Global Journal of HUMAN-SOCIAL SCIENCE: E
Economics
Volume 24 Issue 2 Version 1.0 Year 2024
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-460x & Print ISSN: 0975-587X

Exploring the Relationship between Language of Instruction and Academic Achievement among Primary School Students: Evidence from UAE

By Imen Hentati

Tunis-University of Tunis El Manar

Abstract- We explore the impact of the language of instruction on academic performance of Emirati students who are expected to learn mathematics and science in and through a second language. Using TIMSS 2019 standardized tests and employing propensity score technique we explain the differences in achievement in mathematics and science between two groups of young Emirati children living in an Arabic dialect–dominated environment who get instruction in English and who receive instruction in classical Arabic. Our findings indicate that students who receive instruction in English have lower odds of attaining higher achievement scores.

Keywords: TIMSS, native language, second language, students' performance, propensity score technique.

GJHSS-E Classification: FOR Code: 130204

EXPLORINGTHERELATIONSHIPBETWEENLANGUAGEOFINSTRUCTIONANDACADEMICACHIEVEMENTAMONGPRIMARYSCHOOLSTUDENTSEVIDENCEFROMUAE

Exploring the Relationship between Language of Instruction and Academic Achievement among Primary School Students: Evidence from UAE

Imen Hentati

Abstract- **We explore the impact of the language of instruction on academic performance of Emirati students who are expected to learn mathematics and science in and through a second language. Using TIMSS 2019 standardized tests and employing propensity score technique we explain the differences in achievement in mathematics and science between two groups of young Emirati children living in an Arabic dialect–dominated environment who get instruction in English and who receive instruction in classical Arabic. Our findings indicate that students who receive instruction in English have lower odds of attaining higher achievement scores.**

Keywords: TIMSS, native language, second language, students' performance, propensity score technique.

I. Introduction

Since its founding in 1971, the United Arab Emirates (UAE) have transformed from a desolate desert into a bustling metropolis thanks to money earned from oil profits. The United Arab Emirates has attained international status in fields including business, tourism, aviation, and architecture by using international standards to gauge its objectives. Nonetheless, comparable educational advancements still lag behind the nation's social and economic growth (Harold, 2005). Despite strong government support for the education sector, the quality of education as measured by student performance in international evaluations such as Trends in International Mathematics and Science Study (TIMSS) remains below international standards.

The language capital is an aspect that may affect student performance. Language capital, or more simply "the mother tongue," is defined in the literature as the set of skills that are acquired during childhood with no particular effort and strengthened in school (Chiswick and Miller, 1995). Arab students in UAE speak two varieties of Arabic; the classical Arabic known as « fushaa » and the different dialects spoken at home (the mother tongue). In such "dialect-dominated" environment, students are unlikely to be fluent in Arabic. Thus, language acquisition efficiency will be significantly limited (Boutieri, 2012). It has been shown that learning a first language requires at least 12 years of experience (Collier, 1989). By the time they reach this age, arabic-speaking pupils have not finished this process, and it might take them much longer to learn the fundamentals of the language.

In addition, a "new school model" has been announced by the government of Abu Dhabi in the United Arab Emirates for state kindergarten and early primary grades with effect from the academic year 2010-2011, and extending annually thereafter to cover all grade levels. Amongst a range of ongoing pedagogic, curricular and school leadership reforms, a major new departure is the introduction of English as an additional medium of instruction alongside the existing medium of Arabic. The reason to introduce English as a medium of learning for all Emirati children at such a young age is the belief that learning a second language at a younger age is beneficial. Despite this widespread view, international research on the issue is "surprisingly ambiguous" (Saville-Troike, 2006), with some review studies showing that adolescents are better language learners than young children, and others showing the opposite (Marinova-Todd et al., 2000). Cummins' (1984) asserts that there is an interdependence between the first language (L1), or the mother tongue, and the second language (L2), which also influences academic achievement (Cummins, 1978). According to the language interdependence hypothesis, L2 development depends in part on the developmental level of the L1, and students with lower levels of L1 and L2 proficiency are more likely to experience academic difficulties in school (Cummins, 1984).

Bilingual education in the Arab world has received little attention (Al-Khatib, 2006), and there has consequently been little research on language-in-education issues in the region. This paper aims to bridge this gap. Using the UAE as a case study, we want to explain the disparity in academic achievements of two groups of young students living in a dialect-dominated environment. The first is taught in English, while the second is taught in classical Arabic. More specifically, building on Cummins' theoretical background, we aim to provide (a) evidence of language interdependence in the Emirati context, and (b) its association with academic performance as measured by the standardized TIMSS 2019 test.

Author: Ph.D in Family Economics. She is an assistant professor at the Faculty of Economics and Management of Tunis-University of Tunis El Manar. She is a researcher at the Laboratory for Research on Quantitative Development Economics (LAREQUAD)
e-mail: imen_hinteti@yahoo.fr

Our work adds to the empirical literature linking academic performance to school language. It was implemented with three objectives in mind. The first is to expand research on the variables that influence academic performance. The second is to test the language interdependence hypothesis in the Emirati context. The third is to consider future language policies in a multilingual context.

II. Context of the Research Study

Since the early 1990s, the UAE has witnessed remarkable transformation in all aspects of socio-economic and political life in a short period of time and this rapid development has been made possible by an influx of an expatriate population. Martin (2003) reports that in 1968, nationals of the United Arab Emirates (UAE) comprised 63% of the population, but by 1975, after the discovery of oil, nationals comprised only 36%. Along with the commercialization of oil and migration of expatriate laborers came the proliferation of the English language to the UAE. English has historically played an essential role for the modernized Arab world. However, English, and the values it represents, has become a controversial issue within a culture dominated by Islamic traditions (Al Mahrooqi & Denman, 2015).

English might act as a connecting bridge to the outside, international community, but as such it replaces Arabic as the primary means of communication among local people. In the UAE society today, English not only as a medium of instruction in education, but also in mass media and communication technology, such as satellite TV stations, computers, and the Internet. Knowledge of English has become the key to better jobs, especially in the private sector: "*A good command of English is regarded as one way of gaining access into the private sector workforce*"(Troudi & Al Hafidf, 2017). That is why everyone feels the need to learn English. In certain settings, such hospitals and private businesses, the usage of English is becoming so imperative that Arabic is now considered secondary. This important role of English explains the increasing focus on using English as a medium of instruction in UAE schools.

Most parents who can afford it are keen to send their children to the mushrooming private and international schools that start English instruction from kindergarten. People are more and more eager to have an English education which they believe is the prerequisite to professional and social success (Troudi and Jendli, 2011).

It is generally understood that individuals' attitudes towards foreign languages are influenced by social, economic, and political factors. In the UAE, the spread of English in education has been made more rapid as a consequence of decisions taken by the government, private institutions, and individuals. Because policy makers in the region tend to link English to modernization, it has been aggressively taught in schools at all levels (Hanani, 2009).

A major challenge facing the new school model is a linguistic one: Arabic diglossia which refers to two varieties of the same language being used alongside each other for different functions. Modern Standard Arabic (MSA) is used in reading, writing, and formal spoken situations, while colloquial Arabic is used for speaking. According to Badry (2004) "the linguistic situation in every Arabic speaking country is characterized by diglossia where literate speakers select from several varieties of Arabic depending on the setting, interlocutors, and occasion of the communication". Furthermore, Al Sharhan (2007), in his study of educational languages in the United Arab Emirates, showed that Emirati children need to master not only two but three Arabic languages (trilingual): Gulf Colloquial Arabic; Modern Standard Arabic which is the language of instruction in schools and the language of written media; and Classical Arabic, required for learning the Quran. Emirati schoolchildren learning their first language faced significant challenges in mastering all registers of the language (Abu-Libdeh, 1996).

This complex linguistic situation is likely to contribute to the low levels of first language literacy in UAE, and is thought to be an aggravating factor in poor general educational attainment levels (Maamouri, 1998).

III. Theoretical Background

Bilingual or multilingual people are those who use two or more languages in their everyday lives. In order for someone to be considered bilingual, he/she needs to have an appropriate functional ability in both languages. Such control of the two languages is referred to as balanced bilingualism. However, most bilinguals use each language for different purposes, in different circumstances, and with different people in their everyday lives (Fishman, 1972). Bilingualism comes in three flavors: sub-coordinate, coordinate, and compound. Compound bilinguals acquire two languages in the same setting, combining their two verbal expressions to form one thought. Coordinated bilinguals learn both languages in distinct settings, such as school and home, so their vocabulary is part of two distinct and autonomous systems. One language predominates in a sub-coordinate multilingual (D'acierno, 1990).

There is also a difference between subtractive and additive bilingualism (Cummins, 2000). When the first language continues to be developed and the first culture continues to be valued while the second language is added this is known as additive bilingualism. According to research cited by Cummins (1994), kids who study in additive bilingual environments tend to be more successful than those whose original language and culture are completely or partially devalued by their schools and society and are instead substituted by a powerful second language.

For these reasons scholars have called for the need to use the mother tongue in the primary stages of education in order to improve the quality of education as well as to preserve the language, and to provide a solid foundation in the students' native language. This approach prepares them better for learning a second language (Baker, 2001; Cummins, 1999, 2000). Cummins (1984) argues for the need to make a distinction between the academic (cognitive academic language proficiency; CALP) and the conversational (basic interpersonal communicative skills; BICS) dimensions of language proficiency. BICS refers to the development of conversational fluency in the second language, which is a language needed to interact socially with other people and is an earlier development. CALP describes the use of language in decontextualized academic situations which include listening, speaking, reading, and writing. The above hypothesis means that a student's ability to converse in a second language (BICS) is not an indicator of his/her ability to engage in academic skills (CALP) in that second language.

Cummins (2001) elaborates the relationship between L1 and L2 development in his Development Interdependence hypothesis. According to him, cross-lingual proficiency can aid in the growth of intellectual and cognitive abilities. Cummins claims that if a minimal level of L1 academic and cognitive development is met, then cognitive and literacy abilities developed in the mother tongue, or L1, will translate between languages. Therefore, reading in L1 can help with both the continuation of first language development and the growth of literacy and knowledge in L2. This implies that the learner may struggle to become bilingual if the threshold of cognitive proficiency in L1 is not met. Research has shown that it usually takes at least five years for second language learners to catch up academically to their native English-speaking peers, but conversational fluency in English is often attained within two years of intensive exposure to the language (Cummins, 1999). Cummins (2001) further argues in his Threshold hypothesis that there are threshold levels of linguistic competence that bilingual children must achieve in both of their languages to take advantage of the benefits of bilingualism.

The findings of several investigations support Cummins's hypothesis (1979). According to Papanastasiou (2000), students who took TIMSS evaluations in a second language experienced lower achievements. Similarly, Herbert et al. (2002) showed that students who received their instruction in English (second language) instead of Chinese are more disadvantaged than native students. In the same vein, Brock-Utne (2007) emphasized that using English as a medium of instruction in Tanzania hinders students' learning. Likely, Samuelson and Freedman (2010) showed that using English as the only language of instruction, would not always allow students to participate successfully in the global economy because many of them will not develop a solid command of academic literacy in their mother tongue.

The Cummins (1979) theoretical framework has been beneficial for second language education from a theoretical standpoint. His main argument is that bilingualism can be advantageous for academic and cognitive purposes when LI abilities are sufficiently developed. In this paper, we will examine the linguistic interdependence in the setting of the United Arab Emirates by comparing the academic achievement of two groups of students whose mother tongue is "Arabi" (Dialect). One group is receiving instruction in English, while the other is receiving instruction in Arabic.

IV. Timss Data and Methods

a) *TIMSS Data*

The primary aim of the TIMSS (Trends in International Mathematics and Science Study) assessments is to establish a reliable and valid measure of knowledge and skills in mathematics and science. These assessments hold significant value within the global education community and are designed to align with the curricula of participating countries. Introduced in 1994/1995, the initial implementation of TIMSS involved 45 countries, with Kuwait being the only Arab participant at that time. However, as time has progressed, the participation of Arab nations in TIMSS evaluations has increased, encompassing both fourth and eighth-grade levels.

In the latest TIMSS survey conducted in 2019, a total of 64 countries and 8 benchmarking participants took part. Notably, 9 Arab countries participated at the fourth-grade level, while 11 Arab countries participated at the eighth-grade level. The TIMSS 2019 assessment witnessed the involvement of 6 Gulf countries, namely Bahrain, Oman, Qatar, the United Arab Emirates, Saudi Arabia, and Kuwait. In addition, Dubai played a role as a benchmarking participant. To ensure meaningful comparisons across countries, the TIMSS assessments are strategically scheduled to take place at the end of the academic year. This timing allows for greater cross-country comparability by minimizing potential variations in the content and timing of instruction across different educational systems.

The choice to use the United Arab Emirates as a sample for our study investigating the impact of language of instruction on academic achievement for primary students is justified for multiple reasons. Firstly, UAE has demonstrated a strong commitment to education and have made significant investments in their educational systems. The UAE's Ministry of

Education (MOE) has started a number of educational reforms since the early 1990s, at a period of significant social and economic development in the country as well as growing influence from foreign cultures (UNESCO, 2016). In 2009, the Abu Dhabi Education Council (ADEC) which is the government organization in charge of managing public education in Abu Dhabi, developed the New School Model (NSM). Amongst a range of ongoing pedagogic, curricular and school leadership reforms, a major new departure is the introduction of English as an additional medium of instruction alongside the existing medium of Arabic. Furthermore, the UAE has a distinctive linguistic diversity due to the various ethnic groups that comprise its population (Syed, 2003). Because of this intricate language map, English has come to be used as the common language across the various ethnic groups. This English invasion has also been expedited by the United Arab Emirates' rapid advancements in business and communication technology.

b) Sample Design and Exclusion

TIMSS is a globally recognized assessment program that employs meticulous sampling procedures to ensure the accuracy and reliability of its estimates, striving to provide an accurate representation of the larger population. In order to maintain consistency and fairness across different cycles of TIMSS, stringent exclusion criteria have been established at both the school and student levels, as documented by Martin et al. (2020). At the school level, certain criteria are employed to exclude specific schools from the dataset. Schools located in remote areas or with an extremely small student population, such as four or fewer students in the target grade, are excluded. Additionally, schools that deviate significantly from the mainstream educational system or exclusively cater to children with special needs are also excluded. These exclusion criteria ensure that the dataset primarily consists of schools that align with the typical educational context. Moreover, international guidelines are implemented to determine within-school exclusions. These guidelines take into account students with functional disabilities, students with intellectual disabilities, and students who are non-native language speakers. By excluding these specific student groups, the aim is to maintain a standardized assessment environment and ensure that the results accurately reflect the abilities of the target population. It is important to note that these exclusion criteria are designed to enhance the validity and reliability of the assessment outcomes. In our specific sample, the overall exclusion rate is 5,6%, demonstrating the adherence to these exclusion guidelines while still maintaining a robust and representative dataset for analysis *(Table A-1).*

The reason for selecting fourth grade students is that at this grade most uses of language in school is cognitively demanding and context-reduced, which means that students have to rely primarily on linguistic cues to meaning and may in some cases require suspending knowledge of the "real world" in order to appropriately interpret the logic of the communication (Cummins & Swain, 1986). Since our aim is to test the linguistic interdependence hypothesis and its link to students' performance, we define two groups of students: those who took the tests in Arabic known as the control group and those who took the tests in English known as the treatment group. In order to get more accurate results and knowing that only native speakers are allowed to participate in TIMSS (according to the students' exclusion criteria discussed above), we restricted the participants who took the test in English to those whose both parents were born in the country. By doing so, we excluded students who were born from mixed marriages (having a native parent), and speak Arabic and English at home. Our sample is then composed of 3092 students wehere the proportion of students who took the tests in English amounts to 32,47%.

V. Methods

a) Empirical Model and Technique

OLS regression is first used to estimate the average treatment effect of the use of English as a medium of instruction. The variable treatment is introduced in *Equation (1)* as a dummy variable. Recall that the treated group consists of students receiving English instruction, while the control group consists of students receiving Arabic instruction.

$$P_{i,c,s} = \alpha_0 + \alpha_1 F_{i,c,s} + \alpha_2 T_{i,c,s} + \varepsilon_{i,c,s} \qquad (1)$$

Where P_{ics} is the score of student i in class c at school s. F_{ics} is a vector of individual and family background characteristics. T_{ics} is the treatment variable which is a binary variable that takes the value 1 for treated observations and 0 for control observations. To draw valid inferences, we use the students' sampling weights. The aim of OLS regression is to examine if there are any significant differences in performance between the treatment and control groups. To gain better understanding of performance differences, and because it is not straightforward to directly compare the outcomes for these two groups because those who choose English education may differ from those who choose Arabic, we use the propensity score matching technique. The propensity score matching (PSM) is a quasi-experimental method in which the researcher creates an artificial control group by matching each treated unit with a nontreated unit with similar characteristics. PSM, in particular, computes the probability of a unit enrolling in a program based on observed characteristics. This is the propensity score. Then, based on the propensity score, PSM assigns treated units to untreated units. PSM is based on the assumption that untreated units can be compared to treated units based on some observable characteristics, as if the treatment had been fully randomized (Rubin, 2001).

The use of the matching technique is to control for the potential confounding influence of pretreatment variables (individual and family variables). We utilize a logit model to predict children's propensity for the treatment group. Following that, we use the Nearest Neighbor Matching to pair cases in both groups based on their likelihood of experiencing a treatment. We use matching with replacement to identify neighbor cases (Frisco et al., 2007).

b) Variables

This section describes the outcome variables used for the purpose of this study as well as the covariates.

Outcome variables: The dependent variables are overall performance in mathematics and science. In TIMSS, student achievement is represented by sets of five plausible values: Math 1, Math 2, Math 3, Math 4 and Math 5 for mathematics achievement and Science 1, Science 2, Science 3, Science 4 and Science 5 for science achievement. Plausible values are imputed values drawn from the estimated ability distributions (Martin et al., 2016; Mislevy, 1991). Plausible values are generated by making use of all available background data of the students. Plausible values are not intended to be estimates of individual student scores, but rather are imputed scores for like students- students with similar response patterns and background characteristics in the sampled population- that may be used to estimate population characteristics correctly (Martin et al., 2020). A detailed review of the plausible values methodology is given in Mislevy (1991).

Covariates: The core explanatory variables are individual and family background characteristics. These variables and their coding are listed in *Table A2.* Controlling for socioeconomic status (SES) and pre-education is important as they may influence the rate at which a second language is learned. Individual factors include the students' age, gender, and the number of years in pre-primary education. Family factors include parents' education and educational resources at home. The latter was constructed based on the data reported by students and their parents regarding the number of books and other study materials in their homes, the parents' levels of education (National Academies of Sciences, Engineering, and Medicine, 1997) and the parents' employment. Cut scores were used to define students into three categories: students with Many Resources, students with Some Resources and students with Few Resources. All the nominal variables were introduced in the model as dummies. For Sex, it takes the value 1 for females and zero for males. Parents' education level takes four categories: some primary, lower secondary, upper secondary, post secondary and university or higher. The category some primary is considered as a reference category. For home resources, students are assigned to three categories which are the following: students with many resources, students with some resources and students with few resources. The category few resources is considered as a reference category.

c) Descriptive Statistics

In what follows, we provide a comprehensive overview of the descriptive statistics pertaining to the outcome variables *(Table 1 and 2)* and covariates employed *(Table 3).* For continuous variables, we compute the means to summarize the central tendency of the data. On the other hand, categorical variables, such as gender or socioeconomic status, are characterized by distinct categories or groups. In order to depict the distribution of these variables, we calculate proportions.

Table 1: Descriptive statistics for Mathematics Achievements

Variables	Mean of Total Sample	Mean of Control Group	Mean of Treatment Group	Effect Size (Cohen's d)
Math1	480.668	492.112	456.868	-0.396
Math 2	481.054	492.763	456.705	-0.409
Math 3	481.271	492.183	458.576	-0.386
Math 4	480.936	491.953	458.025	-0.386

Math 5	479.333	491.600	453.821	-0 .426
Number of Observations	3092	2088	1004	

Table 2: Descriptive statistics for Science Achievements

Variables	Mean of Total Sample	Mean of Control Group	Mean of Treatment Group	Effect Size (Cohen's d)
Science 1	480.616	503.133	433.787	-0.703
Science 2	479.929	503.619	430.66	-0.739
Science 3	478.889	500.441	434.0671	-0.675
Science 4	478.269	500.715	431.589	-0.690
Science 5	479.363	501.947	432.396	-0.691
Number of Observations	3.092	2088	1004	

Within both the treated and control groups, the mean values of the outcome variables hover around the global average of approximately 500 points. The negative sign associated with the effect size signifies that learning in English is less effective than learning in Arabic. The effect size in mathematics and science performance is relatively important (for further insights on the range of variation, refer to Borenstein, 2009, and Hattie, 2009).

Table 3: Descriptive statistics of covariates for all students

Variables	Mean/Proportion	Min	Max
Age	9.612	8.42	11
Gender	0 .553	0	1
Pre-Education	1.566	0	4
Parents' education			
Some primary	0.0268	0	1
Lower secondary	0.0252	0	1
Upper secondary	0 .213	0	1
Post-secondary	0 .177	0	1
University or higher	0 .558	0	1

Resources			
Few resources	0.0124	0	1
Some resources	0 .890	0	1
Many resources	0 .0976	0	1
Number of Observations	3092		

The students included in the sample exhibit an average age of 9.61 years. In addition, the sample is approximately evenly divided between boys and girls, ensuring a balanced representation of both sexes in the analysis. Students typically received two years of pre-primary education. Furthermore, 55 percent of parents in UAE have a university degree or higher and more than 98 percent of students have access to some home educational resources.

VI. Results and Discussion

The following section provides a comprehensive analysis of the results obtained from both Ordinary Least Squares (OLS) regression and propensity score matching techniques. OLS regression analysis was conducted independently five times for each evaluation. To derive the final parameter of interest, Rubin's rules (Rubin, 1987) were applied, which enable the combination of results from multiple regression analyses to generate a comprehensive and reliable estimate. The findings from the OLS regression analysis are presented in *Table 4,* while *Table 5* presents the results obtained from propensity score matching.

The OLS results reveal a significant negative association between taking the test in english and academic achievement. Students who took the tests in English have significantly lower achievement in mathematics and science than their peers who passed the tests in Arabic.

Consistent with previous research (Ammermüller et al., 2005; Chiu & Khoo, 2005), the analysis reveals that family variables have a significant positive impact on students' performance. These findings emphasize the influential role of family factors in shaping academic outcomes. Moreover, when examining individual student characteristics, the coefficient of gender demonstrates a significant influence in evaluations; Boys outperform girls in mathematics, however, girls outperform boys in science. Preschool education participation has a positive and significant influence on overall performance for both evaluations which is in line with previous studies (Holla et al., 2021).

Table 4: OLS Results

Variables	Mathematics Achievement	Science Achievement
Treat	-45,436*** (3,918)	-81,918*** (4,835)
Age	28,496*** (4,180)	37,453*** (4,7630)
Sex	-6,083* (3,280)	5,848* (3,609)
Pre Education	3,897*** (1,379)	2,466*** (1,539)
Parents' Education		
Lower secondary	3,130 (14,713)	5,695 (16,004)
Upper secondary	-6,639 (11,745)	6,601 (12,884)

Resources		
Some resources	27,306** (13,593)	15,991 (16,735)
Many resources	42,786*** (14,227)	35,430** (17,537)
Constant	177,0532*** (42,794)	90,805* (48,426)

Notes: Standard errors are in parentheses.
**** p<.01, ** p<.05, * p<.1*

To further substantiate the finding between language of instruction and academic achievement, we employ a propensity score model. The results of this analysis are presented in *Table 5.*

Table 5: ATET Nearest Neighbor Results

ATE TREAT (1 vs. 0)	Mathematics	Science
Overall Performance	-51,003*** (4,128)	-88,844*** (5,106)

Notes: Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

Students who received English-language mathematics instruction (Treatment group) had their overall scores reduced by 51 points, respectively, when compared to students who passed the Arabic-language Mathematics test (Control group). Similarly, students who received science instruction in English (Treatment group) had their overall scores fall by 88 points, respectively, as compared to students in the control group. This suggests that performance at this grade level is lower when education is offered in English. All the differences are statistically significant at the 1% level. Our findings reveal that students whose mother tongue is Arabic and with early English learning experience score significantly lower in TIMSS evaluations than their peers who take the tests in Arabic. The differences are more noticeable in science than in mathematics. The learning of mathematics and science requires a variety of linguistic skills that second language learners may not have mastered by the age of nine. Even though mathematics language seems to be abstract, the study of mathematics (especially at the primary level) begins with the study of real-world problems and requires the application of the language. Therefore, the language plays an important role in conveying mathematical knowledge to students and in knowing how abstraction is interpreted (Ferrari, 2003). Likely, the difference in science proficiency between the treatment and control group is manifest. This provides evidence that a first language enhances more science learning than a second language. At grade 4, science context domains are more linked to real-world situations and thus to "home language." Students need to develop a deep understanding of science concepts, make connections among concepts and apply concepts in explaining natural phenomena or real-world situations. Students in science classes must be engaged in science inquiry, have to negotiate ideas and justify claims based on evidence. It has been shown that sometimes, and for effective instruction in science, teachers focus on students' home language as an instructional support. They use students' home language to explain science terms (Goldenberg, 2013).

A key premise in the literature is that age is significant in child language acquisition, whether in L1 or L2 (Oliver and Azkarai, 2017). Given that the critical period of twelve years is required to be fully proficient in the first language (Collier, 1989), students in UAE at the age of nine (fourth grade) have not yet begun to complete full cognitive development in the first language and so do their peers in their native languages. Nonetheless, the difference is that English native speakers are approaching the completion of L1 acquisition, whereas students whose "mother tongue" is Arabic (Dialect) require additional years to improve their Arabic language skills. So they lag behind their peers in terms of years of L1 acquisition. The mismatch between the language of the home and the language of the school makes those students less proficient in "modern Arabic."

VII. Conclusion and Recommendations

In this study, we investigate the effect of the language of instruction (English in this case) on academic achievement of Emirati-speaking students who are expected to learn science and math in a second language (L2) before achieving a sufficient level of proficiency in it.

Based on Cummins's (1984) theoretical framework, which links language proficiency to academic achievement, we offer some insights regarding the connection between Emirati students' academic achievement in mathematics and science and their L1 and L2 proficiency. In order to explain the disparity in academic performance in science and mathematics between two groups of young children living in a dialect-dominated environment who got instruction in English and who received instruction in Arabic, we used the propensity score technique and the TIMSS 2019 standardized tests.

Our results demonstrate that the differences in performance between the two groups can be explained by the language used during instruction. Young children who started studying L2 before their L1 was fully matured showed overall performance losses as compared to children who started learning Arabic at the same age. The results of this study support the linguistic interdependence hypothesis developed by Cummins (1984). This study helps us to think about future language policy not only in the UAE, but also in Arab countries that are all multilingual.

Strengthening early exposure to the Arabic language (modern Arabic) will minimize the retardation produced by the mismatch between students' mother tongue and the school language. Arabic should be taught and learned in schools in a more practical manner. Additionally, Arabic learners should have access to audio-visual resources, which have the added benefit of advancing the language's cognitive and intellectual components in addition to its fluency. Finally, pupils who are fully proficient in both languages would benefit from the cognitive and academic advantages of bilingualism, while bilinguals who are only sufficiently proficient in one of their languages would not face any such academic drawbacks.

Statements and Declarations:

Funding: There has been no significant financial support for this work that could have influenced its outcome.

Ethical Approval: Not applicable.

Informed Consent: This article does not contain any studies with human participants performed by any of the authors.

Author's Contribution: There is only one author for this article

Conflict of Interest: We know of no conflicts of interest associated with this publication.

Data Availability Statement: https://timss2019.org/international-database/

References Références Referencias

1. Borenstein (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), The handbook of research synthesis and meta analysis (pp. 279–293). Russell Sage Foundation. Available at: https://www.daneshnamehicsa.ir/userfiles/files/1/9-%20The%20Handbook%20of%20Research%20Synthesis%20and%20Meta-Analysis.pdf.
2. Boutieri, C. (2012). In two speeds (A deux vitesses): Linguistic pluralism and educational anxiety in contemporary Morocco. International Journal for Middle East Studies, 44(3), 443–464. https://doi.org/10.1017/S0020743812000414.
3. Chiswick, B. R., and Miller, P. W. (1995). The endogeneity between language and earnings: International analysis. Journal of Labor Economics; 13, 246–288. http://dx.doi.org/10.1086/298374.
4. Chiu, M. M., & Khoo, L. (2005). Effects of resources, inequality, and privilege bias on achievement: Country, school and student level analyses. American Education. Research Journal, 42, 575603. https://doi.org/10.3102/00028312042004575.
5. Collier, V. P. (1989). How Long? A Synthesis of Research on Academic Achievement in a Second Language. TESOL Quarterly, 23, 509–531. https:// doi.org/10.2307/3586923.
6. Cummins, J. (1978). Educational implications of mother tongue maintenance in minority-language groups. The Canadian Modern Language Review, 34, 395-416. https://doi.org/10.3138/cmlr .34.3.395.
7. Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. Review of Educational Research, 49(2), 222-251. https://doi.org/10.3102/00346543049002222.
8. Cummins, J. (1984). Wanted: A theoretical framework for relating language proficiency to academic achievement among bilingual students. In C. Rivera (Ed.). Language proficiency and academic achievement (pp. 2–19). Multilingual Matters. http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED240882.
9. Cummins, J. (2008). "BICS and CALP: Empirical and theoretical status of the distinction." In B. Street & N. H. Hornberger (Eds.). Encyclopedia of language and education, 2nd Edition, Volume 2: Literacy (pp. 71–83). Springer. 10.1007/978-0-387-30424-3_36
10. Cummins, J., & Swain, M. (1986). Bilingualism in education. Longman. https://doi.org/10.4324 /9781315835877.
11. Ferrari, P. L. (2003). Abstraction in mathematics. Philosophical Transactions: Biological Sciences, 358(1435), 1225–1230. 0.1098/rstb.2003.1316.
12. Fishbein, B., Foy, P., & Yin, L. (2021). TIMSS 2019 user guide for the international database (2nd ed.). https://www.iea.nl/publications/user-guides/user-guide-international-database-timss-2019.
13. Frisco, M. L., Muller C, & Frank, K. (2007). Family structure change and adolescents' school

performance: A propensity score approach. Journal of Marriage and Family; 69, 721–741.
14. GCC Education. (2020). Sector report GCC education. https://gfh.com/wp-content/uploads/2020/05 /GFH-Education-Sector-Report-2020.pdf.
15. Goldenberg, C. (2013). Unlocking the research on English learners: What we know—and don't yet know—about effective instruction. The American Educator, 37(2), 4–11. https://eric.ed.gov/?id=EJ1014021.
16. Hattie, John (2009): Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge. https://doi.org/10.4324/9780203887332.
17. Herbert, W. M., Hua, T. K., & Kong, K. C. (2002). Multilevel causal ordering of academic self-concept and achievement: Influence of language of instruction (English compared with Chinese) for Hong Kong students. American Educational Research Journal, 39, 727–763. https://doi.org/10.3102/00028312039003727.
18. Holla, A., Bendini, M., Dinarte, L., & Trako, I. (2021). Is investment in pre-primary education too low? lessons from (quasi) experimental evidence across countries. Policy Research Working Paper, No. 9723. World Bank. http://hdl.handle.net/10986/35894.
19. Martin, M. O., Mullis, I. V. S., Foy, P. & Hooper, M. (2016). TIMSS achievement methodology. In Methods and procedures in TIMSS 2015 (pp. 12.1–12.9). https://timssandpirls.bc.edu/publications/timss/2015-methods.html.
20. Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). Methods and procedures: TIMSS 2019 Technical Report. https://timssandpirls.bc.edu/timss2019/methods/.
21. Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. Psychometrika, 56(2), 177–196. https://doi.org/10.1007/BF02294457.
22. National Academies of Sciences, Engineering, and Medicine. (1997.) Improving schooling for language-minority children: A research agenda. The National Academies Press. https://doi.org/10.17226/5286.
23. Oliver, R., & Azkarai, A. (2017). Review of child second language acquisition (SLA): Examining theories and research. Annual Review of Applied Linguistics, 37, 62–76. https://doi.org/10.1017/S0267190517000058.
24. Papanastasiou, C. (2000). Internal and external factors affecting achievement in mathematics: Some findings from TIMSS. Studies in Educational Evaluation, 26, 1–7. https://doi.org/10.1016 /S0191-491X(00)00002-X.
25. Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. John Wiley & Sons. http://dx.doi .org/10.1002/9780470316696.
26. Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. Health Services & Outcomes Research Methodology, 2, 169–188. 10.1023/A:1020363010465.
27. Samuelson, B. L., & Freedman, S. W. (2010). Language policy, multilingual education, and power in Rwanda. Language Policy, 9(3), 191–215. 10.1007/S10993-010-9170-7.

Appendix

Table A-1: Coverage and Exclusion Rates in UAE

Coverage	Overall exclusion	Schools	Students
100%	5.6%	2.6%	3%

Source: TIMSS (2019)

Table A-2: Variables' List, coding and meaning

List of Variables	Coding	Meaning
Overall Performance Math1-Math5	asmmat 01–05	The 1st to 5th plausible value of Overall Performance in Mathematics
Overall Performance Science1-Science5	asssci 01–05	The 1st to 5th plausible value of Overall Performance in Science
Age	asdage	Quantitative variable which indicates student's age.
Sex	asbg01	Dummy variable which takes the value 1 for female and 0 for male.

Pre-Education	asbh04b	Quantitative variable which indicates the number of years in pre-primary education for each student.
Parents' Education	asdhedup	Categorical variable which reflects parents' education level as follows; some primary, upper secondary, postsecondary and university or higher. The category some primary is considered as reference category.
Resources	asdghrl	A score calculated based on the number of books and other study materials in the students' homes, their parents' level of education, and their parents' employment. Scores were used to define students into three categories: students with Many Resources, students with Some Resources and students with Few Resources. The category few resources is considered as a reference category.

Source: TIMSS (2019)