposed to pondering as to whether this assumption was met or violated.

Figure 22

Two-way Connected Plot of Ordinal Ratios by Predictor



Appendix G: Mathematical Notations of Two-Level Models

Model 1: Null Unconditional Model

The two-level ordinal logistic regression can assume the following form for student *i* in school *j*:

$$\text{Level 1: } \text{logit}[\pi_{kij}(Y > K)] = \ln\frac{(Yij > K)}{(Yij \ K)}) = -a_k + \beta_{0j} \quad (19)$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + u_{0j} \quad (20)$$

$$\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = -a_k + u_{0j} \quad (21)$$

Model 2: Random-Intercept Model with 4 Fixed Parameters

$$\text{Level 1: } \text{logit}[\pi_{kij}(Y > K)] = \ln\frac{(Yij > K)}{(Yij \ K)}) = -a_k + \beta_{0j} \quad (22)$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + u_{0j} \quad (23)$$

$$\beta_{1j} = \gamma_{20} + u_{1j} \quad (24)$$

$$\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + u_{1j}) \quad (25)$$

Model 3: Random-Intercept Model with 6 Fixed Parameters

$$\text{Level 1: } \text{logit}[\pi_{kij}(Y > K)] = \ln\frac{(Yij > K)}{(Yij \ K)}) = -a_k + \beta_{0j} \quad (26)$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + u_{0j} \quad (27)$$

$$\beta_{1j} = \gamma_{20} + u_{1j} \quad (28)$$

$$\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + u_{1j}) \quad (29)$$

Model 4: Random-Covariate Model with 7 Fixed Parameters and 1 Random Parameter

Level 1: $\text{logit}[\pi_{kij}(Y > K)] = \ln\frac{(Yij > K)}{(Yij\ K)} = -a_k + (\beta_{0j} + \beta_{1j}\text{StudentSex}_{ij})$ (30)

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j +$ (31)
$\gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + u_{0j}$

$\beta_{1j} = \gamma_{20} + u_{1j}$ (32)

$\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j +$ (33)
$\gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + u_1\text{jStudentSex}_{ij})$

Model 5: Random-Covariate Model with Cross-Level Interactions

Level 1: $\text{logit}[\pi_{kij}(Y > K)] = \ln\frac{(Yij > K)}{(Yij\ K)} = -a_k + (\beta_{0j} + \beta_{1j}\text{StudentSex}_{ij})$ (34)

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j + \gamma_{03}\text{Toilets}_j +$ (35)
$\gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j + \gamma_{07}\text{StudentAge}_j +$
$\gamma_{08}\text{ClassroomCondition}_j \text{ x StudentSex}_{ij} + \gamma_{09}\text{DeskCondition}_j \text{ x StudentSex}_{ij} +$
$\gamma_{11}\text{Toilets}_j \text{ x StudentSex}_{ij} + \gamma_{12}\text{Textbooks}_j \text{ x StudentSex}_{ij} +$
$\gamma_{13}\text{ExcerciseBooks}_j \text{ x StudentSex}_{ij} + \gamma_{14}\text{CornerLibraries}_j \text{ x StudentSex}_{ij} +$
$\gamma_{15}\text{StudentAge}_j \text{ x StudentSex}_{ij} + u_{0j}$

$\beta_{1j} = \gamma_{20} + \gamma_{21}\text{ClassroomCondition}_j + \gamma_{21}\text{DeskCondition}_j + \gamma_{22}\text{Toilets}_j +$ (36)
$\gamma_{23}\text{Textbooks}_j + \gamma_{24}\text{ExcerciseBooks}_j + \gamma_{25}\text{CornerLibraries}_j + \gamma_{26}\text{StudentAge}_j$
$+ u_{1j}$

$\eta_{ij} = \text{logit}[\pi_{kij}(Y > K)] = (-a_k + u_{0j}) + (\gamma_{01}\text{ClassroomCondition}_j + \gamma_{02}\text{DeskCondition}_j +$ (37)
$\gamma_{03}\text{Toilets}_j + \gamma_{04}\text{Textbooks}_j + \gamma_{05}\text{ExcerciseBooks}_j + \gamma_{06}\text{CornerLibraries}_j +$
$\gamma_{07}\text{StudentAge}_j + \gamma_{20}\text{StudentSex}_{ij} + \gamma_{08}\text{ClassroomCondition}_j \text{ x StudentSex}_{ij} +$
$\gamma_{09}\text{DeskCondition}_j \text{ x StudentSex}_{ij} + \gamma_{10}\text{Toilets}_j \text{ x StudentSex}_{ij} +$
$\gamma_{11}\text{Textbooks}_j \text{ x StudentSex}_{ij} + \gamma_{12}\text{ExcerciseBooks}_j \text{ x StudentSex}_{ij} +$

γ_{13}CornerLibraries$_j$ x StudentSex$_{ij}$ + γ_{14}StudentAge$_j$ x StudentSex$_{ij}$ + u_1jStudentSex$_{ij}$)

Where:

Y_{ij} = inverse log odds of performing at higher academic achievement category for student *i* in school *j*

β_{0j} = random intercept for the *j*-th school (academic achievement for student *i* in school *j* when all parameters = 0)

γ_{01} = log odds of classrooms in bad condition (%) for student *i* in school *j*

γ_{02} = log odds of desks in broken condition (%) for student *i* in school *j*

γ_{03} = log odds of presence of at least one toilet for student *i* in school *j*

γ_{04} = log odds of textbooks available (%) for student *i* in school *j*

γ_{05} = log odds of exercise books available (%) for student *i* in school *j*

γ_{06} = log odds of presence of at least one corner library for student *i* in school *j*

γ_{07} = log odds of age for student *i* in school *j*

γ_{20} = log odds of biological sex for student *i* in school *j*

γ_{08} = log odds of classroom condition and students sex interaction for student *i* in school *j*

γ_{09} = log odds of desk condition and students sex interaction for student *i* in school *j*

γ_{10} = log odds of toilet presence and students sex interaction for student *i* in school *j*

γ_{11} = log odds of textbook availability and student sex interaction for student *i* in school *j*

γ_{12} = log odds of exercise book availability and student sex interaction for student *i* in school *j*

γ_{13} = log odds of corner library presence and students sex interaction for student *i* in school *j*

γ_{14} = log odds of student age and sex interaction for student *i* in school *j*

u_{0j} = fixed intercept for all parameters at level-2

u_{1j} = random slope of biological sex for student *i* in school *j*

Appendix H: Two-Level Ordered Logistic Models (1-4)

Figure 23

TOLR Model 1: Unconditional Null Hypothesis

```
. meologit tot_pri_ord_30_pct if e(sample)|| school_name_pri: || child_name_pri:

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3469.8018

Refining starting values:

Grid node 0:   log likelihood = -3033.3407

Fitting full model:

Iteration 0:   log likelihood = -3033.3407  (not concave)

Iteration 59:  log likelihood = -2802.4663

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                 chi2()           =          .
Log likelihood = -2802.4663                      Prob > chi2      =          .
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
/cut1	-1.131743	.2006385			-1.524987	-.7384983
/cut2	1.59398	.2013463			1.199348	1.988611
school_name_pri						
var(_cons)	3.645683	.5945761			2.648232	5.018821
school_name_pri>child_name_pri						
var(_cons)	1.94e-33	4.69e-18			.	.

```
LR test vs. ologit model: chibar2(01) = 1334.67       Prob >= chibar2 = 0.0000
```

Figure 23

TOLR Model 1: Unconditional Null Hypothesis (Odds Ratio Version)

```
. meologit, or

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                 chi2()           =          .
Log likelihood = -2802.4663                      Prob > chi2      =          .
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
/cut1	-1.131743	.2006385			-1.524987	-.7384983
/cut2	1.59398	.2013463			1.199348	1.988611
school_name_pri						
var(_cons)	3.645683	.5945761			2.648232	5.018821
school_name_pri>child_name_pri						
var(_cons)	1.94e-33	4.69e-18			.	.

```
LR test vs. ologit model: chibar2(01) = 1334.67       Prob >= chibar2 = 0.0000
```

Figure 24

TOLR Model 2: Random-Covariates Model with Four Fixed Parameters

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official desks_pct_pri i.toilets_pri stu_txtbks_pri_per_student if e(sample)|| school_name_
> pri: || child_name_pri:

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3205.8393
Iteration 2:   log likelihood = -3202.4959
Iteration 3:   log likelihood = -3202.4861
Iteration 4:   log likelihood = -3202.4861

Refining starting values:

Grid node 0:   log likelihood = -2911.4868

Fitting full model:

Iteration 0:   log likelihood = -2911.4868  (not concave)

Iteration 56:  log likelihood = -2773.2393  (backed up)

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(4)      =      61.94
Log likelihood = -2773.2393                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.0075776	.0026931	-2.81	0.005	-.0128559	-.0022993
desks_pct_pri	-.0092808	.0038572	-2.41	0.016	-.0168409	-.0017208
1.toilets_pri	.1674831	.2221073	0.75	0.451	-.2678393	.6028054
stu_txtbks_pri_per_student	.011673	.0019657	5.94	0.000	.0078204	.0155256
/cut1	-.8726621	.2947889			-1.450438	-.2948865
/cut2	1.889077	.2963972			1.308149	2.470005
school_name_pri						
var(_cons)	2.514743	.4302831			1.798247	3.51672
school_name_pri>child_name_pri						
var(_cons)	3.99e-33	1.27e-17			.	.

```
LR test vs. ologit model: chibar2(01) = 858.49        Prob >= chibar2 = 0.0000
```

Figure 25

TOLR Model 2: Random-Covariates Model with Four Fixed Parameters (Odds Ratio Version)

```
. meologit, or

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(4)      =      61.94
Log likelihood = -2773.2393                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	.9924511	.0026727	-2.81	0.005	.9872264	.9977034
desks_pct_pri	.9907621	.0038216	-2.41	0.016	.9833002	.9982806
1.toilets_pri	1.182325	.2626031	0.75	0.451	.7650307	1.827238
stu_txtbks_pri_per_student	1.011741	.0019887	5.94	0.000	1.007851	1.015647
/cut1	-.8726621	.2947889			-1.450438	-.2948865
/cut2	1.889077	.2963972			1.308149	2.470005
school_name_pri						
var(_cons)	2.514743	.4302831			1.798247	3.51672
school_name_pri>child_name_pri						
var(_cons)	3.99e-33	1.27e-17			.	.

```
Note: Estimates are transformed only in the first equation.
LR test vs. ologit model: chibar2(01) = 858.49        Prob >= chibar2 = 0.0000
```

Figure 26

TOLR Model 3: Random-Intercepts Model with Six Fixed Parameters

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official desks_pct_pri i.toilets_pri stu_txtbks_pri_per_student stu_exbks_pri_per_student :
> .corner_lib_pri if e(sample)|| school_name_pri: || child_name_pri:

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3203.0276
Iteration 2:   log likelihood = -3199.5718
Iteration 3:   log likelihood = -3199.5613
Iteration 4:   log likelihood = -3199.5613

Refining starting values:

Grid node 0:   log likelihood = -2911.2528

Fitting full model:

Iteration 0:   log likelihood = -2911.2528  (not concave)

Iteration 55:  log likelihood = -2768.9181  (backed up)

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(6)      =      70.74
Log likelihood = -2768.9181                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.0087714	.002776	-3.16	0.002	-.0142123	-.0033305
desks_pct_pri	-.0113939	.0039604	-2.88	0.004	-.0191561	-.0036317
1.toilets_pri	.165344	.2239692	0.74	0.460	-.2736275	.6043154
stu_txtbks_pri_per_student	.0136201	.0020819	6.54	0.000	.0095397	.0177005
stu_exbks_pri_per_student	-.0053055	.0031617	-1.68	0.093	-.0115024	.0008913
1.corner_lib_pri	-.5167931	.2143307	-2.41	0.016	-.9368736	-.0967127
/cut1	-1.440016	.4076548			-2.239005	-.6410272
/cut2	1.331081	.4059479			.5354382	2.126725
school_name_pri						
var(_cons)	2.617693	.4490227			1.87029	3.663773
school_name_pri>child_name_pri						
var(_cons)	3.26e-17	.0000292			.	.

```
LR test vs. ologit model: chi2(2) = 861.29                Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

Figure 27

TOLR Model 3: Random-Intercepts Model with Six Fixed Parameters (Odds Ratio Version)

```
. meologit, or

Mixed-effects ologit regression               Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite               Integration pts.  =          7

                                              Wald chi2(6)      =      70.74
Log likelihood = -2768.9181                   Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	.991267	.0027518	-3.16	0.002	.9858882	.996675
desks_pct_pri	.9886708	.0039155	-2.88	0.004	.9810262	.9963749
1.toilets_pri	1.179799	.2642385	0.74	0.460	.7606153	1.829999
stu_txtbks_pri_per_student	1.013713	.0021104	6.54	0.000	1.009585	1.017858
stu_exbks_pri_per_student	.9947085	.003145	-1.68	0.093	.9885635	1.000892
1.corner_lib_pri	.5964302	.1278333	-2.41	0.016	.391851	.9078168
/cut1	-1.440016	.4076548			-2.239005	-.6410272
/cut2	1.331081	.4059479			.5354382	2.126725
school_name_pri						
var(_cons)	2.617693	.4490227			1.87029	3.663773
school_name_pri>child_name_pri						
var(_cons)	3.26e-17	.0000292			.	.

```
Note: Estimates are transformed only in the first equation.
LR test vs. ologit model: chi2(2) = 861.29                 Prob > chi2 = 0.0000
```

Figure 28

TOLR Model 4: Random-Covariates Model with Seven Fixed Parameters & One Random Parameter

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official desks_pct_pri i.toilets_pri stu_txtbks_pri_per_student stu_exbks_pri_per_student i
> .corner_lib_pri age if e(sample)|| school_name_pri: || child_name_pri: gender, from(XY, skip)

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3183.6193
Iteration 2:   log likelihood = -3179.6597
Iteration 3:   log likelihood = -3179.6466
Iteration 4:   log likelihood = -3179.6466

Refining starting values:

Grid node 0:   log likelihood = -3645.5864

Fitting full model:

Iteration 0:   log likelihood = -3645.5864  (not concave)

Iteration 113: log likelihood = -2762.7286  (backed up)

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(7)      =      83.59
Log likelihood = -2762.7286                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.0092634	.0027449	-3.37	0.001	-.0146433	-.0038835
desks_pct_pri	-.0118168	.0039371	-3.00	0.003	-.0195333	-.0041002
1.toilets_pri	.1086208	.2228583	0.49	0.626	-.3281734	.545415
stu_txtbks_pri_per_student	.0134043	.002075	6.46	0.000	.0093374	.0174712
stu_exbks_pri_per_student	-.0054425	.0031487	-1.73	0.084	-.0116138	.0007288
1.corner_lib_pri	-.5577741	.2141234	-2.60	0.009	-.9774482	-.1381
age	-.0957947	.0272638	-3.51	0.000	-.1492308	-.0423586
/cut1	-2.568222	.516812			-3.581155	-1.555289
/cut2	.212817	.5127257			-.7921069	1.217741
school_name_pri						
var(_cons)	2.556256	.4396638			1.824743	3.581022
school_name_pri>child_name_pri						
var(gender)	2.38e-41	1.51e-23			.	.
var(_cons)	1.69e-33	4.92e-18			.	.

```
LR test vs. ologit model: chibar2(01) = 833.84        Prob >= chibar2 = 0.0000
```

Figure 29

TOLR Model 4: Random-Covariates Model with Seven Fixed Parameters & One Random Parameter (Odds Ratio Version)

```
. meologit, or

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(7)      =      83.59
Log likelihood = -2762.7286                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	.9907794	.0027196	-3.37	0.001	.9854634	.996124
desks_pct_pri	.9882528	.0038909	-3.00	0.003	.9806562	.9959082
1.toilets_pri	1.11474	.2484289	0.49	0.626	.7202381	1.725324
stu_txtbks_pri_per_student	1.013495	.002103	6.46	0.000	1.009381	1.017625
stu_exbks_pri_per_student	.9945723	.0031316	-1.73	0.084	.9884534	1.000729
1.corner_lib_pri	.5724819	.1225818	-2.60	0.009	.37627	.8710116
age	.9086505	.0247733	-3.51	0.000	.8613703	.958526
/cut1	-2.568222	.516812			-3.581155	-1.555289
/cut2	.212817	.5127257			-.7921069	1.217741
school_name_pri						
var(_cons)	2.556256	.4396638			1.824743	3.581022
school_name_pri>child_name_pri						
var(gender)	2.38e-41	1.51e-23			.	.
var(_cons)	1.69e-33	4.92e-18			.	.

```
Note: Estimates are transformed only in the first equation.
LR test vs. ologit model: chibar2(01) = 833.84        Prob >= chibar2 = 0.0000
```

Figure 30

TOLR Model 5: Random-Covariates Model with Cross-Level Interactions

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official desks_pct_pri i.toilets_pri stu_txtbks_pri_per_student
>  stu_exbks_pri_per_student i.corner_lib_pri age gender_class_or gender_desks_or gender_toilets_or gender_texts_or
> gender_exbks_or gender_lib_or gender_age_or if e(sample) || school_name_pri: || child_name_pri: gender, from(XY, s
> kip)

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3178.7048
Iteration 2:   log likelihood = -3174.6307
Iteration 3:   log likelihood = -3174.6177
Iteration 4:   log likelihood = -3174.6177

Refining starting values:

Grid node 0:   log likelihood = -3643.7379

Fitting full model:

Iteration 0:   log likelihood = -3643.7379  (not concave)
Iteration 121: log likelihood =  -2759.797  (backed up)

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(14)     =      89.30
Log likelihood =  -2759.797                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.008222	.0029753	-2.76	0.006	-.0140535	-.0023905
desks_pct_pri	-.0142687	.0043377	-3.29	0.001	-.0227705	-.0057669
1.toilets_pri	.1314131	.243578	0.54	0.590	-.3459909	.6088171
stu_txtbks_pri_per_student	.0125689	.0023323	5.39	0.000	.0079977	.0171401
stu_exbks_pri_per_student	-.0071329	.0035207	-2.03	0.043	-.0140334	-.0002325
1.corner_lib_pri	-.511195	.2342116	-2.18	0.029	-.9702412	-.0521488
age	-.0810361	.0299419	-2.71	0.007	-.1397212	-.022351
gender_class_or	-.0019936	.0023138	-0.86	0.389	-.0065286	.0025414
gender_desks_or	.0034605	.0030771	1.12	0.261	-.0025706	.0094915
gender_toilets_or	-.0257217	.183911	-0.14	0.889	-.3861807	.3347373
gender_texts_or	.0015456	.0020615	0.75	0.453	-.0024949	.0055861
gender_exbks_or	.0031912	.0030322	1.05	0.293	-.0027518	.0091342
gender_lib_or	-.0777522	.1935349	-0.40	0.688	-.4570736	.3015691
gender_age_or	-.0237755	.0248132	-0.96	0.338	-.0724084	.0248574
/cut1	-2.552197	.5179119			-3.567286	-1.537109
/cut2	.2326801	.5138485			-.7744445	1.239805
school_name_pri						
var(_cons)	2.55563	.4395523			1.824302	3.580134
school_name_pri>child_name_pri						
var(gender)	1.66e-33	8.75e-18			.	.
var(_cons)	2.36e-33	5.28e-18			.	.

```
LR test vs. ologit model: chibar2(01) = 829.64        Prob >= chibar2 = 0.0000
```

Figure 31

TOLR Model 5: Random-Covariates Model with Cross-Level Interactions (Odds Ratio Version)

```
. meologit, or

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(14)     =      89.30
Log likelihood =  -2759.797                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	.9918117	.002951	-2.76	0.006	.9860448	.9976123
desks_pct_pri	.9858326	.0042763	-3.29	0.001	.9774868	.9942497
1.toilets_pri	1.140439	.2777858	0.54	0.590	.7075189	1.838256
stu_txtbks_pri_per_student	1.012648	.0023618	5.39	0.000	1.00803	1.017288
stu_exbks_pri_per_student	.9928924	.0034957	-2.03	0.043	.9860646	.9997675
1.corner_lib_pri	.5997784	.140475	-2.18	0.029	.3789916	.9491876
age	.9221604	.0276112	-2.71	0.007	.8696007	.9778969
gender_class_or	.9980084	.0023092	-0.86	0.389	.9934927	1.002545
gender_desks_or	1.003466	.0030878	1.12	0.261	.9974327	1.009537
gender_toilets_or	.9746063	.1792408	-0.14	0.889	.6796477	1.397573
gender_texts_or	1.001547	.0020647	0.75	0.453	.9975082	1.005602
gender_exbks_or	1.003196	.0030419	1.05	0.293	.997252	1.009176
gender_lib_or	.9251936	.1790572	-0.40	0.688	.6331338	1.351979
gender_age_or	.9765049	.0242302	-0.96	0.338	.9301509	1.025169
/cut1	-2.552197	.5179119			-3.567286	-1.537109
/cut2	.2326801	.5138485			-.7744445	1.239805
school_name_pri						
var(_cons)	2.55563	.4395523			1.824302	3.580134
school_name_pri>child_name_pri						
var(gender)	1.66e-33	8.75e-18			.	.
var(_cons)	2.36e-33	5.28e-18			.	.

```
Note: Estimates are transformed only in the first equation.
LR test vs. ologit model: chibar2(01) = 829.64        Prob >= chibar2 = 0.0000
```

Figure 32

TOLR Likelihood Ratio Tests M1-M5

```
. lrtest M1_M_30 M2_M_30

Likelihood-ratio test                                 LR chi2(7)  =     67.10
(Assumption: M1_M_30 nested in M2_M_30)               Prob > chi2 =    0.0000

. lrtest M2_M_30 M3_M_30

Likelihood-ratio test                                 LR chi2(4)  =     42.60
(Assumption: M3_M_30 nested in M2_M_30)               Prob > chi2 =    0.0000

. lrtest M3_M_30 M4_M_30

Likelihood-ratio test                                 LR chi2(4)  =     54.98
(Assumption: M3_M_30 nested in M4_M_30)               Prob > chi2 =    0.0000

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space.  If this is not true, then
      the reported test is conservative.

. lrtest M4_M_30 M5_M_30

Likelihood-ratio test                                 LR chi2(7)  =      5.86
(Assumption: M4_M_30 nested in M5_M_30)               Prob > chi2 =    0.5558

. lrtest M1_M_30 M3_M_30

Likelihood-ratio test                                 LR chi2(3)  =     24.50
(Assumption: M1_M_30 nested in M3_M_30)               Prob > chi2 =    0.0000

. lrtest M1_M_30 M4_M_30

Likelihood-ratio test                                 LR chi2(7)  =     79.48
(Assumption: M1_M_30 nested in M4_M_30)               Prob > chi2 =    0.0000

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space.  If this is not true, then
      the reported test is conservative.
```

The likelihood ratio tests illustrate that model 4, a random-covariates two-level ordered logistic model with seven fixed parameters and one random parameter, appears to display the best model fit with respect to the remainder four other models.

Appendix I: Two-Level Ordered Logistic Models (3.1-3.6)

This section illustrates the formation of Model 3 (i.e. the random-intercepts multilevel logit model) from the introduction of the first fixed parameter to the internalization of all six parameters to garner a better understanding of how individual educational inputs contribute to the model. Evident in the Stata outputs below, the constructed modified versions of M3, namely M3.1 to M3.6, depict the logit estimates of adding one parameter at a time. The concluding portion of this section presents the likelihood ratio tests to compare the goodness of fit of the models, suggesting that M3.4 relatively has the best model.

Table 2

Parameter-by-Parameter Construction of Model 3

Variables	M2.1	M2.2	M2.3	M2.4	M2.5	M2.6
Fixed Effect Estimates						
Threshold 1 (a_1)	-1.426	-1.567	-1.027	-0.873	-1.331	-1.440
	(0.208)	(0.214)	(0.311)	(0.295)	(0.404)	(0.408)
Threshold 2 (a_2)	1.308	1.169	1.714	1.889	1.435	1.331
	(0.207)	(0.213)	(0.312)	(0.296)	(0.403)	(0.406)
Classroom Condition (γ_1)	0.009***	0.009***	0.006**	0.008***	0.007***	0.009***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Desk Condition (γ_2)		0.008**	0.004	0.009**	0.010**	0.011***
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Presence of Toilets (γ_3)			0.528**	0.167	0.15	0.165
			(0.221)	(0.222)	(0.223)	(0.224)
Textbooks (γ_4)				0.012***	0.012***	0.014***
				(0.002)	(0.002)	(0.002)
Exercise Books (γ_5)					-0.005*	-0.005*
					(0.003)	(0.003)
Corner Libraries (γ_6)						-0.517**
						(0.214)
Random Effect Estimates						
School-Level Intercept (t_{00})	3.322	3.205	3.148	2.515	2.600	2.618
	0.549	0.532	0.525	0.430	0.448	0.449
Student-Level Intercept (t_{01})	1.18E-38	2.43E-36	3.47E-34	3.99E-33	9.26E-36	3.26E-17
	9.27E-22	3.47E-19	1.59E-18	1.27E-17	3.84E-19	0.0000292
Model Fit Summary						
LR Test (χ^2)	1201.32***	1159.59***	1127.62***	858.49***	855.8***	861.29***
ICC	0.502	0.494	0.489	0.433	0.441	0.443

AIC	5599.09	5596.13	5592.44	5560.48	5559.663	5557.836
BIC	5623.42	5626.54	5628.93	5603.05	5608.319	5618.657
Degrees of Freedom	4	5	6	7	8	10
Observations	3,236	3,236	3,236	3,236	3,236	3,236
Number of groups	97	97	97	97	97	97

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 33

TOLR Model 3.1: Random-Intercepts Model with One Fixed Parameter

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official if e(sample)|| school_name_pri: || child_name_pri:

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3396.4553
Iteration 2:   log likelihood = -3396.2029
Iteration 3:   log likelihood = -3396.2029

Refining starting values:

Grid node 0:   log likelihood = -2998.1169

Fitting full model:

Iteration 0:   log likelihood = -2998.1169  (not concave)

Iteration 55:  log likelihood = -2795.5451  (backed up)

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(1)      =      14.13
Log likelihood = -2795.5451                     Prob > chi2       =     0.0002
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.009295	.0024727	-3.76	0.000	-.0141413	-.0044487
/cut1	-1.425578	.2077553			-1.832771	-1.018385
/cut2	1.308064	.2069057			.9025361	1.713592
school_name_pri						
var(_cons)	3.322273	.5489419			2.403204	4.592827
school_name_pri>child_name_pri						
var(_cons)	1.18e-38	9.27e-22			.	.

```
LR test vs. ologit model: chibar2(01) = 1201.32       Prob >= chibar2 = 0.0000
```

Figure 34

TOLR Model 3.2: Random-Intercepts Model with Two Fixed Parameters

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official desks_pct_pri if e(sample)|| school_name_pri: || child_name_pri:

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3373.2634
Iteration 2:   log likelihood = -3372.8597
Iteration 3:   log likelihood = -3372.8596

Refining starting values:

Grid node 0:   log likelihood =  -2987.138

Fitting full model:

Iteration 0:   log likelihood =  -2987.138  (not concave)

Iteration 65:  log likelihood = -2793.0632

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(2)      =      19.25
Log likelihood = -2793.0632                     Prob > chi2       =     0.0001
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.0088129	.0024733	-3.56	0.000	-.0136605	-.0039653
desks_pct_pri	-.0078551	.0035159	-2.23	0.025	-.0147462	-.000964
/cut1	-1.567078	.2144749			-1.987441	-1.146715
/cut2	1.169474	.2128125			.7523689	1.586578
school_name_pri						
var(_cons)	3.204889	.5316603			2.315298	4.436282
school_name_pri>child_name_pri						
var(_cons)	2.43e-36	3.47e-19			.	.

```
LR test vs. ologit model: chibar2(01) = 1159.59       Prob >= chibar2 = 0.0000
```

Figure 35

TOLR Model 3.3: Random-Intercepts Model with Three Fixed Parameters

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official desks_pct_pri i.toilets_pri if e(sample)|| school_name_pri: || child_name_pri:

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3354.6293
Iteration 2:   log likelihood = -3354.0272
Iteration 3:   log likelihood = -3354.0269

Refining starting values:

Grid node 0:   log likelihood = -2978.4767

Fitting full model:

Iteration 0:   log likelihood = -2978.4767  (not concave)
Iteration 59:  log likelihood = -2790.2178

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(3)      =      24.85
Log likelihood = -2790.2178                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.0058818	.0027711	-2.12	0.034	-.011313	-.0004506
desks_pct_pri	-.003898	.0038971	-1.00	0.317	-.0115361	.0037401
1.toilets_pri	.5283843	.2207497	2.39	0.017	.0957228	.9610457
/cut1	-1.027016	.3111433			-1.636845	-.4171861
/cut2	1.714449	.3124032			1.102151	2.326748
school_name_pri						
var(_cons)	3.148316	.52457			2.271181	4.364202
school_name_pri>child_name_pri						
var(_cons)	3.47e-34	1.59e-18			.	.

```
LR test vs. ologit model: chibar2(01) = 1127.62       Prob >= chibar2 = 0.0000
```

Figure 36

TOLR Model 3.4: Random-Intercepts Model with Four Fixed Parameters

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official desks_pct_pri i.toilets_pri stu_txtbks_pri_per_student if e(sample)|| school_name_
> pri: || child_name_pri:

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3205.8393
Iteration 2:   log likelihood = -3202.4959
Iteration 3:   log likelihood = -3202.4861
Iteration 4:   log likelihood = -3202.4861

Refining starting values:

Grid node 0:   log likelihood = -2911.4868

Fitting full model:

Iteration 0:   log likelihood = -2911.4868  (not concave)

Iteration 56:  log likelihood = -2773.2393  (backed up)

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(4)      =      61.94
Log likelihood = -2773.2393                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.0075776	.0026931	-2.81	0.005	-.0128559	-.0022993
desks_pct_pri	-.0092808	.0038572	-2.41	0.016	-.0168409	-.0017208
1.toilets_pri	.1674831	.2221073	0.75	0.451	-.2678393	.6028054
stu_txtbks_pri_per_student	.011673	.0019657	5.94	0.000	.0078204	.0155256
/cut1	-.8726621	.2947889			-1.450438	-.2948865
/cut2	1.889077	.2963972			1.308149	2.470005
school_name_pri						
var(_cons)	2.514743	.4302831			1.798247	3.51672
school_name_pri>child_name_pri						
var(_cons)	3.99e-33	1.27e-17			.	.

```
LR test vs. ologit model: chibar2(01) = 858.49        Prob >= chibar2 = 0.0000
```

Figure 37

TOLR Model 3.5: Random-Intercepts Model with Five Fixed Parameters

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official desks_pct_pri i.toilets_pri stu_txtbks_pri_per_student stu_exbks_pri_per_student i
> f e(sample)|| school_name_pri: || child_name_pri:

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3203.1941
Iteration 2:   log likelihood = -3199.7423
Iteration 3:   log likelihood = -3199.7318
Iteration 4:   log likelihood = -3199.7318

Refining starting values:

Grid node 0:   log likelihood = -2911.7903

Fitting full model:

Iteration 0:   log likelihood = -2911.7903  (not concave)

Iteration 51:  log likelihood = -2771.8314  (backed up)

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(5)      =      64.27
Log likelihood = -2771.8314                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.0073597	.0027108	-2.71	0.007	-.0126728	-.0020466
desks_pct_pri	-.0095312	.0038805	-2.46	0.014	-.0171369	-.0019255
1.toilets_pri	.1496128	.2232421	0.67	0.503	-.2879336	.5871592
stu_txtbks_pri_per_student	.0121374	.0019952	6.08	0.000	.0082268	.016048
stu_exbks_pri_per_student	-.0053047	.0031626	-1.68	0.093	-.0115032	.0008939
/cut1	-1.330725	.404114			-2.122774	-.5386762
/cut2	1.434928	.4028491			.6453583	2.224498
school_name_pri						
var(_cons)	2.59957	.4480188			1.854396	3.644186
school_name_pri>child_name_pri						
var(_cons)	9.26e-36	3.84e-19			.	.

```
LR test vs. ologit model: chibar2(01) = 855.80        Prob >= chibar2 = 0.0000
```

Figure 38

TOLR Model 3.6: Random-Intercepts Model with Six Fixed Parameters

```
. meologit tot_pri_ord_30_pct classroom_bad__pct_pri_official desks_pct_pri i.toilets_pri stu_txtbks_pri_per_student stu_exbks_pri_per_student i
> .corner_lib_pri if e(sample)|| school_name_pri: || child_name_pri:

Fitting fixed-effects model:

Iteration 0:   log likelihood = -3469.8018
Iteration 1:   log likelihood = -3203.0276
Iteration 2:   log likelihood = -3199.5718
Iteration 3:   log likelihood = -3199.5613
Iteration 4:   log likelihood = -3199.5613

Refining starting values:

Grid node 0:   log likelihood = -2911.2528

Fitting full model:

Iteration 0:   log likelihood = -2911.2528  (not concave)
Iteration 55:  log likelihood = -2768.9181  (backed up)

Mixed-effects ologit regression                 Number of obs     =      3,236
```

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
school_nam~i	97	8	33.4	532
child_name~i	2,976	1	1.1	4

```
Integration method: mvaghermite                 Integration pts.  =          7

                                                Wald chi2(6)      =      70.74
Log likelihood = -2768.9181                     Prob > chi2       =     0.0000
```

tot_pri_ord_30_pct	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
classroom_bad__pct_pri_official	-.0087714	.002776	-3.16	0.002	-.0142123	-.0033305
desks_pct_pri	-.0113939	.0039604	-2.88	0.004	-.0191561	-.0036317
1.toilets_pri	.165344	.2239692	0.74	0.460	-.2736275	.6043154
stu_txtbks_pri_per_student	.0136201	.0020819	6.54	0.000	.0095397	.0177005
stu_exbks_pri_per_student	-.0053055	.0031617	-1.68	0.093	-.0115024	.0008913
1.corner_lib_pri	-.5167931	.2143307	-2.41	0.016	-.9368736	-.0967127
/cut1	-1.440016	.4076548			-2.239005	-.6410272
/cut2	1.331081	.4059479			.5354382	2.126725
school_name_pri						
var(_cons)	2.617693	.4490227			1.87029	3.663773
school_name_pri>child_name_pri						
var(_cons)	3.26e-17	.0000292			.	.

```
LR test vs. ologit model: chi2(2) = 861.29                Prob > chi2 = 0.0000
```

Figure 39

TOLR Likelihood Ratio Tests M3.1-M3.6

```
. lrtest M2_M_30 M3_M1_30

Likelihood-ratio test                                 LR chi2(6)  =     53.25
(Assumption: M3_M1_30 nested in M2_M_30)              Prob > chi2 =    0.0000

. lrtest M3_M1_30 M3_M2_30

Likelihood-ratio test                                 LR chi2(1)  =      4.96
(Assumption: M3_M1_30 nested in M3_M2_30)             Prob > chi2 =    0.0259

. lrtest M3_M1_30 M3_M3_30

Likelihood-ratio test                                 LR chi2(2)  =     10.65
(Assumption: M3_M1_30 nested in M3_M3_30)             Prob > chi2 =    0.0049

. lrtest M3_M3_30 M3_M4_30

Likelihood-ratio test                                 LR chi2(1)  =     33.96
(Assumption: M3_M3_30 nested in M3_M4_30)             Prob > chi2 =    0.0000

. lrtest M3_M4_30 M3_M5_30

Likelihood-ratio test                                 LR chi2(1)  =      2.82
(Assumption: M3_M4_30 nested in M3_M5_30)             Prob > chi2 =    0.0933

. lrtest M3_M4_30 M3_M6_30

Likelihood-ratio test                                 LR chi2(3)  =      8.64
(Assumption: M3_M4_30 nested in M3_M6_30)             Prob > chi2 =    0.0344

. lrtest M3_M1_30 M3_M4_30

Likelihood-ratio test                                 LR chi2(3)  =     44.61
(Assumption: M3_M1_30 nested in M3_M4_30)             Prob > chi2 =    0.0000

. lrtest M3_M2_30 M3_M3_30

Likelihood-ratio test                                 LR chi2(1)  =      5.69
(Assumption: M3_M2_30 nested in M3_M3_30)             Prob > chi2 =    0.0171
```

The aforementioned likelihood ratio tests suggest that model 3.4, a random-intercepts multilevel ordered logistic model consisting of four educational inputs (classroom condition, desk condition, presence of toilet/s, and textbook availability), serves as the best model relative to the five versions of model 3. Thus, we can further argue that the inclusion of all six parameters in the random intercepts model does not produce a better fit, rather a more restrictive model with these four parameters can be a sufficient alternative.

Appendix J: Intraclass Correlation Coefficient Calculations

The intraclass correlation coefficient (ICC) is a critical statistic that can allow us to, in the context of this study, quantify the degree of homogeneity of academic achievement across schools. However, the standard method of estimating the ICC in multilevel linear regression cannot be applied in a multilevel logistic regression given that the student and school level variances are not directly expressed on the same scale and are, thus, not comparable. Specifically, the school-level residual variance (u_{0j}) follows on a logistic scale, but the student level residual variance (u_{1ij}) is situated on a probability scale. For an accurate representation of the ICC, Snijders and Bosker (1999) put forth an alternative approach, which computes ICC through the conversion of student level variance from a probability scale to a logistic scale, based on the assumption that the propensity of reaching a higher category of academic achievement is a continuous latent variable underlying the ordinal response. Given that the unobserved student variable mirrors a logistic distribution with variance (u_{1ij}) equal to $\pi^2/3$, the ICC can be calculated using the following formula:

$$ICC = \frac{(u0j)}{(u0j) + (\frac{\pi 2}{3})}$$

Using this formula, this study hand-calculated the ICC values of M1-M5 and M3.1-M3.5, the results of which can be illustrated in Table 3 below.

Table 3

ICC Calculations Based on the Linear Threshold Model Method

Models	*Individual Level Variance*	*ICC*
M1	3.646	0.525
M2	2.515	0.433
M3	3.148	0.433
M4	2.556	0.437
M5	2.556	0.437
M3.1	3.322	0.502
M3.2	3.205	0.494

M3.3	3.148	0.489
M3.4	2.515	0.433
M3.5	2.600	0.441
M3.6	2.618	0.443

Appendix K: Missing Values

Figure 40

Missing Values in Two-Level Ordinal Logit Model 4

```
 * Missing Values
 drop if missing(tot_pct_pri_ord)
0 observations deleted)

 drop if missing(classroom_bad__pct_pri_official)
243 observations deleted)

 drop if missing(desks_pct_pri)
296 observations deleted)

 drop if missing(toilets_pri)
0 observations deleted)

 drop if missing(stu_txtbks_pri_per_student)
0 observations deleted)

 drop if missing(stu_exbks_pri_per_student)
27 observations deleted)

 drop if missing(corner_lib_pri)
58 observations deleted)

 drop if missing(age)
38 observations deleted)

 drop if missing(gender)
7 observations deleted)
```

Appendix L: English, Mathematics, and Science Assessments



Select the school level	◯ Primary ◯ JSS
Every school that is visited for this exercise must be from the list of randomly selected schools. Permission to visit any replacement school must come directly from FME. Once FME receives the request, you will receive the name of the replacement school. If a replacement school is chosen, please choose 'Other' for the State and then 'Other' again for LGA and then write down the details.	
State	◯ Abia ◯ Akwa-Ibom ◯ Borno ◯ Cross-River ◯ Edo ◯ Jigawa ◯ Katsina ◯ Kebbi ◯ Ondo ◯ Plateau ◯ Rivers ◯ Taraba ◯ Zamfara ◯ Bauchi ◯ Other
LGA	◯ Other ◯ Ohafia ◯ Ini

1

Nigeria Student Learning Assessment

What assessment are you conducting?	◯ English ◯ Maths ◯ Science
Please press the GPS button	
Child ID (Usually, 1 - 10 for English, 11 - 20 for Maths and 21 - 30 for Science) *However, if a child has already taken another assessment today, please use that ID again*	
Has the child already taken another assessment today?	◯ Yes ◯ No
Child's name	
Is this student from the teacher's P4 class who is being interviewed today?	◯ Yes ◯ No
Child's age	
Child's gender	◯ Male ◯ Female ◯ No Response
Ask the child - Did you have breakfast this morning?	◯ Yes ◯ No

Instructions

Instructions for the enumerator :
You will administer the assessment verbally to the learners one at a time. You should read the instructions printed below to the learner, but do not help them to answer the question. You should read from this tablet. The learners will either give their answers orally or write their answers on paper. There are three assessments: English, mathematics and science. The learner has thirty five (35) minutes to complete each assessment. You should speak in English. The learner must answer in English but can ask for clarity in any other language. Before you begin, explain the assessment to the learner. Below is the introduction that will be used with each learner. Present the introduction as written to each learner.

Nigeria Student Learning Assessment

Instructions to be given to the learner :	
Good morning/afternoon! My name is _____. I'm happy to meet you! We are here to find out about the things you are learning at school, so thank you very much for helping me with that. Remember that this is not an exam or a test. You don't need to worry if you don't know some things. If you don't want to do this, you do not have to, and you can go back to your classroom at any time. There is nothing bad that can happen to you because of this exercise. Your answers will not be shared with your teacher, parents or anyone else. I will read questions for you and you can point to the correct answer in some instances or say the correct answer when you are asked. You may be asked to write a word or a sentence. *(Present answer sheet to the learner)*	
Have you understood what I have explained and that your answers will not be shown to your parents, teacher, friends or anyone in your school?	◯ Yes ◯ No
Are you happy to answer some questions about yourself and your family?	◯ Yes ◯ No
Are you happy to take the assessment in the way that I have explained?	◯ Yes ◯ No
Primary English	
If the child does not give consent, please take him/her back to the classroom and replace the child with another randomly selected child.	
1. Enumerator : Ask the learner to identify three letters correctly. For example, "Can you show me the letter 'e' please?" *Enumerator note : You must say this in English. Do not point out the letters to the learner. Repeat this 3 times picking different letters from the square. Record the number of letters correctly identified.*	**l** **e** **h** **f** **u** **v** **a** **s** **c**
How many letters did the child identify correctly?	

2. Enumerator : Ask the learner to identify three words correctly. For example, "Can you show me the word 'greet' please?"

Enumerator note : You must say this in English. Do not point out the words to the learner. Repeat this 3 times picking different words from the square. Record the number of words correctly identified.

under	**respect**	**story**
bananas	**green**	**greet**
fruit	**father**	**outside**

How many words did the child identify correctly?

3. Enumerator : Point at the first picture and ask the learner "What is this?" Repeat for all other pictures.

Enumerator note : Record the number of pictures correctly named.

Bucket/pail, shirt/t-shirt are all correct responses



How many pictures did the child correctly identify correctly?

4. Enumerator: Ask the learner to read this sentence. Follow the sentence as the child reads it. (10 words)

Enumerator note : Select the words that are not read correctly below.

- ☐ When
- ☐ I
- ☐ grow
- ☐ up,
- ☐ I
- ☐ want
- ☐ to
- ☐ be
- ☐ a
- ☐ teacher.

How many words did the child read correctly?

How many seconds did the child take?

Reading Passage

5. Enumerator: Please ask the learner to read the passage out loud. Please remember to start the timer as soon as the child starts reading. (58 words)

Enumerator note : Select the words that are not read correctly below.

- ☐ Sam
- ☐ walked
- ☐ into
- ☐ a
- ☐ shop.
- ☐ He

	☐ cried ☐ Sam. ☐ "You ☐ said ☐ your ☐ dog ☐ doesn't ☐ bite." ☐ "But ☐ that's ☐ not ☐ my ☐ dog!" ☐ replied ☐ John.
How many words did the child read correctly?	
How many seconds did the child take?	
How proficient was the learner's reading?	◯ Very proficient ◯ Somewhat proficient ◯ Not at all proficient
Enumerator: Read the questions to the learner. Tell them that they are allowed to look at the reading text to find the answers. Record their answers. Allow the learner approximately 30 seconds to answer each question. If the learner is unable to answer any question within thirty seconds, move on to the next question.	
6.a. Where did Sam and John meet? *(Correct answer : In a shop)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
6.b. What type of animal was sleeping next to John? *(Correct answer: dog)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
6.c. Why was Sam surprised that the dog bit him?	◯ Because the dog was sleeping ◯ Because John told him that his dog

	does not bite ◯ Because Sam didn't see the dog ◯ Don't know ◯ Not attempted
Listening and Speaking Enumerator: ask the learners to look at pictures in turn, then read the statements and answers. Ask the learner to say the right answer in a full sentence, "He is washing his car because it is dirty." Accept "because it is dirty."	
7. Look at this picture. Mr. David is washing his car. Why is he washing his car? 	◯ He is washing his car because it is clean. ◯ He is washing his car because it is dirty. ◯ He is washing his car because it is broken. ◯ He is washing his car because it is old. ◯ Don't know ◯ Not attempted
8. This is Mrs. Folake. She has put up her umbrella. Why do you think she has put up her umbrella? 	◯ She put up her umbrella because it was raining. ◯ She put up her umbrella because it was dark. ◯ She put up her umbrella because it was morning. ◯ She put up her umbrella because she was pleased. ◯ Don't know ◯ Not attempted
9. This is Joseph. He is greeting his teacher. What time of the day is it?	◯ Evening ◯ Morning ◯ Afternoon ◯ Midnight



- ◯ Don't know
- ◯ Not attempted

10. This is John. He has stepped on Bukola's foot. What did he do?



- ◯ He laughed.
- ◯ He apologised.
- ◯ He ran away.
- ◯ He did not say anything.
- ◯ Don't know
- ◯ Not attempted

11. This is Peter. He gave Agnes a sweet. What did she do?



- ◯ She apologised.
- ◯ She refused it.
- ◯ She took it.
- ◯ She was angry.
- ◯ Don't know
- ◯ Not attempted

12. This is Mr. John. He is leaving his friend's house. What time of day is it?

- ◯ Morning
- ◯ Afternoon
- ◯ Night time
- ◯ Monday



- ◯ Don't know
- ◯ Not attempted

10. This is John. He has stepped on Bukola's foot. What did he do?



- ◯ He laughed.
- ◯ He apologised.
- ◯ He ran away.
- ◯ He did not say anything.
- ◯ Don't know
- ◯ Not attempted

11. This is Peter. He gave Agnes a sweet. What did she do?



- ◯ She apologised.
- ◯ She refused it.
- ◯ She took it.
- ◯ She was angry.
- ◯ Don't know
- ◯ Not attempted

12. This is Mr. John. He is leaving his friend's house. What time of day is it?

- ◯ Morning
- ◯ Afternoon
- ◯ Night time
- ◯ Monday

	◯ One hundred naira ◯ He did not pay anything ◯ Don't know ◯ Not attempted
13.c. What was the size of the loaf that Abu took home?	◯ Small ◯ Big ◯ Medium ◯ Abu did not take a loaf with him ◯ Don't know ◯ Not attempted
13.d. Which of the following is the best word to describe the trader?	◯ Honest ◯ Dishonest ◯ Clever ◯ Stupid ◯ Don't know ◯ Not attempted
13.e. How did Abu feel when the trader tried to trick him?	◯ Happy ◯ Sad ◯ Tired ◯ Angry ◯ Don't know ◯ Not attempted
13.f. Which of the following is the best way to describe what Abu did to the trader?	◯ Clever ◯ Stupid ◯ Happy ◯ Dishonest ◯ Don't know ◯ Not attempted
Silent Reading Passage 1 Enumerator, let the learner read the passage silently.	
Mr Rabbit bought four plantains and took them home. He cooked them himself. Then he called his wife and three children. When they were all sitting at the table, he brought in the four plantains.	

'I have roasted four plantains for you,' he said. 'I will share them out now. One for my wife, two, three, four for my children, and none for me.'

Mr Rabbit looked very unhappy.

He put one plantain on a plate and gave it to his wife. He put another on a plate and gave it to his first child. He put another on a plate and gave it to his second child. He put the last one on a plate and gave it to his third child. 'Eat up your plantains, my dears,' Mr Rabbit said. 'I am very hungry, but there is none for me. I will go without food.' 'Oh no, father, you must have half of mine,' the first child said. He cut his plantain in two and gave half to his father. 'Yes, you must have half of mine too,' the second child said, and gave her father half of hers. 'And you must have half of mine too,' the third child said, and gave Mr Rabbit half of hers. 'And you must have half of mine too,' his wife said. She put half of her plantain onto her husband's plate. Mr Rabbit quickly ate everything on his plate.

14.a. How many children did Mr. Rabbit have?	◯ One ◯ Two ◯ Three ◯ Four ◯ Don't know ◯ Not attempted
14.b. How many people were there in his family?	◯ Two ◯ Three ◯ Five ◯ Six ◯ Four ◯ Don't know ◯ Not attempted
14.c. How many plantains did Mr. Rabbit cook?	◯ Two ◯ Three ◯ Five ◯ Six ◯ Four

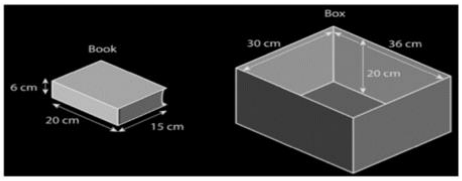
	◯ Don't know ◯ Not attempted
14.d. How many plantains did he eat?	◯ One ◯ Two ◯ Three ◯ Four ◯ Don't know ◯ Not attempted
14.e. Who ate the most plantains?	◯ Mr. Rabbit ◯ His wife ◯ The first child ◯ The third child ◯ Don't know ◯ Not attempted
14.f. Which sentence would you choose from the following to describe Mr. Rabbit?	◯ He was lazy. ◯ He was angry. ◯ He was clever. ◯ He was stupid. ◯ Don't know ◯ Not attempted
14.g. Which sentence would you choose from the following to describe how the children felt about their father not having any plantains?	◯ They were angry. ◯ They were concerned. ◯ They were pleased. ◯ They were angry with their mother. ◯ Don't know ◯ Not attempted
Listening Passage 2 *Enumerator : Read out the passages to the learner and the ask questions below each passage.*	
One day our teacher brought a caterpillar to school in a jar. It was eating green leaves. This tiny caterpillar was going to change into a butterfly. Caterpillars usually turn into butterflies outdoors. But we watched our caterpillar change into a butterfly right in our classroom.	

15.a. What was the caterpillar eating when the teacher first showed it to the children? *(Correct answer : Green leaves **OR** leaves **OR** plants)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
15.b. Where did the children watch the caterpillar turn into a butterfly? *(Correct answer : In their classroom **OR** in a jar **OR** at school)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Our teacher told us our caterpillar started out as a tiny egg. The mother butterfly laid the egg on a leaf. The mother butterfly chose the leaf of a plant that the caterpillar would eat. When the caterpillar hatched out of the egg, it was hungry. It ate its own eggshell! Then it started to eat green plants right away. The caterpillar's job was to eat and eat, so it would grow. Each day it got bigger. The caterpillar ate and grew for 12 days.	
15.c. Why did the egg need to be on a leaf? *(Correct answer : So it would have food **OR** the caterpillar could eat it **OR** the caterpillar would be hungry **OR** to eat it)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
15.d. What was the first thing that the caterpillar did when it hatched?	◯ chose a leaf ◯ ate its eggshell ◯ ate green plants ◯ laid an egg ◯ Don't know ◯ Not attempted
Primary Math	
1. Enumerator : Ask the learner to identify three numbers correctly. For example, ask "Can you show me the number '23' please?" Repeat this 3 times, picking 1 one-digit number and 2 two-digit numbers from the square. Do not point out the number to the learner. *Enumerator note : Record the number of digits correctly identified in the box.*	**55** | **3** | **23** **4** | **12** | **6** **34** | **21** | **9**

How many numbers was the child able to identify correctly?	
2. Enumerator : Please ask the learner to write down the numbers in this list in order from smallest to largest. *(Correct answer : 2, 17, 55, 117, 123, 987)*	**117 987 17 2 55 123**
Was the child able to write down the numbers in the right order?	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3. Enumerator : Please read each question and its associated answers to the learner and say "Work out the answer for each of these questions and write it on paper. You are free to skip any question you are unable to answer."	
3.a. 8 + 7 *(= 15)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.b. 28 + 27 *(= 55)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.c. 335 + 145 *(= 480)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.d. 8 - 5 *(= 3)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.e. 57 - 49 *(= 8)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted

3.f. 7 x 8 *(= 56)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.g. 37 x 40 *(= 1480)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.h. 214 x 104 *(= 22256)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.i. 6/3 *(= 2)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.j. 75/5 *(= 15)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
4. 4 × ..... = 28 *(= 7)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
5. 3 + 8 = ... + 6 *(= 5)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
6. In which of the numbers below does the 8 have the place value of 800?	◯ 1,468 ◯ 2,587 ◯ 3,809 ◯ 8,634 ◯ Don't know ◯ Not attempted

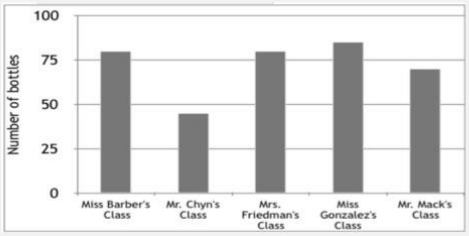
7. Which number is 152 more than 5,432?	◯ 6,432 ◯ 5,584 ◯ 5,444 ◯ 5,433 ◯ Don't know ◯ Not attempted
8. Which of these below gives a value close to 9 × 22?	◯ 5 x 20 ◯ 5 x 25 ◯ 10 x 20 ◯ 10 x 25 ◯ Don't know ◯ Not attempted
9. Pick a number that is larger than 5 and is smaller than 6.	◯ 6.1 ◯ 5.5 ◯ 4.5 ◯ 4.7 ◯ Don't know ◯ Not attempted
10. Enumerator : Please read the following sentence : "Write down the following sums and work them out on your worksheet." a. 80 ÷ 5 *(= 16)* b. 80 ÷ 4 *(= 20)* c. 90 ÷ 3 *(= 30)* *Enumerator : Mark correct only if all three answers are correct.*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
11. Enumerator: Please read the following sentence: "A box holds 26 oranges. How many oranges are there in 10 boxes? Please write down the number." *(Correct answer: 260)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
12. Enumerator: Please read the following sentence: "What is the next number in the sequence? Please write down the number." 48, 24, 12, 6, ___ *(Correct answer: 3)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
13. Amina writes letters to 12 of her friends. One third of the letters are 1 page each and a further third of the letters are 2 pages each. The	◯ 12 ◯ 14 ◯ 24

remaining letters are half a page each. How many pages has she written altogether?	◯ 36 ◯ Don't know ◯ Not attempted
14. A train left Abuja at 8:45 a.m. It arrived in Kaduna 2 hours and 18 minutes later. What time did it arrive in Kaduna?	◯ 11:15 a.m. ◯ 11:13 a.m. ◯ 11:03 a.m. ◯ 10:53 a.m. ◯ Don't know ◯ Not attempted
15. Yusuf is packing books into this rectangular box. All books are the same size. What is the largest number of books that will fit inside the box? *(Correct answer: 12)*  	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
16. Paint comes in 5 liter cans. Segun needs 37 liters to paint his house. How many cans must he buy?	◯ 5 ◯ 6 ◯ 7 ◯ 8 ◯ Don't know ◯ Not attempted
17. Hadija ate ½ of a cake, and Jane ate ¼ of the cake. How much of the cake did they eat altogether?	◯ 1/2 ◯ 3/4 ◯ 3/5 ◯ 3/8 ◯ Don't know ◯ Not attempted
18. The distance from Lagos to Kano is 100 km and a bus travels at 50 km an hour all the way. How long will it take the bus to get from Lagos to Kano?	◯ 1 1/2 hours ◯ 2 hours ◯ 3 hours ◯ 3 1/2 hours ◯ Don't know ◯ Not attempted

19. The school playground is a square. The playground is 100 meters long. Ruth walks all the way around the edge of the playground. How far does she walk?

- ◯ 100 meters
- ◯ 200 meters
- ◯ 400 meters
- ◯ 10,000 meters
- ◯ Don't know
- ◯ Not attempted

20. Central School had a bottle collection. Children in each class brought empty bottles to school. The principal made a bar graph of the number of bottles from five classes. Which class collected 45 bottles?



- ◯ Miss Barber's class
- ◯ Mr. Chyn's class
- ◯ Mrs. Friedman's class
- ◯ Mr. Mack's class
- ◯ Don't know
- ◯ Not attempted

21. Here is a calendar for December. Mary's birthday is on December 2. She is going on a trip exactly 3 weeks later. On what date will she go on the trip?

December						
S	M	T	W	T	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

- ◯ December 16th
- ◯ December 21st
- ◯ December 23rd
- ◯ December 30th
- ◯ Don't know
- ◯ Not attempted

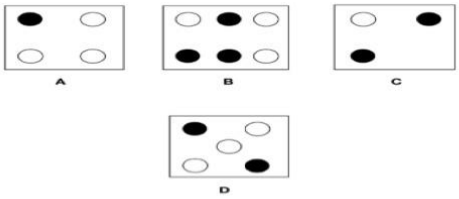
22. This chart shows temperature readings made at different times on four days. When was the highest temperature recorded?

TEMPERATURES					
	6 a.m.	9 a.m.	Noon	3 p.m.	6 p.m.
Monday	15°	17°	20°	21°	19°
Tuesday	15°	15°	15°	10°	9°
Wednesday	8°	10°	14°	13°	15°
Thursday	8°	11°	14°	17°	20°

- ◯ Noon on Monday
- ◯ 3 pm on Monday
- ◯ Noon on Tuesday
- ◯ 3 pm on Wednesday
- ◯ Don't know
- ◯ Not attempted

23. In which figure are one-half of the dots black?

- ◯ Figure A
- ◯ Figure B
- ◯ Figure C
- ◯ Figure D



- ◯ Don't know
- ◯ Not attempted

24. On a map, 1 centimeter represents 4 kilometers on the land. The distance between two towns Kaduna and Abuja on the map is 8 centimeters. How many kilometers apart are the two towns?

- ◯ 2
- ◯ 8
- ◯ 16
- ◯ 32
- ◯ Don't know
- ◯ Not attempted

25. E stands for the number of pencils Olu had. Obi gave Olu 3 more pencils. How many pencils does Olu have now?

- ◯ 3 ÷ E
- ◯ E + 3
- ◯ E – 3
- ◯ 3 × E
- ◯ Don't know
- ◯ Not attempted

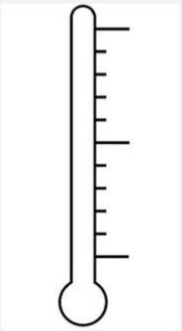
Primary Science

Enumerator: Please read each question and its associated answers to the learner and say "Work out the answer for each of these questions. You are free to skip any question you are unable to answer."

Here are some pictures showing things that affect the weather.

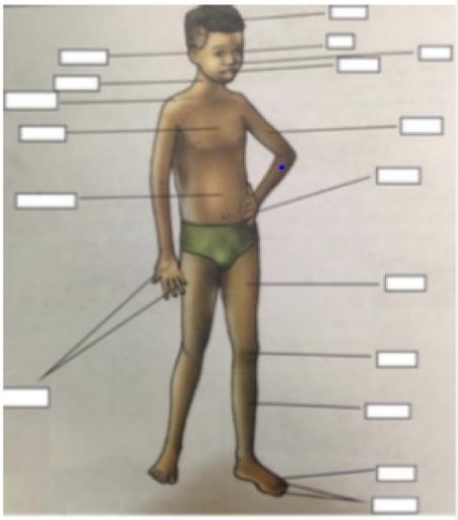
1.a. What is the name of this instrument?

Enumerator note: Record the learner's answer.



- ◯ Watch
- ◯ Ruler
- ◯ Wind vane
- ◯ Thermometer
- ◯ Don't know
- ◯ Not attempted

1.b. What is this instrument used for?	◯ To tell the time of the day ◯ To tell the direction of the wind ◯ To tell how much it has rained ◯ To tell how hot it is ◯ Don't know ◯ Not attempted
1.c. What is the name of this instrument? *Enumerator note: Record the learner's answer.* 	◯ Ruler ◯ Watch ◯ Wind vane ◯ Rain gauge ◯ Don't know ◯ Not attempted
1.d. What is this instrument used for?	◯ To tell the time of the day ◯ To tell the direction of the wind ◯ To tell how much it has rained ◯ To tell how hot it is ◯ Don't know ◯ Not attempted
2. The boy in this picture is called Abdul. We are going to see if we can name the parts of Abdul's body. I am going to read the names of different body parts to you.	

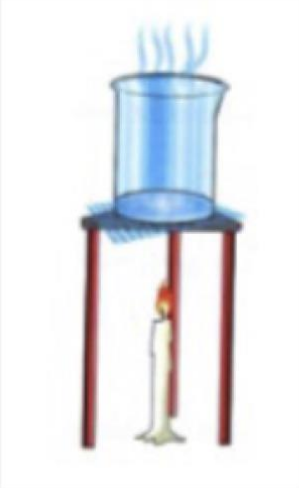


Let us do an example: Ear. Show me Abdul's ear. *Enumerator note: Point out the correct answer if the learner is wrong here.* Leg. Show me Abdul's leg. *Enumerator note: Point out the correct answer if the learner is wrong here.*	
Thank you. Let's start. *Enumerator note: Do not point out the correct answer to the learner now.*	
Eye	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Mouth	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Hand	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted

Fingers	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Stomach	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Shoulder	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Chest	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Shin	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Head	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Toe	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Neck	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Foot	◯ Correct ◯ Incorrect

	◯ Don't know ◯ Not attempted
Arm	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Nose	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Knee	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.a. If we put an empty plastic bottle in a bowl of water, will it float or sink? *(Correct answer: Float)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
3.b. If we put a key into a bowl of water, will it float or sink? *(Correct answer: Sink)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
4.a. Look at this picture. A farmer left a bag on the grass for several days.  He returned and picked up his bag. What do you think has happened to the grass that was under the bag?	◯ It grew taller ◯ It died ◯ It got wet ◯ It grew wider ◯ Don't know ◯ Not attempted
4.b. Why do you think this has happened?	◯ The grass was not cut ◯ The grass was contaminated

	◯ The grass did not get any sunlight ◯ The grass was burned ◯ Don't know ◯ Not attempted
5. Children in a Primary 4 class in Nigeria have measured the rainfall for a week. They made this table. Look at the table and answer the questions that follow	Daily Rainfall Table Days of the Week / Amount of Rain (in millimetres, mm) Monday — 1 mm Tuesday — 5 mm Wednesday — 2 mm Thursday — 4 mm Friday — 3 mm Saturday — 0 mm Sunday — 1 mm
5.a. Based on the information in this table, does it rain everyday?	◯ Yes ◯ No ◯ Not enough information in the table to decide ◯ None of the above ◯ Don't know ◯ Not attempted
5.b. What was the highest amount of rainfall in one day?	◯ 5 mm ◯ 1 mm ◯ 3 mm ◯ 2 mm ◯ 0 mm ◯ Don't know ◯ Not attempted
5.c. What was the lowest amount of rainfall in one day?	◯ 5 mm ◯ 1 mm ◯ 3 mm ◯ 2 mm ◯ 0 mm ◯ Don't know ◯ Not attempted
6. Look at this picture. What word in science can we use to describe what is happening to the water?	◯ Condensation ◯ Filtration ◯ Evaporation ◯ Excavation ◯ Don't know ◯ Not attempted



7.a. Abu is putting his finger into a pot of water. Study the picture. Do you think the water is –



- ◯ Cold
- ◯ Sweet
- ◯ Bitter
- ◯ Warm
- ◯ Don't know
- ◯ Not attempted

7.b. Anam is putting her finger into a bowl of water. Study the picture. Do you think the water is –

- ◯ Warm
- ◯ Sweet
- ◯ Cold
- ◯ Bitter
- ◯ Don't know
- ◯ Not attempted



8.The picture shows living things and non-living things. Point out all the living things.

Read out the options to the learner.



Is the sun a living thing? *(Correct answer: No)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Are rocks living things? *(Correct answer: No)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Are frogs living things? *(Correct answer: Yes)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Are trees living things? *(Correct answer: Yes)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Are clouds living things? *(Correct answer: No)*	◯ Correct ◯ Incorrect

	◯ Don't know ◯ Not attempted
Is water a living thing? *(Correct answer: No)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Are turtles living things? *(Correct answer: Yes)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Are flies living things? *(Correct answer: Yes)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
9. Some of the materials I am about to say will burn and some will not. Tell me which materials will burn. *Read out the options to the learner.*	
Will water burn? *(Correct answer: No)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Will wood burn? *(Correct answer: Yes)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Will sand burn? *(Correct answer: No)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
Will petrol burn? *(Correct answer: Yes)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted

Will air burn?

(Correct answer: No)

- ◯ Correct
- ◯ Incorrect
- ◯ Don't know
- ◯ Not attempted

10. Some animals are very rare. For example, there are very few tigers. If the only tigers left are female, what will most likely happen?

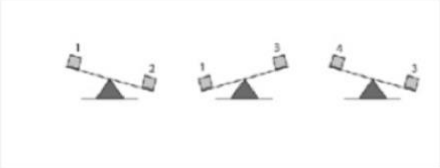
- ◯ The females will find another type of male animal to mate with and produce more tigers.
- ◯ The females will mate with each other and produce more tigers.
- ◯ The females will only be able to produce female tigers.
- ◯ The females will not be able to produce more tigers, and they will die out.
- ◯ Don't know
- ◯ Not attempted

11. Water that has its salt removed before it can be used as drinking water is most likely to come from .....

- ◯ underground
- ◯ a river
- ◯ a lake
- ◯ a sea
- ◯ Don't know
- ◯ Not attempted

12. Stephanie has a balance and four cubes (1, 2, 3, 4). The cubes are made of different materials. She put two cubes at a time on the balance and observes these results.

What can she conclude about the weight of cube 2?



- ◯ It is heavier than cubes 1, 3, and 4.
- ◯ It is heavier than cube 1 but lighter than cubes 3 and 4.
- ◯ It is heavier than cube 3 but lighter than cubes 1 and 4.
- ◯ It is heavier than cube 4 but lighter than cubes 1 and 3.
- ◯ Don't know
- ◯ Not attempted

13. *Read the following passage out to the learner.*

Weather affects people in many ways. People wear clothes according to the weather. Warm clothes such as jackets are worn when it is cold. During dry seasons the weather is hot. People

Weather Elements	Good Effects	Bad Effects
Rainfall	It makes people's crops grow	
Wind		
Sunshine		

wear light clothes. When it rains, people use umbrellas to stop them from getting wet.

Rainfall provides water. This helps people to carry out agriculture. They grow different types of crops. They also keep animals such a cattle and goats. Sometimes rainfall is heavy. It destroys people's houses, plants and animals.

Wind helps to bring rain. As the wind blows, it makes clouds move from one place to another. Strong winds turn windmills. They are used to draw water from a well. But sometimes wind also destroys people's houses and crops.

Sunshine helps people to dry their crops and clothes. Crops and clothes are put under the sun to dry. Sometimes there is a lot of sunshine. It may lead to drought. This destroys people's crops. But it can help plants to make their own food and to grow strong.

Look at this table. I am going to read some sentences to you. You must show me where in the table I should write (place) the sentence. I have done the first one for you. Look. Rainfall is good because it makes people's crops grow well. This sentence goes here. Now, let's do the rest.

13.a. If there is too much of this, it makes plants dry up. *(Correct answer : Sunshine - Bad effects)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
13.b. If it is too strong, it can blow your house down. *(Correct answer : Wind - Bad effects)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
13.c. It helps plants to make their own food. *(Correct answer : Sunshine - Good effects)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
13.d. It turns windmills that help people to draw water. *(Correct answer : Wind - Good effects)*	◯ Correct ◯ Incorrect

Nigeria Student Learning Assessment

	◯ Don't know ◯ Not attempted
13.e. It causes flooding. *(Correct answer : Rain - Bad effects)*	◯ Correct ◯ Incorrect ◯ Don't know ◯ Not attempted
JSS	
Hello! Welcome. We are here to find out about the things you are learning at school, so thank you very much for helping me with that. Remember that this is not an exam or a test. You don't need to worry if you don't know some things. If you don't want to do this, you do not have to, and you can go back to your classroom at any time. There is nothing bad that can happen to you because of this exercise. Your answers will not be shared with your teacher, parents or anyone else.	
Do you understand what has been explained to you and that no one in this school, your parents or your friends will see the answers to this exercise?	◯ Yes ◯ No
Please call the administrator to help you!	
Do you agree to answer a few questions about yourself and your family?	◯ Yes ◯ No
Do you agree to participate in this exercise?	◯ Yes ◯ No
Please write your name	
Please select your gender	◯ Male ◯ Female ◯ No Response
Please write your age	
Please write the name of your school	
Did you have breakfast this morning?	◯ Yes ◯ No
The questions are mostly multiple choice questions. Please select the right answer from the	

Appendix M: Teacher Interview & School Survey (Extracts)

Teacher Interview & School Survey

Identification

Version: October 19th, 2019 (2 PM)

Enumerator, please fill in this page before handing the tablet to the teacher.

What kind of interview are you conducting? *

- ◯ Teacher Interview
- ◯ School Survey/ School Principal Interview

School Level *

- ◯ Primary
- ◯ JSS

Every school that is visited for this exercise must be from the list of randomly selected schools. Permission to visit any replacement school must come directly from FME. Once FME receives the request, you will receive the name of the replacement school. If a replacement school is chosen, please choose 'Other' for the State and then 'Other' again for LGA and then write down the details.

State *

- Abia
- Adamawa
- Akwa-Ibom
- Anambra
- Bauchi
- Bayelsa
- Benue
- Borno
- Cross-River
- Delta
- Ebonyi
- Edo
- Ekiti
- Enugu
- Fct-Abuja
- Gombe
- Imo
- Jigawa
- Kaduna
- Kano
- Katsina
- Kebbi
- Kogi
- Kwara
- Lagos
- Nasarawa
- Niger
- Ogun
- Ondo
- Osun
- Oyo
- Plateau
- Rivers
- Sokoto
- Taraba
- Yobe
- Zamfara
- Other

LGA *

School Name *

Please write the name of the school, province and local government authority (LGA) *

Please select your name *

Please enter your name here *

GPS co-ordinate - please press the button below while you are outside the school.

latitude (x.y °)

longitude (x.y °)

altitude (m)

accuracy (m)



Is this school from the original list of schools selected for the assessment? *

- Yes
- No
- No response/ Don't know

Why was a replacement school chosen? *

At what time did you arrive at the school? *

hh:mm

Teacher's Information

Good morning! My name is __. We are working with authorization from the Federal Ministry of Education. We are conducting a survey about education in various states of Nigeria. The information we collect will help the government to plan for better education services. This school was randomly selected for the survey. The data is not representative of any teacher, any school, any LGA or even any state. All the data will be anonymised. No information about how you respond to this interview will be shared with your school principal or anyone else. *

Do you have any questions? *(If yes, answer the questions)*

May we begin?

- ◯ Yes
- ◯ No
- ◯ No response/ Don't know

Your age *

Your gender *

- ◯ Male
- ◯ Female
- ◯ Unknown

What grades do you normally teach this academic year (2019/2020)? *Please select all that apply* *

- ☐ P1
- ☐ P2
- ☐ P3
- ☐ P4
- ☐ P5
- ☐ P6

What grades do you normally teach this academic year (2019/2020)? *Please select all that apply* *

- ☐ JSS 1
- ☐ JSS 2
- ☐ JSS 3

What grades did you normally teach last academic year (2018/2019)? *Please select all that apply* *

- [ ] P1
- [ ] P2
- [ ] P3
- [ ] P4
- [ ] P5
- [ ] P6

What grades did you normally teach last academic year (2018/2019)? *Please select all that apply* *

- [ ] JSS 1
- [ ] JSS 2
- [ ] JSS 3

What is the subject you specialize in? *

- ◯ English
- ◯ Maths
- ◯ Science
- ◯ Other

If other, please specify. *

Since which year have you been teaching? (Example. 2013) *

What is your position at school? *

- ◯ Head teacher/Principal
- ◯ Deputy head teacher
- ◯ Senior teacher
- ◯ Teacher (government)
- ◯ Teacher (paid contract)
- ◯ Teacher (volunteer)
- ◯ Other

If other, please specify. *

What is the highest level of education you have completed? *

- ◯ Primary
- ◯ Secondary
- ◯ Diploma/Certificate
- ◯ University degree (bachelor)
- ◯ University degree (master)
- ◯ None

Which year did you complete your highest level of training (year)? *

Did you teach in this school last year? *

- ◯ Yes
- ◯ No
- ◯ No response/ Don't know

How well can you *read* English? *

----- ◯ ◯ ◯

Not at all well Somewhat well Very well

How well can you *understand* English? *

----- ◯ ◯ ◯

Not at all well Somewhat well Very well

How well can you *speak* English? *

----- ◯ ◯ ◯

Not at all well Somewhat well Very well

Instructions

Please DO NOT move to the next page before you are told. The enumerator will tell you when to move to the next page to begin the questions.

You have 1 hour and 20 minutes to complete all questions.

Answer all questions. For each question, put your answer in the space provided below/next to each question.

Please ask the enumerator any questions you might have.

Water *

Construction and maintenance *

Teachers' meals *

The sum of all individual category budget expenditure is greater than the total budget you mentioned at the beginning. Please go back and correct the numbers.

What is the local contribution (PTA) on to the school budget last year (in Naira)? *

Please type '0' if there was no contribution made.

Which of the following items did you receive from any government or non-government source during the academic year 2018/19? *

- [ ] Desks
- [ ] Books
- [ ] Paper/notebooks
- [ ] Pens/pencils/writing implements
- [ ] Chalk
- [ ] Sponges
- [ ] Blackboards or chalkboards
- [ ] Electronic equipment of any kind
- [ ] Sports equipment
- [ ] Instructional kits or materials (e.g., science kit)
- [ ] Maps/posters/instructional materials for walls
- [ ] Uniforms
- [ ] Food (for student meals)
- [ ] Food (for students to take home)
- [ ] Sanitary towels
- [ ] Other
- [ ] None

If other, please specify. List one item only. *

Please list the quantities supplied below. Please write '999' if the respondent doesn't know the answer.

Desks *

Books *

Paper/notebooks *

Pens/pencils/writing implements *

Chalk *

Sponges *

Blackboards or chalkboards *

Electronic equipment of any kind *

Sports equipment *

Instructional kits or materials (e.g., science kit) *

Maps/posters/instructional materials for walls *

Uniforms *

Food (for student meals) *

Food (for students to take home) *

Sanitary towels *

Classroom Observation

Enumerator, now you may accompy the principal to a classroom visit. Answer these questions directly after visiting a classroom - preferably a P4 or a JSS2 classroom.

How many boys are in class? *

How many girls in class? *

Ask the pupils to hold up their textbooks for the lesson you just observed.

How many pupils in the class have a textbook for the lesson you observed? *

Ask the pupils to hold up their exercise books.

How many pupils in the class have an exercise book? *

Is there a "corner library" in the class or additional available books for pupils? *

◯ Yes ◯ No ◯ No response/ Don't know

Is there a blackboard or whiteboard in the class? *

◯ Yes ◯ No ◯ No response/ Don't know

Is there chalk/marker to write on the blackboard/whiteboard available during the lesson? *

◯ Yes ◯ No ◯ No response/ Don't know

Does the classroom have a working electricity connection (e.g. electric light)? *

◯ Yes ◯ No ◯ No response/ Don't know

Was children's work displayed on the walls? *

◯ Yes ◯ No ◯ No response/ Don't know

Other than children's work, were there other materials like charts displayed on the wall? *

◯ Yes ◯ No ◯ No response/ Don't know

How would you classify the hygiene in the classroom? *

Extremely clean and well maintained — Reasonably clean and maintained — Not very clean or maintained

How many desks are in the classroom? *

How many desks in the classroom are *broken*? *

Does the blackboard have sufficient contrast for reading what is written on the board? *

Enumerator, use stencil and write letters on the board with chalk/marker. Then observe from the back of the classroom.

◯ Yes ◯ No ◯ No response/ Don't know

Is there sufficient light for reading text from the *front* of the classroom? *

◯ Yes ◯ No ◯ No response/ Don't know

Is there sufficient light for reading text from the *back* of the classroom? *

◯ Yes ◯ No ◯ No response/ Don't know

Is the class held outside? *

◯ Yes ◯ No ◯ No response/ Don't know

Please take a photo of the classroom

Click here to upload file. (< 10MB)

Infrastructure

Enumerator, physically inspect toilets and classrooms personally and then record the answers below based on your observation.

Does the school have toilets for pupils? *

◯ Yes ◯ No ◯ No response/ Don't know

Are the toilets designated for boys and girls? *

◯ Yes ◯ No ◯ No response/ Don't know

How many boys' toilets does the school have? *

How many girls' toilets does the school have? *

Are the pupil toilets clean? *

- Yes
- No
- No response/ Don't know

Are the pupil toilets private (have doors or separating entry way wall)? *

- Yes
- No
- No response/ Don't know

Are the pupil toilets accessible (i.e., unlocked, not overflowing, in working condition)? *

- Yes
- No
- No response/ Don't know

What type of pupil toilet facility is used at the school? *

Select one option only. If there is more than one option, select the most common one.

- Flush to piped sewer system
- Flush to septic tank
- Flush to pit latrine
- Flush to somewhere else
- Flush to unknown place
- Ventilated Improved Pit latrine (VIP)
- Pit latrine with slab
- Pit latrine without slab/open pit
- Composting toilet
- Bucket
- Hanging toilet/hanging latrine
- No facilities or bush or field
- Other

If other, please specify. *

Please take a photograph of a toilet

Click here to upload file. (< 10MB)

What is the source of drinking water at the school? *

Select one option only. If there is more than one option, select the most common one.

- No water source
- Piped onto facility grounds
- Public taps/standpipe
- Tube well/borehole
- Protected dug well
- Protected spring
- Cart with small tank/drum
- Tanker truck
- Buy from vendors
- Bottled water
- Unprotected dug well
- Unprotected spring
- Rainwater
- Surface water
- Other

If other, please specify. *

How many classrooms does the school have for pre-primary classes? *

How many pre-primary classrooms are in good condition? *

How many classrooms does the school have for primary classes? *

How many primary classrooms are in good condition? *

How many classrooms does the school have for JSS classes? *

How many JSS classrooms are in good condition? *

How many classrooms does the school have for Special Needs classes? *

How many Special Needs classrooms are in good condition? *

Any comments?

You don't have to leave any comments, if you don't have anything you must say!

Please submit, do not select 'Save as Draft' as the form will never reach us!