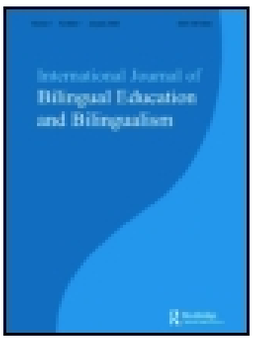

International Journal of Bilingual Education and Bilingualism

ISSN: 1367-0050 (Print) 1747-7522 (Online) Journal homepage: https://www.tandfonline.com/loi/rbeb20

A burden or a boost: The impact of early childhood English learning experience on lower elementary English and Chinese achievement

Si Chen, Jing Zhao, Laura de Ruiter, Jing Zhou & Jinzhen Huang

To cite this article: Si Chen, Jing Zhao, Laura de Ruiter, Jing Zhou & Jinzhen Huang (2020): A burden or a boost: The impact of early childhood English learning experience on lower elementary English and Chinese achievement, International Journal of Bilingual Education and Bilingualism, DOI: 10.1080/13670050.2020.1749230

To link to this article: https://doi.org/10.1080/13670050.2020.1749230

Published online: 15 Apr 2020.

INTERNATIONAL JOURNAL OF BILINGUAL EDUCATION AND BILINGUALISM
https://doi.org/10.1080/13670050.2020.1749230

A burden or a boost: The impact of early childhood English learning experience on lower elementary English and Chinese achievement

Si Chen [a], Jing Zhao [b], Laura de Ruiter [c], Jing Zhou[d] and Jinzhen Huang[e]

[a]Graduate School of Education, Harvard University, Cambridge, MA, USA; [b]Department of English, Sun Yat-sen University, Guangzhou, People's Republic of China; [c]Department of Psychology, Tufts University, Medford, MA, USA; [d]Faculty of Education, East China Normal University, Shanghai, People's Republic of China; [e]Kunshan Kid Castle English Training School, Kunshan, People's Republic of China

ABSTRACT

We examined the effect of learning English as a foreign language in early childhood on children's later English and Chinese achievement and attitude toward English learning in mainland China. A total of 892 children were divided into two groups according to whether they had any English learning experience before they entered elementary school. To reduce self-selection bias in children's early English learning experience, we employed propensity score matching (PSM) and coarsened exact matching (CEM) techniques and generated balanced samples based on children's demographic, parent-child interactional and socio-economic characteristics. We found that the experience of early childhood English learning alone positively contributed to later English language and Chinese language achievement and also to attitudes toward English learning. The findings provide important evidence that limited exposure to English before elementary school does not harm Chinese learning. They also have educational implications that encourage educators to continue helping students develop genuine interest in English learning and maintain their high level of motivation starting at a very young age.

ARTICLE HISTORY
Received 6 July 2019
Accepted 23 March 2020

KEYWORDS
Young learners; English as a foreign language; Chinese; motivation; age of acquisition

Introduction

Across the globe, policy makers are mandating earlier starting ages for foreign language learning in schools (Enever 2012). In many cases, this foreign language is English, because of its dominating role as a lingua franca world-wide (Butler and Le 2018). For example, in South Korea, English used to be taught only in secondary school and at universities, but since the 1990s has been a regular subject in the 3rd grade of primary schools (Song 2018). In Japan, English used to be first introduced as a school subject in 7th grade, but since 2011, instruction in public schools now starts in 5th grade (Hashimoto 2011). Mexico in 2009 also expanded English language education in public schools from the previously 7th–9th grades to starting in kindergarten (Sayer 2018), resulting in an increase in compulsory English instruction from three years total to ten years.

The main motivation for these early starts is the assumption that 'earlier is better,' – in other words, that age of acquisition is an important and robust predictor in foreign language learning success. It is

CONTACT Jing Zhao zhaoj67@mail.sysu.edu.cn
This article has been republished with minor changes. These changes do not impact the academic content of the article.

assumed that if children start learning English from a very young age, they will reach higher proficiency levels compared to starting at a later age. Consequently, they would have a head start in today's globalized world. On the other hand, there are also concerns attached to early foreign language learning: that learning another language before the first language (L1) is fully acquired could have detrimental effects on the L1, or on both the L1 and the L2 (Cummins 1979). Another concern is that beginning foreign language instruction early could negatively affect children's motivation and attitude toward foreign language learning and increase language anxiety.

Against this background, this paper provides new evidence for three issues: the effect of age of acquisition on very young learners' English achievement, effects of early English language learning (ELL) on the L1, and effects of ELL on learners' language learning attitudes. The data come from China, the country with the largest population of English as a foreign language (EFL) learners (Bolton 2003; He and Zhang 2010; Hu 2007). China belongs to the countries in the 'third wave of early language learning' identified by Johnstone (2009), that is, countries that started introducing early language learning at the beginning of the twenty-first century (mostly Asian countries). However, the situation in China is different from that of its neighbors. Specifically, elementary schools are not allowed to teach English before 3rd grade, which shifts all early language learning (before age six) to the private sector. This poses a particular methodological challenge for researchers who want to study the effects of ELL, as parents who have the financial means to send their children to private institutions are more likely to come from a background with a higher socio-economic status (SES), which is a factor known to positively affect academic success (Sirin 2005; von Stumm et al. 2019; Zhang, Hu, Ren, & Zhang, 2019). To address this potential confound, we employ modern statistical techniques called sample matching techniques, which allow us to go beyond purely correlational findings and draw causal inferences.

In the remainder of this introduction, we will first give an overview of what is currently known about the effects of age on foreign language learning, crosslinguistic transfer (i.e. effects of the L2 on the L1), and attitudes toward foreign language learning. We will then consider environmental factors such as SES and discuss some of the methodological problems that arise from these factors. Finally, we will describe the current educational and societal situation in China and similar contexts with respect to English language learning.

Age effects in foreign language learning

A widely held belief among the general public is that starting early is a major way not only to give children a head start in learning a foreign/second language but also to ensure native-like proficiency of the L2. The assumption is based on the Critical Period Hypothesis (CPH) (Lenneberg 1967), which states that language must be learned before puberty if normal, native-like levels are to be achieved. Although the CPH was originally formulated for L1 acquisition, over the years a substantial body of evidence has emerged that has found clear negative correlations between age of acquisition (AoA) and L2 proficiency (for overviews, see Nikolov and Djigunovic 2006; Singleton 2001), supporting the notion that there is a developmental window for language learning in general. While some late learners do achieve native-like proficiency in the L2 (Birdsong 2007; Marinova-Todd 2002; Marinova-Todd, Marshall, and Snow 2000), many do not (e.g. Han and Selinker 2005; Johnson et al. 1996). Research indicates that older learners typically have disadvantages compared to younger learners in oral and written grammaticality judgment tasks (Birdsong 1992; McDonald 2006) and listening and pronunciation skills (DeKeyser and Larson-Hall 2005). However, the 'the younger, the better' assumption has largely been based on studies of naturalistic L2 acquisition (i.e. immersive contexts), and generalizing those findings to instructional settings is problematic (Muñoz 2014). In fact, the majority of studies have found that learners who started later are either as good as or better than early learners in a number of measures (e.g. Al-Thubaiti 2014; Larson-Hall 2008; Pfenninger and Singleton 2016). In a meta-analysis of 42 studies, Huang (2016) found that older learners had advantages over younger learners both in the short and in the long term. Pfenninger and Singleton (2016) found

that early starters in instructed foreign language settings initially showed some advantage in receptive vocabulary and written lexical richness but were caught up soon by late starters, indicating that late starters had better rate of EFL learning, and late starters outperform early starters on oral and written accuracy.

Studies conducted specifically in East Asia have produced mixed results. A large-scale study in South Korea compared language performance and language learning attitudes in a cohort of high-school students that received English instruction in primary school vs. a cohort that did not (Butler 2015). The younger learners outperformed the older learners both in language measures and in attitudinal indicators such as their motivation to learn English. Among Japanese college students, Larson-Hall (2008) found that those that had received English instruction before age 12 were better than later learners in a phonemic task, but that there was no advantage in grammaticality judgment. The result held true even after controlling for the total hours of input.

Taken together, some studies find advantages for earlier starters, but the majority do not. One possible explanation for the observed differences between younger and older learners is that young children use different learning mechanisms compared to teenagers or adults. DeKeyser and Larson-Hall (2005) suggest that children learn more implicitly, while adults learn more explicitly. According to DeKeyser (2003), children's ability to learn implicitly decreases between early childhood and puberty. As a consequence, learners past puberty would have to rely largely on explicit learning strategies and at what age children begin to lose their ability to learn implicitly may be subject to individual variation (Dixon et al. 2012). Younger children may not be able to benefit from more cognitively oriented learning that is common in classrooms (Pfenninger and Singleton 2017). They may need substantial language input in a supportive learning environment and plenty of opportunities to experience the language (Muñoz 2011).

Crosslinguistic transfer between L1 and L2

Those who are cautious about getting children started learning English at an early age – often parents – are sometimes concerned that learning English may be getting in the way of a child's acquisition of their native language. The Linguistic Interdependence Hypothesis (Cummins 1979) highlights that second language competence is partly dependent on the level of competence already achieved in the first language, and that L2 exposure can have detrimental effects on L1 development: '[F]or children whose L1 skills are less well developed in certain respects, intensive exposure to L2 in the initial grades is likely to impede the continued development of L1' (Cummins 1979, 233). A number of studies on young learners in naturalistic settings have produced support for linguistic interdependence (crosslinguistic transfer) in various language skills, including general language proficiency, vocabulary, narrative skills, and reading (Hao et al. 2019; Hwang, Shin, and Hartsuiker 2018; Sparks, Patton, and Luebbers 2019). Working with very young EFL learners aged three to six years, Sun et al. (2016) found that Chinese vocabulary was predictive of English vocabulary. Two of their measures were paradigmatic and syntagmatic knowledge, which they measured by asking children to provide definitions for words (e.g. 'What is a dog? Tell me everything you know about a dog.') and identifying a word's semantic category (e.g. 'Is "dog" an animal or a fruit?'). They found that Chinese paradigmatic knowledge was a significant predictor of English paradigmatic and syntagmatic knowledge. The results support Snow and Kim (2007), who argue that a large vocabulary in the L1 usually means that the child has more concepts and better metalinguistic skills, which is conducive to learning words in the L2. It should be noted that previous research indicates that these transfer effects are modulated by a number of factors. The effects are generally stronger with higher levels of L1 proficiency (Sparks, Patton, and Luebbers 2019), stronger for language skills that involve smaller learning challenges (vocabulary) compared to those requiring a broader range of knowledge (oral language), and stronger when the L1 and L2 are typologically similar (Proctor et al. 2010). So far, the findings thus support the hypothesis that L1 skills are predictive of L2 acquisition. Importantly, however, previous research does not suggest that L2 acquisition negatively affects L1 acquisition.

In cases in which early language learners lag behind later language learners in L1 literacy, these differences usually disappear quickly (Pfenninger and Singleton 2017). Li et al. (2012) compared Chinese children participating in an English-Chinese immersion program in primary school with non-immersion students. They found that immersion students outperformed non-immersion peers in phonological awareness, and that English phonological awareness was predictive of English reading comprehension, but no crosslinguistic transfer was found from Chinese phonological awareness to English reading. While non-immersion students outperformed immersion students on Chinese phonological awareness in Grade 2, the difference disappeared in Grade 4, again suggesting that early and intensive L2 exposure does not lead to long-term disadvantages in the L1.

SES and learning English as a foreign language

A host of environmental (also called external) factors have also been found to be significantly related to young foreign language learners' acquisition, especially in instructional settings (Muñoz 2014). These factors include quantity and quality of the input, amount of practice, level of motivation, type of instruction, and social variables, such at home literacy environment, parental education, and parental beliefs about English education (Dixon et al. 2012). Socio-economic status in particular is related to young EFL learners' language learning outcomes (Butler 2014; Butler and Le 2018; Hoff 2013).

Butler (2014) examined the effect of Chinese parents' SES and their behaviors and beliefs about English education on their children's English language learning, and how such relationships may differ across different grades. She found that parents from all socio-economic backgrounds believe that effort should account for about 70% of a child's success in acquiring English, as opposed to language learning aptitudes and other factors. Compared to lower-SES parents, higher-SES parents supported their children's learning both through more direct behaviors (e.g. helping their children with their studies) and more indirect behaviors (e.g. having many English-language books at home). While these differences did not have an effect on the children's listening and reading performance in elementary school, a relationship between SES and speaking abilities started to emerge at the fourth-grade level, which was one year after children began to receive formal EFL education. In another longitudinal study on parental SES and student learning English in China, Butler and Le (2018) found that for middle school students, parents' educational level was positively associated with students' English learning outcome across grades. Parents' income was correlated with English learning only at the higher and the lower ends of the income distribution. Students whose parents were in the middle-income group did not show this correlation of income with their English learning outcome. This study provided important insights into the longitudinal correlation between SES and Chinese students' English learning. However, the small sample size and the fact that the study did not control for self-selection bias constrain the conclusions that can be drawn from it. Specifically, it is not possible to conclude that parental income is *causing* changes in English grades.

Attitude in foreign language learning

It is generally believed that motivation in language learning mediates language learning behaviors and therefore has a great impact on language learning outcome (Csizér and Dörnyei 2005). In the foreign language context, where input is minimal and learning happens mainly through formal instruction, Lamb (2007) found students' interest deteriorates as they proceed through school and particular when the teachers use teacher-centered approaches with few communicative activities. In the Chinese context, university students reported higher prevention instrumentality (e.g. learning English in order to pass an exam) compared with their peers who were studying in an ESL context and were members of the L2 community (Li 2014). Although foreign language motivation was found to be closely related to achievement among young learners (Hu and McGeown 2020), explaining

approximately 40% of the variance, many studies have reported a clear decline in students' intrinsic motivation and an increase in anxiety as they proceeded from lower to upper grades (Butler and Le 2018; Carreira 2011; Hu and McGeown 2020).

It is worth noting that a large-scale study of English learners in Chinese secondary schools (You and Dörnyei 2016) found that students' intended efforts in learning English are more due to more interest in the language (e.g. 'I find learning English really interesting') than to instrumental motivation (e.g. 'I study English because I will need it for future studies'). Therefore, to maintain young learners' interest in language learning may be more important than to emphasize academic performance. It is, however, unclear, if an early start in ELL affects attitudes: García Mayo and Garcia Lecumberri (2003) found that that early starters tend to have positive attitudes toward language learning, but Larson-Hall (2008) did not find any influence of starting early on language attitude in Japanese learners of English.

Furthermore, most of the studies that have examined the relationship between Chinese students' motivation and learning have been conducted in upper primary grades (Butler 2017), middle school (Butler and Le 2018) and universities (Liu and Thompson 2018; You and Dörnyei 2016; Zheng 2012). We know very little about language attitudes and motivation of very young learners learning English as a foreign language. Some studies have furthermore found gender effects in attitudes toward English language learning: Liu and Thompson (2018) found that among Chinese university students female learners had stronger ideal L2-self (e.g. 'I can imagine myself studying and living abroad and using English for communicative purposes') than male students. It is unclear if similar differences may also be present in young learners.

English language learning in China

English education from kindergarten to tertiary institutions attracts immense investment from the government, educators, parents, and learners. In 2001, the *Guideline for Promoting English Teaching in Elementary Schools* made English a required subject in elementary schools from Grade 3 (Ministry of Education 2001). Under this guideline, the Ministry of Education of China specifically asked elementary schools to reduce Chinese lessons by one hour and include at least four English teaching sessions per week in the curriculum. Since then, teaching EFL to young learners has grown in popularity around China. While Chinese public elementary schools are not allowed to introduce English earlier than 3rd grade, children in private institutions have started to learn English at increasingly younger ages. It is reported that about 210 million childhood EFL learners are taking English lessons in more than 50,000 private English institutes in mainland China (Sun et al. 2016).

There seem to be several reasons behind this 'unprecedented passion to learn English,' as described in Feng (2012, 364), for very young learners. First, Chinese parents believe that 'earlier is better,' and are afraid of missing the 'critical period' for learning a second language. Second, achievement in English is considered a vital component of an individual's overall competence and competitiveness and an important tool for educational advancement and career opportunity (Li, 2014; You and Dörnyei 2016); therefore, the instrumental motivation is high. Investment in English learning at a younger age is considered effective for producing more benefit at lower cost. Third, with the use of English not constrained to traditional classroom use and to taking exams, intrinsic motivation has actually increased to an unparalleled level (Feng 2012). Some families are able to create English-speaking communities within the larger Chinese-speaking society. The amount of authentic input increases through interacting with native-English speaking persons, use of imported materials like books, audio and video clips, apps, and participation in one-on-one online classes, such as *vipkid* and *51talk*.

Despite the rising popularity of early English instruction in private settings, there has been open discussion in the media in recent years about reducing and even banning the earlier introduction of English education through government policies. For example, the Ministry of Education mandated that preschools and kindergartens do not teach content that is supposed to be taught at elementary

school level (Ministry of Education 2017). The policies and the surrounding debate reflect the issues discussed above, in particular the concern that starting too early could have negative effects. The time spent learning English would diminish time spent learning Mandarin. Other concerns come also into play. It is feared that the shortage of qualified EFL teachers would degrade the quality of EFL education, with two potential consequences: First, the introduction of EFL education across the board would reduce the quality of early childhood and elementary school education in general. Second, low-quality EFL education may lead to negative learning experiences, which could in turn reduce children's motivation and enthusiasm for English learning. Despite these concerns, many parents are eager to have their children learn English from a very young age, and achieve this through home teaching, private tutoring, and early childhood EFL cram schools for children between three and six years, and sometimes for those even younger.

Methodological issues in current EFL research

Previous studies have shown that that children's English attainment is often confounded with environmental factors (Butler 2014, 2015; Butler and Le 2018; Sun et al. 2016), such as parents with higher education levels being more likely to provide more resources for their children's language learning. This is a problem when the aim is to study the causal link between early EFL education and academic achievement. Without experimental controls (e.g. a randomized controlled trial), it is difficult to argue that early childhood EFL education plays a causal role in students' later academic achievement and learning attitude. As discussed above, students may perform better in English and Chinese tests because their families provide more resources on education in general, not specifically because they have early childhood EFL education. For example, Knell et al. (2007) studied the benefits of children participating in an immersion program. The authors administered a series of language tests to primary school students in northwestern China, half of which were enrolled in an immersion program. They found that the immersion students performed significantly better on all English measures and they were especially better at speaking and listening in English. In addition, the immersion students performed better on Chinese character recognition. However, the immersion program was implemented in a school that was originally designated for government officials only and later opened to a small portion of local elites. Thus, the children at that school almost certainly came from a different socio-economic background than those in other schools, and may have benefitted particularly because of their supportive home environment (see above). It is thus unclear to what extent similar gains through immersion would occur in other socio-economic contexts. Similarly, in the study comparing English language learning gains of early and late starters in Japan mentioned earlier (Larson-Hall 2008), the authors did not measure SES and other environmental factors, either.

A traditional method to explore the predictive power of early childhood EFL education on later academic achievement is to control covariates (e.g. SES). However, even when controlling for many critical covariates, classic regression cannot solve the confounding issue, since the outcome variable is still involved in the process of controlling the covariates. An alternative to controlling covariates is matching. Matching methods aim at achieving a balanced sample to compare two groups (e.g. children with early EFL education and children without) while the whole process is blinded to the outcome (e.g. the students' EFL achievement). By preprocessing the data through matching, we approximate a randomized controlled trial in order to draw causal inference from observational data (see Stuart 2010 for more information on matching). Propensity score matching (PSM), developed by Rosenbaum and Rubin (1983), is the most popular matching method; however, other matching methods could also be used to address specific research questions.

The present study

The purpose of this study is to estimate the effect of early childhood English learning experiences on lower elementary students' English and Chinese academic achievement with minimal bias. To

address the issue of confounding, we use both propensity score matching and coarsened exact matching to address selection bias due to children's observed demographic and parent-child interactional and socio-economic characteristics. After achieving an ideal balance between the treatment and the comparison group, we condition on student's age, gender, and SES indicators, and answer three research questions:

(1) Does early English learning experience improve first and third grade elementary students' English learning outcome?
(2) Does early English learning experience undermine first and third grade elementary students' Chinese learning outcome?
(3) Does early English learning experience have an impact on first and third grade elementary students' attitude toward English learning?

Method

Participants

Collaborating with the District Bureau of Education, we collected participants' information in one school district of a central urban area of a typical metropolitan city in eastern China. Six public elementary schools (22% of the public schools within the district) agreed to participate in this study. Note that some public schools in Chinese urban areas have been named 'key/experimental school' to indicate their richness of education resources (more funding, higher teacher-student ratio, and better teaching quality). In the school recruitment procedure, we avoided the two key schools in the district. Therefore, none of the schools in our sample was a key elementary school or experimental school. All the schools in our sample were public schools managed by the District Bureau of Education. For each grade, their curriculum, teacher training, daily class plan, daily homework, and the final examinations of each semester were centralized. Thus, we assumed that the variation in teaching quality between the schools was limited, and that the scores of the exam were comparable between schools. We collected cross-sectional data from the first grade and the third grade (7- to 9-year-olds). In total, 892 students from 26 classes and their parents participated in this study.

We collected the English and Chinese test scores of all students. Eight hundred and thirty-one (N = 831) students and parents completed the questionnaire. All students and parents were Mandarin Chinese speakers. Table 1 shows the characteristics of the sample before propensity score matching and coarsened exact matching.

Table 1. Sample characteristics by early English learning status (Unmatched).

	Full Sample $N = 892$	With early English learning experience $N = 564$	Without early English learning experience $N = 328$
Female	46.49%	46.45%	43.6%
Age	8.06 (1.37)	8.05 (1.41)	8.10 (1.31)
Mother education			
High(≥College)	42%	47%	33%
Low(≤High school)	58%	53%	67%
Parent English level			
High	41%	46%	39%
Low	59%	54%	61%
Parent-child level			
Often book reading	40%	44%	37%
Not often book reading	60%	56%	63%
Often teach English	37%	47%	22%
Not often teach English	63%	53%	78%
Family income	24,385 (16,098)	26,031 (16,223)	21,939 (15,572)

Approximately 63% ($N = 564$) of the parents reported that their child attended English learning classes before elementary school. A comparison between those children who had early childhood EFL education experience and those who had no early childhood EFL education experience showed that these groups did indeed differ from each other. The data confirmed our suspicion that children who had early childhood EFL education experience may come from higher-SES backgrounds. In addition, slightly more girls had early childhood EFL education experiences. Mothers of children with early childhood EFL education experiences had a significantly higher education level (47% had college or higher education level) and significantly higher English proficiency level (46% reported a high level of English proficiency) than mothers of children without early childhood EFL education experiences (33% had college or higher education level; 39% reported a high level of English proficiency). A greater percentage of parents of children with early childhood EFL education experiences reported a higher frequency of child-parent interaction (44%), compared to parents of children without early childhood EFL education experiences (37%). The average family income of children with early childhood EFL education experiences was 0.26 standard deviations above that of children without early childhood EFL education. In general, children with early childhood EFL education were more likely to come from higher-SES families. This is clear evidence that sophisticated matching models are needed to answer the research questions and draw causal inference.

Measures

All data for this research came from parent and student questionnaires and direct child examination scores gathered from elementary schools. The outcome, predictor, and control variables were used in one of two methodological steps: (1) propensity score matching and coarsened exact matching models, or (2) multiple regression models.

Outcome measures

English test score. The English test score of each child in this study is the unified English test score of the school district in the fall semester of the 2013–2014 school year. The school district had a unified examination of English, Chinese, and mathematics at the end of each school year. Students and schools were not notified of the raw score of the test. The final scores that students and schools received were standardized by the mean and standard deviation of the raw scores of the whole school district and converted into four ranked levels: A (excellent), B (satisfactory), C (sufficient), and D (fail). In this sample, 68% of students received As. 32% students received a B, C, or D. We then recoded the English test score as a binary variable[1]: An English test score = 1 means that the student got an A; an English test score = 0 means that the student got a B, C, or D.

Chinese test score. The Chinese test scores of this study are the unified Chinese test scores of the school district in the fall semester of the 2013–2014 school year. As with the English scores in this study, we standardized four-ranked levels of students' Chinese test scores: A (excellent), B (satisfactory), C (sufficient), and D (fail). In this sample, we had 74% students get an A. We recoded the Chinese test score as a binary variable: a Chinese test score = 1 means the student got an A; a Chinese test score = 0 means that the student got a B, C, or D.

Attitude. Attitude is a binary variable capturing students' attitude toward English learning. In the student questionnaire, we asked the students: 'Do you enjoy learning English?' We code the answers of these students. Attitude = 1 means that the student answered 'Yes, I enjoy learning English'; Attitude = 0 means that the student answered 'No, I don't enjoy learning English.' In total, 77% students answered with Attitude = 1.

Predictor.

Early childhood EFL education. The predictor in this study is the binary variable early childhood EFL education. In the parental questionnaire, we asked parents if the student had any English learning experience before she/he entered elementary school. We emphasized that the early English learning could happen in school (kindergarten) or outside school (e.g. after school centers or weekend projects). Early childhood EFL education was coded as 1 if the parent answered 'Yes'; early childhood EFL education was coded as 0 if the parent answered 'No.' In our sample, 64% parents reported their children had early childhood EFL education.

Control variables.

We selected critical control variables that matter to young children's English and Chinese academic achievement, and which may influence parents' choice of whether or not to provide their children with early childhood EFL education. Theses control variables included children's age, their gender, the mother's education level, the mother's level of English proficiency, family income, shared book reading time, and parental teaching of English to children. Mother's education level is a dichotomous variable. A score of 1 indicates that the mother has a college or above degree. A score of 0 indicates that the mother has a high school or lower diploma. Mother's level of English proficiency is self-reported English proficiency level. A score of 1 means that the mother's English skill is equal to or above that of the College English Test Band 4 (CET-4) level. A score of 0 means the mother's English level is below the CET-4. Family income is a self-reported variable of family annual income. Shared book reading time is also a dichotomous variable. A score of 1 means that the parents reported that they 'often spend time reading picture books with my child,' while a score of 0 means they do 'not often spend time reading picture books with my child.' We asked parents to report if they 'often teach (their) child English' (=1) or 'do not often teach (their) child English' (=0).

Analytic approach

The ultimate goal of the matching method is to achieve a balanced sample prior to further statistical analysis. In an ideal yet impossible randomized experiment, we could randomly assign children to either receive or not receive early childhood EFL education and then estimate the effect of early childhood EFL education on students' later achievements without bias. In an observational study, we can instead use matching methods to construct two comparable samples of children (i.e. with or without early childhood EFL education). Matching allows having two groups of students that are similar with respect to demographic factors and family SES, and only differ with respect to early EFL education experience.

Multiple matching techniques

There are a number of different matching techniques, which all use different methods to reach a balanced sample. Propensity score matching (PSM) is the most widely used matching technique for nonexperimental causal studies (Dehejia and Wahba 2002). However, when a study's sample size is not large, a recommended practice is to adopt an additional, different matching method and test if the results are robust (Caliendo and Kopeinig 2008). We follow this recommendation and use two techniques: PSM and coarsened exact matching (CEM). Below we describe in more detail how we implemented these techniques in this study.

Our descriptive data confirms that students who had early childhood EFL education were from high-SES families. Hence, we use matching techniques to control for the potential confounding influence of pretreatment variables (i.e. children's age, gender, family SES) by reducing the imbalance between the treatment and control group. The variables in the two matching methods are the same.

Propensity score matching (PSM). We first generated propensity scores using all the control variables in Table 1. In the propensity score algorithm, we treated early childhood EFL education as a function of child and parental background features.

Coarsened exact matching (CEM). In addition to PSM, we also used CEM to reduce the selection bias of the sample. Similar to the PSM, the CEM modeled a balanced sample of the children with or without early childhood EFL education using covariates of child and parent's background information.

After we obtained a balanced sample of the 'treatment' (with early childhood EFL education) and 'control' (without early childhood EFL education) groups, we compared the two groups to allow causal inference regarding the effect of early childhood EFL education on later academic achievement and English learning attitude (see next section). A balanced sample, in our case, meant that those students who received early childhood EFL education and those who did not needed to be as similar as possible on the mean and standard deviation of critical covariates. Table 2 presents the balanced samples we obtained with the two matching methods.

Regression

Once we had a balanced sample from the PSM and CEM processes, we used logistic regression models to compare the students' English and Chinese achievement, as well as their English learning attitude. Since our outcome variables were binary, we used logit regression. In the logit models the log odds of our binary outcomes (English test score, Chinese test score, attitude) were modeled as a linear combination of the predictor variable (early childhood EFL education), controlling for covariates such as age, mother education, and family income. For illustrative purposes, we will also present the results of the regression using the unmatched sample.

Results

Balance check

Before the main analysis, we checked the balance of each covariate we used in the PSM and CEM models between the treatment and control groups to ensure that the matching had indeed generated a balanced sample. Before matching, six covariates had significant differences. This indicated that early childhood EFL education depends on the characteristics of young children and parents themselves. Girls are more likely to learn English prior to elementary school. Children with parents

Table 2. Effects of early English learning experience on elementary school English test scores.

	Model 1. Unmatched sample	Model 2. Propensity score matching	Model 3. Coarsened exact matching
Early English learning	0.72***	0.64***	1.00***
	(0.18)	(0.14)	(0.23)
Age	−0.31***	−0.39***	−0.33***
	(0.07)	(0.06)	(0.10)
Gender	0.70***	0.92***	1.00***
	(0.18)	(0.15)	(0.25)
Mother education	0.84***	1.01***	1.39***
	(0.23)	(0.21)	(0.34)
Mother English level	0.41*	0.75***	0.72*
	(0.20)	(0.16)	(0.29)
Parent-child interaction	0.23	0.05	−0.04
	(0.18)	(0.15)	(0.26)
Home teach English	−0.04	0.09	−0.23
	(0.18)	(0.15)	(0.24)
Family income	0.00	−0.00001**	0.00
	(0.00)	(0.000005)	(0.00)

with higher education, family income, and English proficiency have a higher likelihood of being enrolled in early childhood EFL programs.

Both PSM and CEM generated well-balanced datasets. A t-test showed that the mean differences between all the covariates were minimal. Based on the mean and standard deviation of each covariate, CEM performed slightly better than PSM in our sample in that the two samples were more similar in after CEM than after PSM.

RQ1: does early English learning experience improve first and third grade elementary students' English learning outcomes?

We present three models using different datasets (the unmatched sample, the PSM-matched sample, and the CEM-matched sample). According to Model 1 (see Table 2), which uses the unmatched full sample, early childhood EFL education significantly predicts a student's English test score after controlling for all covariates. For a student who has received early childhood EFL education, the log odds of scoring an A in the English test increase by 0.72 ($p < 0.001$). Model 2 and Model 3 use matched samples from the PSM and CEM techniques, respectively. After controlling for the student's age, gender, maternal education level, mother's English level, parent-child interaction, and family income, we find that for students who have received early childhood EFL education, the log odds of achieving an A in the English test significantly increase by 0.64 (Model 2, $p < 0.001$) and 1.00 (Model 3, $p < 0.001$). This means that students who started to learn English prior to elementary school were more likely to have an A score on their English exam in elementary school. The coefficients in Model 2 and Model 3 were slightly different because the sample size and matching technique were not identical. Yet in both models the chances of getting an A was significantly higher for students who had received early childhood EFL education.

RQ2: does early English learning experience improve first and third grade elementary students' Chinese learning outcomes?

We present three models to answer the second research question in Table 3. Model 4, Model 5, and Model 6 are fitted with the unmatched sample, the PSM dataset, and the CEM dataset, respectively. In Model 4, the naive estimation shows that the early English learning experience significantly predicts the log odds of having an A in the Chinese exam. In Model 5 and Model 6, after controlling for the student's age, gender, maternal education level, mother's English level, parent-child interaction, and family income, the PSM model shows that for students who have received early childhood English

Table 3. Effects of early English learning experience on elementary school Chinese test scores.

	Model 4. Unmatched sample	Model 5. Propensity score matching	Model 6. Coarsened exact matching
Early English learning	0.38*	0.38**	0.57*
	(0.18)	(0.15)	(0.24)
Age	−0.22**	−0.32***	−0.26*
	(0.07)	(0.06)	(0.10)
Gender	0.41*	0.40**	0.55*
	(0.18)	(0.15)	(0.25)
Mother education	0.61**	0.63***	0.70*
	(0.23)	(0.21)	(0.33)
Mother English level	0.48*	0.37*	0.47
	(0.21)	(0.17)	(0.30)
Parent-child interaction	0.13	−0.22	0.03
	(0.19)	(0.15)	(0.27)
Home teach English	0.10	0.21	0.48
	(0.18)	(0.15)	(0.25)
Family income	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)

education, the log odds of achieving an A in the Chinese test increase by 0.38 ($p < 0.01$), while the CEM model shows an increase by 0.57 ($p < 0.05$). All the coefficients are statistically significant.

RQ3: does early English learning experience have an impact on first and third grade elementary students' attitude toward English learning?

To answer our third research question about the impact of early childhood EFL education on students' attitude toward English learning, we fit Model 7 ('naive'), Model 8 (PSM) and Model 9 (CEM) (See Table 4). After controlling for the student's age, student's gender, maternal education level, mother's English level, parent-child interaction, and family income, the estimates of the unmatched sample (Model 7) and CEM sample (Model 9) showed no statistically significant difference between students with or without early childhood EFL education. In contrast, the PSM model (Model 8) showed that for students who had received early childhood EFL education, the log odds of a student saying that they enjoy learning English significantly increase by 0.11 ($p < 0.05$) compared to students who had not received early childhood EFL education.

Discussion

Researchers, parents and policy-makers world-wide are engaged in an ongoing debate around the potential benefits and risks associated with early foreign language learning. Many countries have begun to mandate earlier starting ages in institutional settings (i.e. compulsory primary and secondary education), believing that 'earlier is better.' In China, in contrast, public elementary schools are not allowed to teach English earlier than 3rd grade. Despite – or because of – this policy, English education in China is a growing industry, and some parents have their children learn English from increasingly young ages, hoping that this will provide them with an advantage in a competitive educational and work environment. At the same time, many parents and educators are concerned about potential adverse effects of early English education, worrying that exposure to another language could have negative effects on children's learning of Chinese and on their attitude toward learning English.

In this study, we explored the impact of early childhood EFL program experience on Chinese students' elementary school English and Chinese achievements, as well as on their attitude toward English learning. Importantly, the study controlled for environmental factors such as socioeconomic status, correcting for the fact that children who receive early English education tend to come from families with higher levels of education and higher income. Using two matching techniques, propensity score matching and coarsened exact matching, we prepared a well-balanced dataset to reduce

Table 4. Effects of early English learning experience on English learning attitude.

	Model 7. Unmatched sample	Model 8. Propensity score matching	Model 9. Coarsened exact matching
Early English learning	0.09	0.11*	0.40
	(0.07)	(0.05)	(0.24)
Age	0.08	0.003	−0.01
	(0.07)	(0.02)	(0.10)
Gender	0.14*	0.21***	0.34
	(0.06)	(0.05)	(0.25)
Mother education	0.13	0.24**	0.70*
	(0.09)	(0.07)	(0.33)
Mother English level	0.12	0.19***	0.47
	(0.08)	(0.05)	(0.30)
Parent-child interaction	0.21***	0.20***	0.03
	(0.06)	(0.05)	(0.27)
Home teach English	0.06	0.11*	0.48
	(0.06)	(0.05)	(0.25)
Family income	−0.000004*	−0.00001	0.00
	(0.000002)	(0.000002)	(0.00)

this self-selection bias and investigated the causal link between early childhood English learning experience and academic achievements and learning attitude. Our results show that compared with children who did not receive early childhood English education, students with early childhood English learning experience have significantly higher chances of obtaining an A (the highest score) in both English and Chinese exams. This shows that learning English from an early age can indeed improve English achievement. Starting early often means early exposure to the target language, and early exposure (to the target language) implies more time spent learning the language, and it moderates the effect of age on foreign language learning (Muñoz 2011). In the Chinese context, by the time EFL education is first introduced at the beginning of Grade 3, a child who attends a private school with early EFL education (starting in kindergarten at age 3) will have already received approximately 400–1000 h of English instruction, assuming one lesson is 50 min and they take 2–5 lessons per week. However, future studies should also examine the characteristics of early English learning experience that children under go, in particular the amount of time they spend learning the language. According to Unsworth et al. (2015), children who spent less than 60 min per week learning a foreign language did not make more progress than children who did not spent any time learning a foreign language.

Recall from our introduction that the main concern shared by many parents and educators in China is that students who learn more English from a young age will lag behind their peers who grow up in a monolingual Chinese environment in their development of Chinese academic skills. Hoff (2013) stated that it is inevitable for bilingual learners to have less input in any one language because their input is shared by the two languages. While this is incontrovertibly true, our data show that concerns that English learning would negatively affect Chinese are not warranted. This finding is consistent with Li et al. (2012) and Knell et al. (2007), who found that children's Chinese learning in English-Chinese immersion program was not undermined by the fact they received more input in English.

Another concern often voiced in the public discourse in China is that early English instruction may have a detrimental effect on children's attitude toward learning English later on. Again, our data show that this concern is not justified. Early childhood English learning experience has either no effect or a positive effect on students' attitudes toward learning English when they are in elementary school. Parental factors also influence children's motivation and attitude to learn a foreign language. Parents with higher SES are more likely to provide opportunities for their children to experience the language in more authentic contexts (Butler 2015; Song 2018) and therefore, children may become more motivated to learn English. In the unmatched sample (which does not control for SES), we did not find any difference in attitude between early and late learners of English, echoing Larson-Hall's (2008) findings for Japanese learners of English. However, with the matched sample (controlling for gender and SES), attitude toward learning English was positively associated with early English learning experience. This shows again the importance of controlling for environmental factors such as SES when wanting to make claims about the effects of early language learning.

In a recent review Bialystok (2018) argues that 'there is no credible evidence that bilingual education adds or creates burden for children, yet it is incontrovertible that it provides the advantage of learning another language and possibly the cognitive benefits of bilingualism' (676). Our findings seem to support this: Children who had received English language instruction at an earlier age performed better in both English and Chinese exams and may have a more positive attitude towards learning English. We did not find any adverse effects of early English language instruction. Our findings are proof that learning English from a young age can have measurable benefits for Chinese children, without measurable downsides.

However, we caution that our findings do not imply that early English education is beneficial under all circumstances. Our study did not measure the quality of the English instruction, nor did it examine the type of instruction the children received. We note that our study was conducted in Shanghai, an international and economically better-off metropolitan city. This likely correlates both with higher quality of instruction and with a more stimulating and varied English learning

environment in general. Children may receive not only private tutoring or learn through commercial online apps like *Jiligula* and *ABCmouse*, but also experience English in their daily lives, for example through interaction with other English speakers, native or otherwise, or when visiting commercial places such as Disneyland. Access to digital media seems to have a positive impact on children's motivation to learn a foreign language even if this foreign language is rarely used in their immediate contexts (Al-Nofaie 2018). In Shanghai the boundary between a traditional foreign language setting and a second language setting (i.e. one in which the target language is actually spoken in the immediate environment) can become blurry. This provides a special setting, which does not necessarily reflect the reality of early English instruction in other parts of China.

Because our study did not investigate English instruction directly, we can only conclude that the type of instruction the children received was conducive to their English learning, without knowing what this meant in practice. Yet based on what is known about how children learn language in general we can make some assumptions on what characteristics 'good' early English instruction should probably have. Children acquire their first language out of communicative necessity. They learn language in order to communicate and interact with others, and their learning is influenced by the input they receive. We posit that their second language learning should take this into account. That means that children should learn English in a way that serves a communicative purpose, through interaction and communication in their daily lives, and much less through rote memorization or pattern drills. That said, future studies need to investigate the effect of different types of instruction empirically before any definitive recommendations can be made.

The type and quality of the instruction may also be the reason why the children in our study did not develop any negative attitudes toward English learning. Again, this result may not be generalizable to other areas of China, where less good instruction could potentially lead to more negative attitudes. This may also point to a possible link between positive/interesting experiences in English learning and attitudes towards learning English.

Despite these caveats and open questions, our findings do show that early English language experience can be an instrument to improve Chinese children's English (and Mandarin) proficiency. We will briefly discuss the implications for the larger context of English language education in China, especially as English has become as an academic subject, the success of which is closely related to more educational and professional opportunities (Butler 2015).

Implications

In many non-English speaking countries, the government's policy toward the role of English in society and also the procedure for implementing these policies in the educational system have great impacts on learners' acquisition of English (Sayer 2018; Song 2018; Tollefson 1989). In China, the government has begun to regulate when schools can start to teach English. At the moment, schools may not start teaching English at a grade level earlier than 3rd grade, with the argument that this measure will increase 'equal access' (Butler 2015). However, as our study showed, higher SES parents pay for private EFL classes and also provide more direct and indirect assistance in their children's English learning. This phenomenon is widely seen in big cities like Beijing, Shanghai, and Guangzhou. As a consequence, there is increasing diversity – and some may even say disparity – in EFL education in China.

An alternative to delaying the learning of English could be to implement early English learning programs nationwide. There are a number of practical concerns. Both proponents and opponents of early English intervention in contexts like China have realized that there are many constraints to fully implementing successful programs. Early childhood EFL programs are known to be influenced by issues such as language planning, language resource allocation, curriculum development, and EFL teacher education (Dixon and Zhao 2016; Muñoz 2014). If lawmakers and educators want to improve the quality of early childhood EFL programs, they should prolong their time of instruction, develop learning materials and methods that meet the students' needs and are within their scope of cognitive

abilities, and select and train qualified teachers that are skilled in both English instruction and other subject areas (Zhao et al. 2016). Teacher professional development in particular should be a long-term goal. Also, our findings indicate that some efforts may be necessary to address a gender disparity: Boys may need more attention than girls in maintaining motivation, as not only did more girls attend early childhood EFL education programs, they were also more likely to score A in the English and Chinese tests in elementary school. This may also help to explain the finding of Liu and Thompson (2018) that female students have higher motivation to learn English than male students in universities. On the parents' side, providing a rich home literacy environment, such as having more English books, creating more opportunities for parent-child interaction, and increasing shared book reading will also help with children's motivation and learning outcomes. Parents could be supported in their efforts by schools and teachers.

Limitations and directions for future studies

This study has three limitations. First, the measures for English achievement are not a direct measure; only letter grades are available. Future study should measure students' learning outcomes with measures with better score reliability and validity. Second, our study has only investigated the effect of having early childhood EFL education experience on the lower elementary grades' Chinese and English scores. Future studies should include delayed assessments to examine whether the effect of early English learning will persist in higher elementary grades. Finally, as discussed above, our study was conducted in a well-resourced and economically better-off metropolitan city (Shanghai). It is hard to make the same claims about or suggestions for other provinces or cities in China that do not have similar resources. Future studies should address this issue by examining the cost effect of investing so much in early childhood EFL education in different socio-cultural contexts. Learning experiences and learning outcomes of socio-economically and educationally disadvantages learners should be included in furthering this line of research.

Note

1. The reason for coding rank order data into dichotomous data is that the number of students who received C and D was very small, and the students who got A and not A can be considered to be systematically different. Students who received B, C, and D were largely at the same proficiency level.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Ministry of Education of the People's Republic of China [grant number 20YJC740102], the Fundamental Research Funds for the Central Universities [grant number 18wkzd12], and Shenzhen iRead Foundation [grant number 15000-71210046].

Notes on contributors

Si Chen is Postdoctoral Fellow of Education at Harvard Graduate School of Education. Among her recent publications are Longterm effects of China's One Village One Preschool Program on Elementary Academic Achievement (2019, Early Childhood Research Quarterly), and The Efficacy of a School-Based Book-Reading Intervention on Vocabulary Development of Young Uyghur Children: A Randomized Controlled Trial (2018, Early Childhood Research Quarterly). Her research is primarily focused on how language/literacy learning curriculums and environments for young children can be designed to support children's development.

Jing Zhao is Associate Professor of English at Sun Yat-sen University, China. Among her recent publications are Advantage in Character Recognition among Chinese Preschool Children with Autism Spectrum Disorder (2019, Journal of

Autism and Developmental Disorders), Contribution of Phonological, Morphological and Orthographic Awareness to English Word Spelling: A Comparison of EL1 and EFL Models (2017, Contemporary Educational Psychology), and English-Medium Instruction in Chinese Universities: Perspectives, Discourse and Evaluation (2017, Routledge, with L. Quentin Dixon). Dr. Zhao's research focuses on the influence of children's individual variation, school, and environmental factors on their L2 language and literacy development.

Laura de Ruiter is Research Assistant Professor at Tufts University. Her research focuses on child language acquisition across different languages. Among her recent publications are Interaction of information structure and clause order on children's processing of adverbial sentences (2020, Cognition), Iconicity affects children's comprehension of complex sentences: the role of semantics, clause order, input, and individual differences (2018, Cognition), and Quantifying sources of variability in infancy research using the infant-directed speech preference (2020, Advances in Methods and Practices in Psychological Science).

Jing Zhou is Professor of Education at East China Normal University, Faculty of Education. Dr. Zhou's research focuses on the language and literacy development of young children in China.

Jinzhen Huang is a director of Kunshan Kid Castle English Training School, China.

ORCID

Si Chen http://orcid.org/0000-0003-2025-1547
Jing Zhao http://orcid.org/0000-0002-1325-8698
Laura de Ruiter http://orcid.org/0000-0002-6590-4073

References

Al-Nofaie, H. 2018. "The Attitudes and Motivation of Children Towards Learning Rarely Spoken Foreign Languages: A Case Study From Saudi Arabia." *International Journal of Bilingual Education and Bilingualism* 21: 451–464. doi:10.1080/13670050.2016.1184612.

Al-Thubaiti, K. A. 2014. "Age of L2 Learning Makes no Difference in Instructed Settings: Input Matters Most." In *Teaching and Learning English in the Arabic-Speaking World*, edited by K. M. Bailey, and R. M. Damerow, 162–177. New York, NY: Routledge.

Bialystok, E. 2018. "Bilingual Education for Young Children: Review of the Effects and Consequences." *International Journal of Bilingual Education and Bilingualism* 21: 666–679. doi:10.1080/13670050.2016.1203859.

Birdsong, D. 1992. "Ultimate Attainment in Second Language Acquisition." *Language* 68: 706–755. doi:10.2307/416851.

Birdsong, D. 2007. "Nativelike Pronunciation among Late Learners of French as a Second Language." In *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, edited by O. Bohn, and M. J. Munro, 99–116. Amsterdam: John Benjamins.

Bolton, K. 2003. *Chinese Englishes: A Sociolinguistic Survey*. Cambridge, England: Cambridge University Press.

Butler, Y. G. 2014. "Parental Factors and Early English Education as a Foreign Language: A Case Study in Mainland China." *Research Papers in Education* 29: 410–437. doi:10.1080/02671522.2013.776625.

Butler, Y. G. 2015. "English Language Education among Young Learners in East Asia: A Review of Current Research (2004–2014)." *Language Teaching* 48: 303–342. doi:10.1017/S0261444815000105.

Butler, Y. G. 2017. "The Dynamics of Motivation Development among Young Learners of English in China." In *Early Language Learning: Complexity and Mixed Methods*, edited by J. Enever, and E. Lindgren, 167–185. Bristol: Multilingual Matters.

Butler, Y. G., and V. N. Le. 2018. "A Longitudinal Investigation of Parental Social-Economic Status (SES) and Young Students' Learning of English as a Foreign Language." *System* 73: 4–15. doi:10.1016/j.system.2017.07.005.

Caliendo, M., and S. Kopeinig. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys* 22 (1): 31–72. doi:10.1111/j.1467-6419.2007.00527.x.

Carreira, J. M. 2011. "Relationship Between Motivation for Learning EFL and Intrinsic Motivation for Learning in General among Japanese Elementary School Students." *System* 39 (1): 90–102. doi:10.1016/j.system.2011.01.009.

Csizér, K., and Z. Dörnyei. 2005. "The Internal Structure of Language Learning Motivation and its Relationship with Language Choice and Learning Effort." *The Modern Language Journal* 89: 19–36. doi:10.1111/j.0026-7902.2005.00263.x.

Cummins, J. 1979. "Linguistic Interdependence and the Educational Development of Bilingual Children." *Review of Educational Research* 49 (2): 222–251.

Dehejia, R. H., and S. Wahba. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84 (1): 151–161. doi:10.1162/003465302317331982.

DeKeyser, R. 2003. "Implicit and Explicit Learning." In *The Handbook of Second Language Acquisition*, edited by C. Doughty, and M. Long, 313–348. Malden, MA: Blackwell.

DeKeyser, R., and J. Larson-Hall. 2005. "What Does the Critical Period Really Mean?" In *Handbook of Bilingualism: Psycholinguistic Approaches*, edited by J. F. Kroll, and A. M. B. de Groot, 88–108. New York, NY: Oxford University Press.
Dixon, L. Q., and J. Zhao. 2016. "Bilingual Language Development." In *The Development of Language*, 9th ed., edited by J. B. Gleason, and N. B. Ratner, 285–307. New York, NY: Pearson.
Dixon, L. Q., J. Zhao, J.-Y. Shin, S. Wu, J.-S. Su, R. Burgess-Brigham, M. U. Gezer, and C. E. Snow. 2012. "What we Know About Second Language Acquisition: A Synthesis From Four Perspectives." *Review of Educational Research* 82: 5–60. doi:10.3102/0034654311433587.
Enever, J. 2012. "Current Policy Issues in Early Foreign Language Learning." *CEPS Journal* 2 (3): 9–26.
Feng, A. 2012. "Spread of English Across Greater China." *Journal of Multilingual and Multicultural Development* 33 (4): 363–377. doi:10.1080/0143632.2012.661435.
García Mayo, M., and M. Garcia Lecumberri, eds. 2003. *Age and the Acquisition of English as a Foreign Language*. Clevedon, UK: Multilingual Matters.
Han, Z., and L. Selinker. 2005. "Fossilization in L2 Learners." In *Handbook of Research in Second Language Teaching and Learning*, edited by E. Hinkel, 455–470. Mahwah, NJ: Lawrence Erlbaum Associates.
Hao, Y., L. M. Bedore, L. Sheng, and E. D. Pena. 2019. "Narrative Skills in two Languages of Mandarin-English Bilingual Children." *International Journal of Speech-Language Pathology* 21: 325–335. doi:10.1080/17549507.2018.1444092.
Hashimoto, K. 2011. "Compulsory 'Foreign Language Activities' in Japanese Primary Schools." *Current Issues in Language Planning* 12: 167–184. doi:10.1080/14664208.2011.585958.
He, D., and Q. Zhang. 2010. "Native Speaker Norms and China English: From the Perspective of Learners and Teachers in China." *TESOL Quarterly* 44: 769–789. doi:10.5054/tq.2010.235995.
Hoff, E. 2013. "Interpreting the Early Language Trajectories of Children From low-SES and Language Minority Homes: Implications for Closing Achievement Gaps." *Developmental Psychology* 49: 4–14. doi:10.1037/a0027238.
Hu, G. 2007. "The Juggernaut of Chinese-English Bilingual Education." In *Bilingual Education in China: Practices, Policies and Concepts*, edited by A. Feng, 94–126. Clevedon, UK: Multilingual Matters.
Hu, X., and S. McGeown. 2020. "Exploring the Relationship Between Foreign Language Motivation and Achievement among Primary School Students Learning English in China." *System* 89, doi:10.1016/j.system.2020.102199.
Huang, B. H. 2016. "A Synthesis of Empirical Research on the Linguistic Outcomes of Early Foreign Language Instruction." *International Journal of Multilingualism* 13: 257–273. doi:10.1080/14790718.2015.1066792.
Hwang, H., J. A. Shin, and R. J. Hartsuiker. 2018. "Late Bilinguals Share Syntax Unsparingly Between L1 and L2: Evidence From Crosslinguistically Similar and Different Constructions." *Language Learning* 68: 177–205. doi:10.1111/lang.12272.
Johnson, J., K. Shenkman, E. Newport, and D. Medin. 1996. "Indeterminacy in the Grammar of Adult Language Learners." *Journal of Memory and Language* 35: 335–352. doi:10.1006/jmla.1996.0019.
Johnstone, R. 2009. "An Early Start: What are the Key Conditions for Generalized Success?" In *Young Learner English Language Policy and Implementation: International Perspectives,* J. Enever, J. Moon, and U, Raman (eds.), 31–42. Reading, UK: Garnet Education.
Knell, E., H. Qiang, M. Pei, Y. Chi, L. S. Siegel, L. Zhao, and W. Zhao. 2007. "Early English Immersion and Literacy in Xi'an, China." *The Modern Language Journal* 91: 395–417. doi:10.1111/j.1540-4781.2007.00586.x.
Lamb, M. 2007. "The Impact of School on EFL Learning Motivation: An Indonesian Case Study." *TESOL Quarterly* 41: 757–780. doi:10.1002/j.1545-7249.2007.tb00102.x.
Larson-Hall, J. 2008. "Weighing the Benefits of Studying a Foreign Language at a Younger Starting age in a Minimal Input Situation." *Second Language Research* 24: 35–63.
Lenneberg, E. H. 1967. *Biological Foundations of Language*. New York, NY: Wiley & Sons.
Li, Q. 2014. "Differences in the Motivation of Chinese Learners of English in a Foreign and Second Language Context." *System* 42: 451–461. doi:10.1016/j.system.2014.01.011.
Li, M., J. R. Kirby, L. Cheng, L. Wade-Woolley, and H. Qiang. 2012. "Cognitive Predictors of English Reading Achievement in Chinese English-Immersion Students." *Reading Psychology* 33: 423–447. doi:10.1080/02702711.2010.528663.
Liu, Y., and A. Thompson. 2018. "Language Learning Motivation in China: An Exploration of the L2MSS and Psychological Reactance." *System* 72: 37–48. doi:10.1016/j.system.2017.09.025.
Marinova-Todd, S. H. 2002. "Know Your Grammar: What the Knowledge of Syntax and Morphology in an L2 Reveals About the Critical Period for Second/Foreign Language Acquisition." In *Age and the Acquisition of English as a Foreign Language*, edited by M. P. G. Mayo, and M. L. G. Lecumberri, 59–73. New York, NY: Multilingual Matters.
Marinova-Todd, S. H., D. B. Marshall, and C. E. Snow. 2000. "Three Misconceptions About age and L2 Learning." *TESOL Quarterly* 34: 9–34. doi:10.2307/3588095.
McDonald, J. L. 2006. "Beyond the Critical Period: Processing-Based Explanations for Poor Grammatically Judgment Performance by Late Second Language Learners." *Journal of Memory and Language* 55: 381–401. doi:10.1016/j.jml.2006.06.006.
Ministry of Education of China. 2001. 教育部《关于积极推进小学开设英语课程的指导意见》[Guidelines for promoting English language learning curriculum in elementary schools]. http://www.moe.gov.cn/srcsite/A26/s7054/200101/t20010120_166075.html.

Ministry of Education of China. 2017. 教育部《幼儿园办园行为督导评估指标与要点》[key points of supervision and inspection on the management of kindergartens]. http://www.moe.gov.cn/srcsite/A11/s6500/201705/t20170512_304460.html.

Muñoz, C. 2011. "Input and Long-Term Effects of Starting age in Foreign Language Learning." *International Review of Applied Linguistics in Language Teaching* 49: 113–133. doi:10.1515/iral.2011.006.

Muñoz, C. 2014. "Contrasting Effects of Starting age and Input on the Oral Performance of Foreign Language Learners." *Applied Linguistics* 35: 463–482. doi:10.1093/applin/amu024.

Nikolov, M., and J. M. Djigunovic. 2006. "Recent Research on age, Second Language Acquisition and Early Foreign Language Learning." *Annual Review of Applied Linguistics* 26, doi:10.1017/S0267190506000122.

Pfenninger, S. E., and D. Singleton. 2016. "Affect Trumps age: A Person- in-Context Relational View of age and Motivation in SLA." *Second Language Research* 32: 311–345. doi:10.1177/0267658315624476.

Pfenninger, S. E., and D. Singleton. 2017. *Beyond age Effects in Instructional L2 Learning: Revisiting the age Factor*. Clevedon, UK: Multilingual Matters.

Proctor, C. P., D. August, C. Snow, and C. D. Barr. 2010. "The Interdependence Continuum: A Perspective on the Nature of Spanish–English Bilingual Reading Comprehension." *Bilingual Research Journal* 33: 5–20. doi:10.1080/15235881003733209.

Rosenbaum, P. R., and D. B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70: 41–55.

Sayer, P. 2018. "Does English Really Open Doors? Social Class and English Teaching in Public Primary Schools in Mexico." . *System* 73: 58–70. doi:10.1016/j.system.2017.11.006.

Singleton, D. 2001. "Age and Second Language Acquisition." *Annual Review of Applied Linguistics* 21: 77–89. doi:10.1017/S0267190501000058.

Sirin, S. R. 2005. "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research." *Review of Educational Research* 75: 417–453. doi:10.3102/00346543075003417.

Snow, C., and Y. Kim. 2007. "Large Problem Spaces: The Challenge of Vocabulary for English Language Learners." In *Vocabulary Acquisition: Implications for Reading Comprehension*, edited by R. K. Wagner, A. E. Muse, and K. R. Tannenbaum, 123–139. New York, NY: Guilford Press.

Song, J. 2018. "English Just is not Enough!: Neoliberalism, Class, and Children's Study Abroad among Korean Families." *System* 73: 80–88. doi:10.1016/j.system.2017.10.007.

Sparks, R. L., J. Patton, and J. Luebbers. 2019. "Individual Differences in L2 Achievement Mirror Individual Differences in L1 Skills and L2 Aptitude: Crosslinguistic Transfer of L1 to L2 Skills." *Foreign Language Annals* 52: 255–283. doi:10.1111/flan.12390.

Stuart, E. A. 2010. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science* 25: 1–21. doi:10.1214/09-STS313.

Sun, H., R. Steinkrauss, J. N. Tendeiro, and K. de Bot. 2016. "Individual Differences in Very Young Children's English Acquisition in China: Internal and External Factors." *Bilingualism: Language and Cognition* 19: 550–566. doi:10.1017/S1366728915000243.

Sun, H., R. Steinkrauss, M. Wieling, and K. de Bot. 2016. "Individual Differences in Very Young Chinese Children's English Vocabulary Breadth and Semantic Depth: Internal and External Factors." *International Journal of Bilingual Education and Bilingualism* 21: 405–425. doi:10.1080/13670050.2016.1178706.

Tollefson, J. W. 1989. "The Role of Language Planning in Second Language Acquisition." In *Language Planning and English Language Teaching*, edited by C. Kennedy, 23–41. Cambridge, England: Cambridge University Press.

Unsworth, S., L. Persson, T. Prins, and K. de Bot. 2015. "An Investigation of Factors Affecting Early Foreign Language Learning in the Netherlands." *Applied Linguistics* 36: 527–548. doi:10.1093/applin/amt052.

von Stumm, S., E. Smith-Woolley, Z. Ayorech, A. McMillan, K. Rimfeld, P. S. Dale, and R. Plomin. 2019. "Predicting Educational Achievement From Genomic Measures and Socioeconomic Status." *Developmental Science*, Online Advance. doi:10.1111/desc.12925.

You, C. J., and Z. Dörnyei. 2016. "Language Learning Motivation in China: Results of a Large-Scale Stratified Survey." *Applied Linguistics* 37: 495–519. doi:d10.1093/applin/amu046.

Zhao, J., R. M. Joshi, L. Q. Dixon, and L. Huang. 2016. "Chinese EFL Teachers' Knowledge of Basic Language Constructs and Their Self-Perceived Teaching Abilities." *Annals of Dyslexia* 66: 127–146. doi:10.1007/s11881-015-0110-2.

Zhang, X., B. Hu, L. Ren, and L. Zhang. 2019. "Family socioeconomic status and Chinese children's early academic development: Examining child-level mechanisms." *Contemporary Educational Psychology* 59: 101729. doi:10.1016/j.cedpsych.2019.101792

Zheng, Y. 2012. "Exploring Long-Term Productive Vocabulary Development in an EFL Context: the Role of Motivation." *System* 40: 104–119. doi:10.1016/j.system.2012.01.007.