Early Childhood Research Quarterly 40 (2017) 163–173

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

journal homepage: www.elsevier.com/locate/ecresq

Are structural quality indicators associated with preschool process quality in China? An exploration of threshold effects

Bi Ying Hu[a], Xitao Fan[a], Yan Wu[b,*], Ning Yang[c]

[a] University of Macau, China
[b] Guangdong University of Foreign Studies, China
[c] South China Normal University, China

ARTICLE INFO

Keywords:
The Classroom Assessment Scoring System
Teacher-child interaction
Threshold analysis
Chinese kindergarten
Classroom quality

ABSTRACT

In this study, the generalized additive modeling (GAM) was used to explore possible threshold effects on multiple program structure quality indicators (class size, child-to-teacher ratio, teaching experience, teacher salary) in relation to the classroom teacher-child interaction quality indicators as measured by the Classroom Assessment Scoring System (CLASS) in a sample of Chinese kindergarten classrooms. One hundred eighty classrooms in 60 preschools were selected for the study, based on a stratified sampling procedure. The results were mixed, and revealed threshold effects on some structural indicators. The findings are discussed in the context of Chinese early childhood education practice. As the first study of its kind in a Chinese context, the findings could have some implications for early childhood education policy and practice despite some limitations of the study. Future research may consider better design and more representative sample for studying these issues.

1. Background

An encouraging finding from experimental and longitudinal investigations in the U.S. is that high-quality early childhood education (ECE) could serve as a protective factor for children in poverty (Arteaga, Humpage, Reynolds, & Temple, 2014; Campbell et al., 2014; Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; McCormick et al., 2006; Reynolds, Temple, Ou, Arteaga, & White, 2011; Schweinhart et al., 2005; Vitiello, Moas, Henderson, Greenfield, & Munis, 2012; Williford, Maier, Downer, Pianta, & Howes, 2013). Evidence from developing countries also found that quality ECE contributes to better school attendance and improves children's mental development, motor development, behavior and schooling trajectory (Baker-Henningham & López Bóo, 2010). Encouraged by the empirical evidence about the positive effects of quality ECE on children's development and growth, policy makers worldwide have sought to include high-quality ECE initiatives in their policy reforms. International organizations worldwide including OECD, UNESCO, UNICEF, and the World Bank considered access to, monitoring of, and improvement of ECE quality through quality accountability systems as top priorities for international policy platforms.

Policy makers in most countries including China, often set thresholds for structural quality measures (e.g., class size, child-to-teacher ratio) in their quality accountability systems for assessing overall quality of a program. These thresholds, such as rating score for a program's quality, promotion and funding of programs, have high-stakes implications in practice. Under such an approach, we assume meaningful threshold values exist on a structural variable (e.g., years of teaching, class size), above or below which, its relation with ECE program quality either strengthens or weakens. Researchers are thus encouraged to explore such possibilities for kindergarten structural indicators in relation to ECE quality (Zaslow et al., 2010). The approach of setting thresholds on a structural indicator, however, has only seen limited empirical research; in most cases, thresholds were set with little or no empirical evidence (Le, Schaack, & Setodji, 2015).

1.1. ECE situation in the Chinese context

In China, the goal of providing high-quality ECE for all Chinese young children is clear in recent national initiatives and reform agendas. Most children of three to six year olds attend early childhood programs called kindergarten, which is a full-day program offering three levels of classes: K1 (3–4 years old), K2 (4–5 years old), and K3 (5–6 years old). K3 is equivalent to kindergarten class in the U.S. In 2010, the Chinese government promulgated the *Compendium for China's Mid- and Long-term Education Reform and Development (2010–2020)* (Central People's Government of the People's Republic of China, 2010), which declared that by 2020, 75% of children should receive three

* Corresponding author at: Department of Applied Psychology, Guangdong University of Foreign Studies, Guangzhou, 510420, China.
E-mail address: wuyan209@gdufs.edu.cn (Y. Wu).

http://dx.doi.org/10.1016/j.ecresq.2017.03.006
Received 12 May 2016; Received in revised form 23 February 2017; Accepted 29 March 2017
Available online 25 April 2017

years of ECE. Ministry of Education of People's Republic of China (2012) then issued a national document stressing the importance of improving the quality of ECE. Many reform efforts have been made, such as free admission, subsidized tuition and vouchers in rural areas, aiming to increase socioeconomically disadvantaged children's access to ECE services (Hu, Zhou, Li, & Roberts, 2014). As a result, China has made significant progress in increasing the number of children attending kindergartens. The quality of these programs, however, remains a serious concern (Hu et al., 2014; Rao, Sun, Zhou, & Zhang, 2012; Wong, Luo, Zhang, & Rozelle, 2013). Calls have been made that future research efforts for improvement of ECE quality should provide empirical evidence on appropriate thresholds for some critical structural indicators (e.g., class size, child-to-teacher ratio, teaching experience, expenditure per teacher on salaries) that are highly likely to be associated with ECE quality (e.g., Hu, Mak, Neitzel, Li, & Fan, 2016). Therefore, in this study, we explored threshold effects for some common structural indicators in the Chinese context that could impact the quality of ECE classroom teaching.

1.2. Structural vs. process quality of ECE

To better support the ECE quality initiatives, it is necessary to develop a common understanding among international professionals about what constitutes ECE quality, the components of such quality, and effective ways of measuring the quality. Quality is not a static concept; rather, it varies across time and contexts (Hu, 2015; Moss & Dahlberg, 2008; Tobin, 2005). Consequently, some scholars do not support a universal standard for quality (Tobin, 2005). However, for the sake of scholarly communication and policy implementation, most scholars agree that quality should be divided into two main domains: structural and process. Structural quality refers to some indicators that could be regulated or controlled by the relevant government agency and/or the program such as class size and child-to-teacher ratio, and some program-level variables such as expenditures on teachers' salaries, training, and program facilities. These indicators are often determined in specific cultural and social contexts and represent distal indicators of ECE quality that may have an indirect effect on children development and growth (Bryant, Zaslow, & Burchinal, 2010; Phillips, Mekos, Scarr, McCartney, & Abbott-Shim, 2000). Process quality, in contrast, refers to the indicators that may have a direct effect on program-level and classroom-level teaching/learning quality (Burchinal, Howes, & Kontos, 2002). Most definitions of process quality refer to specific interactive activities that children experience in the classroom, most commonly teacher-child interactions, peer interactions and language stimulation by the teacher.

One frequently examined process quality indicator is teacher-child interaction quality, measured by the widely known and used Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008). The CLASS measures three aspects of teacher-child interaction: emotional support, classroom organization, and instructional support. Emotional support, which fosters children's willingness to participate in classroom activities, assesses the positive and warm relationship between teacher and child (Hamre & Pianta, 2005). Class organization is mainly concerned with teachers' use of proactive behavior management strategies to keep students engaged in learning and play, which helps children to develop better self-regulation skills (Rimm-Kaufman, La Paro, Downer, & Pianta, 2005). Instructional support refers to how teachers effectively promote children's higher-order thinking skills and language abilities in classroom instructional activities. It should be noted that both emotional support and classroom organization underpin high quality instructional support (Pianta & Hamre, 2009).

Both structural and process quality indicators have been included in many different systems in the U.S. (Schaack, Tarrant, & Boller, 2012; Zellman & Perlman, 2008). The general belief is that structural variables may not only be correlated with process quality, but more importantly, they may also be precursors to process quality (Cryer, Tietze, Burchinal, Leal, & Palacios, 1999; Hu, Mak et al., 2016; Phillips et al., 2000; Phillipsen, Burchinal, Howes, & Cryer, 1997; Pianta et al., 2005; Slot, Leseman, Verhagenm, & Mulder, 2015). Studies on relations between structural and process quality have been conducted across many cultures (Cryer et al., 1999; Hu, Mak et al., 2016; Slot et al., 2015). For example, in a cross-cultural comparison study (Cryer et al., 1999) that used the Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms, Clifford, & Cryer, 1998) for assessing process quality, child-to-teacher ratio was reported to be related to higher process quality in the U.S., but not in Portugal and Spain. In Netherlands, De Kruif and colleagues reported that better child-to-teacher ratio was related to higher process quality, but they did not find a similar relation between teachers' education and process quality (as cited in Slot et al., 2015). In China, using the Chinese adapted version of the same environmental rating scale, researchers also found that child-to-teacher ratio was a strong predictor for both "Provisions for Learning" and "Teaching and Interaction" components of program quality (Hu, Mak et al., 2016).

Under the framework of teacher-child interaction quality, process quality was reflected by the indicators in CLASS. By using these CLASS indicators, researchers showed that many structural variables (e.g., class size, child-to-teacher ratio, teacher's degree, years of teaching, expenditures) were associated with process quality. For example, teachers who scored higher on process quality indicators tend to have more years of experiences and/or higher education degrees (Phillips et al., 2000; Phillipsen et al., 1997; Pianta et al., 2005; Slot et al., 2015). Hu, Fan, LoCasale-Crouch, Chen, and Yang (2016), using a latent profile analysis approach, found teachers' experiences, education, and some other professional status variables were significantly distinctive across four profiles of teacher-child interaction patterns. On the other hand, Slot et al. (2015) failed to find a relation between either class size or child-to-teacher ratio with emotional and educational process quality. Similarly, Pianta et al. (2005) did not find better child-to-teacher ratio made any difference for process quality. Finally, it is noteworthy that several studies reported that better teacher salary was related to higher level of classroom process quality (Phillips et al., 2000; Phillipsen et al., 1997).

In summary, the findings from previous studies on the associations between structural quality and process quality have not been consistent (Early et al., 2006, 2007). We surmise that the relations between ECE structural quality indicators and process quality indicators might not be linear. Recently, some researchers have used a different approach in examining this issue and are exploring whether there are baseline and ceiling thresholds for kindergarten structural indicators (such as years of teaching and class size) in relation to ECE quality (Zaslow et al., 2010).

1.3. Research on threshold effects

For the relations between process quality indicators (e.g., teacher-child interaction quality as represented by emotional support, classroom organization and instructional support) and children's social and academic outcomes, there has been some active research that explored whether a critical threshold value existed for an ECE process quality indicator, above or below which the relation between the quality indicator and children's social and academic outcomes was different (Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Burchinal, Vernon-Feagans, Vitiello, & Greenberg, 2014; Finch, Johnson, & Phillips, 2015; Leyva et al., 2015; Weiland, Ulvestad, Sachs, & Yoshikawa, 2013). These studies provided some evidence that there may exist critical threshold values on the process quality indicators in relation to the children's outcomes.

For the relations between ECE program's structural quality indicators and the process quality indicators, however, there has been very limited research on possible critical threshold/ceiling values. Le et al.

(2015) attempted to identify cut-points (i.e., threshold values) for structural quality indicators (child-to-teacher ratio, number of ECE credits completed by teachers, and teaching experience) in relation to global program quality as represented by ECERS-R total score. Meaningful threshold/ceiling values were identified for structural indicators such as class size, years of teaching, and teachers' training in ECE (i.e., number of credit hours taken for ECE training). The authors recommended that in future studies, efforts should be made to identify threshold and ceiling points of the structural program features in relation to teacher-child interaction quality as measured by the CLASS.

1.4. Study context and hypothesized thresholds for structural indicators

As described earlier, the three-year kindergarten is the most prevalent form of ECE in China. Chinese kindergartens have some unique structural features. One notable feature is the fairly large class size and high child-to-teacher ratio (Hu & Li, 2012; Hu & Szente, 2009). The *Chinese Kindergarten Bylaw* and the *National Guideline for Kindergarten Curriculum* recommend class sizes of 20–25 for K1 (3–4 year olds), 25–30 for K2 (4–5 year olds), and 30–35 for K3 (5–6 year olds). In addition, in each classroom, two instructional staff and one teaching assistant or three instructional staff is recommended to ensure basic quality. To be specific, the highest child-to-teacher ratio for classrooms should be 8:1 for K1, 10:1 for K2, and 11:1 for K3. Moreover, the recommended child to total number of instructional staff ratio in full-day program should be from 5:1 to 7:1 (Ministry of Education of People's Republic of China, 2013). Chinese ECE professionals have been fighting to reduce the class size and increase the number of staff in kindergarten classrooms, however, in reality class size can range from 20 to as high as 60 or 70 in rural private programs, and child-to-teacher ratio is much larger because of the lack of additional staff and assistants. It is critical that researchers provide empirical evidence for national guideline on class sizes and child-to-teacher ratio so that local governments can be supported with evidence when strengthening their structural quality monitoring systems. Therefore, we hypothesize that, in Chinese kindergartens, a ceiling threshold exists for class size such that the relation between class size and teacher-child interaction quality (i.e., the smaller the class size, the better the teacher-child interaction quality) only exists below the threshold value, but not beyond it. Similarly, we hypothesize a ceiling threshold for the relation between child-to-teacher ratio and teacher-child interaction quality (i.e., the smaller the ratio, the better the interaction quality), and the relation would weaken or disappear beyond the value. Identifying such a threshold may have important policy and practical implications, as such empirical findings could help local government agencies and program directors make policy and programmatic decisions on such issues (e.g., recommended class size, child-to-teacher ratio).

Another issue related to Chinese kindergartens is the wide disparity across kindergartens in terms of kindergartens' financial resources and expenditure, due to multiple societal and other factors, such as public vs. private kindergartens, urban vs. rural regions, different financial conditions across localities, etc. Such disparity across kindergartens can be reflected in teachers' salary, expenditure for teacher training, and expenditure for school facilities and their maintenance, etc., among which teacher salary is often a critical one, as this could have substantial impact on the recruitment and retention of effective and experienced teachers. Because of this consideration, we hypothesize that the relation between teacher's salary and process quality indicators will be minimal until a threshold on teachers' salary is reached, as kindergarten programs are more likely to recruit and retain effective teachers when the salary is competitive in the market.

To date, in the social and cultural context of China, there has been no attempt to identify threshold values for structural quality indicators relative to the ECE quality of classroom experiences. As the threshold values on such structural indicators relevant for classroom quality could play a very meaningful and important role in determining the levels of resources and incentives awarded to ECE programs, empirical findings in this area are becoming increasingly important for both policy and practical considerations. Identification of these thresholds can help policymakers and ECE program directors identify the most cost-effective way to invest and allocate their limited resources and develop effective policies and approaches for improving the quality of ECE classroom experiences for Chinese children. Based on these considerations, the following research questions were examined in the current study.

1.5. Research question 1

Can classroom interaction quality be predicted by structural quality indicators such as class size, child-to-teacher ratio, teaching experience, and teacher salary?

1.6. Research question 2

Are there threshold effects on these structural quality indicators such that their relations with classroom process quality indicators change at these critical threshold values?

2. Methods

2.1. Participants

This study took place in the southern Chinese province of Guangdong from May to June 2014. Guangdong is a province in southern China with a large population of slightly over 100 million, and it has a wide range of economic development levels typical of China's overall social and economic context. Sixty kindergartens were selected randomly to represent the diverse range of ECE programs from areas of varying economic development levels. More specifically, we targeted three areas: an economically advanced urban metropolitan area, an economically mid-level county area with a mix of urban and rural kindergartens, and an economically under-developed rural area.

With the support of local school boards, we first held briefing meetings in each of these three areas about the goals and processes of the research, and the participating kindergartens were invited to participate in the research project. From each of these three areas that varied in terms of economic development, 20 kindergartens were randomly selected for participation, making the total number of participating kindergartens to be 60. As discussed elsewhere, in the current Chinese social context, there tends to be very obvious disparities between urban and rural kindergartens (Hu et al., 2014). Because of this consideration, the kindergartens from the three sampled areas were coded for their location of either being urban or rural, and this location indicator was used as a controlled variable in later analyses.

Within each selected kindergarten, we randomly selected one teacher from a K1 classroom, one teacher from a K2 classroom, and one teacher from a K3 classroom. This selection resulted in a total of 180 teacher participants from 180 kindergarten classrooms.

2.2. Procedure

2.2.1. Teachers' consent and survey collection

Upon receiving consent from the directors of the 60 kindergartens, we asked each director to provide a list of teachers who were willing to participate in the study. Based on the list of teachers within each school who indicated their interest in participation, we randomly selected one teacher to represent each grade level (K1, K2, and K3, as described above) for each kindergarten. As we only sent consent forms and basic demographics questionnaire to the teachers who had already indicated their willingness to participate, all of the 180 selected teachers provided written consent and basic demographics questionnaire, and none of them withdrew during the study.

We sent the *Kindergarten Survey* to the directors of the 60 kindergartens. The directors provided information on various structural features of their kindergartens, ranging from the total amount of public funding, fee schedules, student enrolment, teacher make-up, and kindergarten expenditure. A graduate research assistant collected the completed surveys from the directors and checked with each director to make sure they did not have difficulty understanding the survey questions. Also, the same graduate student collected the basic demographics questionnaire from the classroom teachers.

2.2.2. Training for CLASS use

Eighteen graduate-level ECE or psychology majors attended a four-day CLASS training program offered by *TeachStone* (www.teachstone.com), the company that provides CLASS training, certification and professional development. The training consisted of an introduction to the CLASS theoretical framework and its three domains, followed by coding practice using videotaped samples of teacher-child classroom interactions. After the training, the coders practiced in real Chinese kindergarten classrooms for one month before they took an online test offered by *TeachStone*. All passed the test and became certified CLASS raters.

2.2.3. Observational video data collection and coding

From May to June 2014, the graduate students videotaped teacher-child interactions in the 180 participating classrooms during the first three hours of a typical day. The graduate students extracted five 20-min video clips (i.e., observation cycles) from each 3-h video. The five cycles represented a typical day in Chinese kindergartens, including whole-group teaching, free play, outdoor activities, and routine/care activities such as meals and snacks. Eight pairs of certified CLASS raters, two per group, rated the quality of teacher-child interactions using the double-coding method. The scores for the five cycles for each kindergarten were averaged to create a single score. Each pair of raters rated between 18 and 24 classrooms. To ensure inter-rater reliability, the raters coded five 20-min cycles independently. By using the intra-class correlation coefficient (*ICC*) approach illustrated in Fan and Sun (2014), we obtained the average inter-rater reliability estimates for all the dimensions of the CLASS. The *ICC*s for the ten dimensions of the CLASS ranged from 0.82 to 0.92, suggesting a high degree of inter-rater consistency on these dimensions of the CLASS in the Chinese preschool sample (Hu, Fan, Gu et al., 2016).

2.3. Instruments

2.3.1. Classroom measures

The CLASS (Pianta et al., 2008) is an assessment tool that originated in the U.S., and has recently been validated and applied in international contexts, including Portugal (Cadima, Leal, & Burchinal, 2010), Finland (Pakarinen et al., 2010), Germany (Suchodoletz, Fäschea, Gunzenhausera, & Hamre, 2014), Chile (Leyva et al., 2015), and China (Hu, Fan, Gu et al., 2016). The CLASS defines and assesses the quality of teacher-child interactions in classrooms through three interaction domains: emotional support (including positive climate, negative climate, teacher sensitivity, and regard for student perspectives), classroom organization (including behavior management, productivity, and instructional learning formats), and instructional support (including concept development, quality of feedback, and language modeling).

In the current study, we combined emotional support and classroom organization, as these two subscales are statistically highly correlated ($r = 0.77$, $p < 0.001$). The correlations of instructional support with emotional support and classroom organization were 0.53 and 0.39 respectively. As discussed elsewhere (Hamre, Hatfield, Pianta, & Jamil, 2014), there are also theoretical reasons to combine them, as they both measure motivational aspects of the learning environment that support cognitive and academic interactions. This combined subscale of emotional support and classroom organization was named ESCO. The internal consistency reliability estimates for ESCO and instructional support subscales in this study sample were 0.90 and 0.85, respectively, and the correlation between ESCO and instructional support was 0.49.

The CLASS uses a 7-point rating scale, with scores of 1 or 2 indicating low quality, 3, 4 and 5 indicating medium-level quality, and 6 or 7 indicating high quality. A recent study in China provided evidence to support the validity of the CLASS use for measuring the quality of kindergartens in Chinese cultural contexts (Hu, Fan, Gu et al., 2016).

2.3.2. Survey of kindergarten program features

Each participating kindergarten director completed a brief questionnaire on the basic information of the kindergarten, including geographic location, financial source, the funding agency, the quality rating of the kindergarten by the relevant government agency, and the amount of annual public funding received, if any, the total number of staff and children enrolled, and total expenditure on teachers' salary, etc.

2.3.3. Survey of teacher demographics information

Each classroom teacher completed a questionnaire for information such as age, teaching experiences, years of teaching, and professional background (i.e., major, highest educational level, first degree, frequency and duration of in-service training), class size, number of staff per classroom, and monthly income of the teacher, etc.

2.4. Analytical approach

The first step in identifying thresholds is to identify possible cut-off point(s) on a relevant structure quality indicator (e.g., class size) at which the relation between this indicator and a process quality variable (e.g., instructional support of CLASS) changes (strengthening, weakening, or disappearing). Generalized additive modeling (GAM; Hastie & Tibshirani, 1990) was used to identify possible thresholds on four structural indicators (teaching experience, class size, child-to-teacher ratio, teacher salary) in relation to two classroom process quality indicators (CLASS subscales: ESCO and Instructional support). So our analyses would involve four independent variables (teaching experience, class size, child-to-teacher ratio, teacher salary) and two dependent variables (ESCO and instructional support) in their respective GAM analysis. Then, the approach of piecewise regression was used to validate the possible thresholds.

2.4.1. Identification of thresholds using GAM

GAM offers a nonlinear method for identifying likely thresholds by estimating the relations between an independent variable (i.e., a kindergarten structural indicator variable, such as class size) and a dependent variable (i.e., a classroom process quality indicator variable, such as instructional support) without making any assumptions about whether the relation is linear or non-linear (e.g., quadratic or polynomial). Likely thresholds are identified by graphically identifying the region(s) where there appears to be systematic change in the relation between the two variables, and both baseline and ceiling thresholds are possible.

A GAM is an additive regression model of the form:

$$Y=\beta_0 + \sum_{j=1}^{p} g_j(X_j) + \varepsilon$$

where Y represents the outcome; β_0 is the intercept; g_j represents smooth functions (Hastie & Tibshirani, 1990), which can be plotted to illustrate the marginal relation between the predictor and the response; and ε is random error. Categorical variables can be easily accommodated within the model by the usual regression approach. The GAM model with categorical variables is

$$Y=\beta_0 + \sum_{j=1}^{p} g_j(X_j) + M_\gamma + \varepsilon$$

where M is the design matrix for the variable that will not be modeled additively, and γ are the regression parameters. More details about GAM are available elsewhere (Hastie & Tibshirani, 1990).

GAM provides graphic representation of the non-linear regression trend between a kindergarten structure quality indicator (e.g., class size) and a covariate-adjusted classroom interaction quality outcome variable (e.g., instructional support). More specifically, the x-axis represents the kindergarten structure quality indicator (e.g., class size), and the y-axis represents this indicator's non-linear smoothed contribution to a classroom interaction quality outcome (e.g., instructional support), which can be viewed as a standardized regression coefficient, or scaled effect size, of the relation between the structure quality indicator and the classroom interaction quality outcome.

2.4.2. Validation of the selected thresholds

GAM does not provide the exact cut-off points for thresholds; instead, it provides visual guidance on the score ranges within which thresholds are likely to exist. Therefore, it is critical to confirm the validity of GAM-derived thresholds. The approach of piecewise regression, also known as the spline regression approach (Marsh & Cormier, 2002), can be used to test whether the regression slopes vary across the different regions defined by the thresholds (Marsh & Cormier, 2002), thus providing evidence for the validity of the GAM-derived thresholds (Le et al., 2015; Setodji, Le, & Schaack, 2013).

Furthermore, for each outcome variable, the results from the simple linear regression model (assuming one constant slope) and the piecewise regression model (assuming varying slopes as defined by the thresholds) could be compared statistically. e.g., statistical test, adjusted R^2s from two models, and the Akaike information criterion (AIC; Akaike, 1974) are available as evidence for the viability of the thresholds. For technical details of such piecewise regression model, readers may consult Marsh and Cormier (2002).

In our analysis, the modified Huber-White standard error was used to account for the clustering of classrooms under kindergartens (Freedman, 2006). To enhance interpretability, unstandardized regression coefficients are presented. Effect size (Cohen's d) was computed to represent the gain (in standard deviation unit) on the classroom quality outcome measure corresponding to an increase of one standard deviation in the predictor (Burchinal et al., 2010; Finch et al., 2015; Weiland et al., 2013).

$$d=\frac{B_{Predictor} SD_{Predictor}}{SD_{Outcome}}$$

Our analysis included 177 out of the initial 180 teachers, due to a very small number of missing responses. No differences were found between the teachers with and without missing values for the CLASS variables, indicating that the very little missing data would be very unlikely to cause estimation bias.

3. Results

3.1. Sample characteristics and variable relations

Table 1 presents the descriptive information about the sample (teachers and classrooms) and the CLASS measures (ESCO and instructional support). The average score for CLASS instructional support is very low at 2.13 for this sample of teachers/classrooms, suggesting that there is considerable room for improvement in classroom teacher-child interaction quality. This sample of teachers reported lower salary (41,595 RMB, equivalent to $6,040.59) than the norm, as the average salary in Guangdong province is 75,656 RMB (equivalent to $10,987.06). Table 2 presents the correlations among the study variables.

Table 1
Descriptive Statistics for Teacher Demographic and Classroom-Related Characteristics.

	Sub-group N (%)	Mean	SD	Range
Teacher certification				
Yes	126 (71%)			
No	51 (29%)			
Teacher highest education				
Below high school	9 (5%)			
High school	52 (29%)			
University	116 (66%)			
Location				
Urban	81 (46%)			
Rural	96 (54%)			
Teacher age		31.49	7.28	18–58
Teaching experience (years)		10.31	6.95	1–29
Class size		32.49	7.49	9–58
Child-to-teacher ratio		12.98	4.15	5–29
Teacher salary (RMB)		41,600	27,900	10,800–180,020
ESCO		4.91	0.71	2.77–6.53
Instructional support		2.13	0.61	1.00–3.93

Note. ESCO: combined emotional support and classroom organization. Total $N = 177$. A certified teacher is a teacher who has earned the required teaching credentials from an authoritative source, typically a local/regional governmental education agency or bureau.

As discussed previously, we will first identify the possible thresholds by using GAM, and then, the piecewise regression approach was used to validate the possible thresholds.

3.1.1. Teaching experience

Fig. 1 provides a graphical representation of the relations between teaching experience and the two CLASS subscales: ESCO and instructional support. Fig. 1 suggests that teaching experience may have a ceiling threshold at about 15 years for ESCO (left plot), while it shows an almost flat trend with no obvious threshold for the outcome of instructional support (right plot). In our sample, the majority (75.7%; $n = 134$) of the teachers fell below the cut-off point value of 15 years of teaching experience. Table 3 compares the results of the linear regression model with those of the piecewise regression model on each outcome (ESCO, instructional support) for which GAM analysis suggested the existence of thresholds. For teaching experience, the results of the piecewise regression model suggest that it is a statistically significant predictor of the ESCO score, and the slope change at the point of 15 is confirmed (below 15: $\beta = 0.033$, $p < .05$, effect size = 0.305; above 15: $\beta = -0.071$, $p < .05$, effect size = 0.287). However, the linear regression model failed to show statistically significant relation between teaching experience and ESCO. Statistical test comparing the linear regression model with the piecewise regression model confirmed that the piecewise regression model had better fit for the data ($F = 5.819$, $p = 0.017$). Both the adjusted R^2 (higher value indicating better fit) and AIC (Akaike, 1974; lower value indicating better fit) also indicated the better fit of the piecewise regression model. In short, for the structure quality indicator variable of teaching experience, the threshold value at 15 as suggested by GAM analysis was empirically supported.

3.1.2. Class size

Fig. 2 indicates that class size may have a ceiling threshold of 35 students per class for ESCO outcome (left plot), while it shows an upward trend with no obvious threshold effect for instructional support outcome (right plot). In our sample, 33% ($n = 59$) of classes have more than 35 students. The piecewise regression confirms that a threshold exists at 35 students per class, with ESCO deteriorating statistically ($\beta = -0.047$, $p < 0.05$; effect size = 0.191) in classes with more than 35 students. The piecewise regression model was shown to have better

Table 2
Pearson Correlations among the Study Variables.

		1	2	3	4	5	6	7	8
1	ESCO	–							
2	Instructional Support	0.486**	–						
3	Teacher highest education	0.386**	0.232**	–					
4	Location (1 = urban)	0.355**	0.246**	0.282**	–				
5	Teacher certification (1 = yes)	0.401**	0.222**	0.427**	0.359**	–			
6	Teaching experience (years)	0.295**	0.221**	0.349**	0.303**	0.386**	–		
7	Class size	0.009	0.068	0.090	−0.113	0.100	0.069	–	
8	Child-to-teacher ratio	−0.251**	−0.176*	−0.203**	−0.386**	0.119	−0.125	0.698**	–
9	Teacher salary	0.422**	0.325**	0.320**	0.540**	0.318**	0.440**	−0.021	−0.248**

Note: ESCO: combined emotional support and classroom organization.
* $p < 0.05$.
** $p < 0.01$.

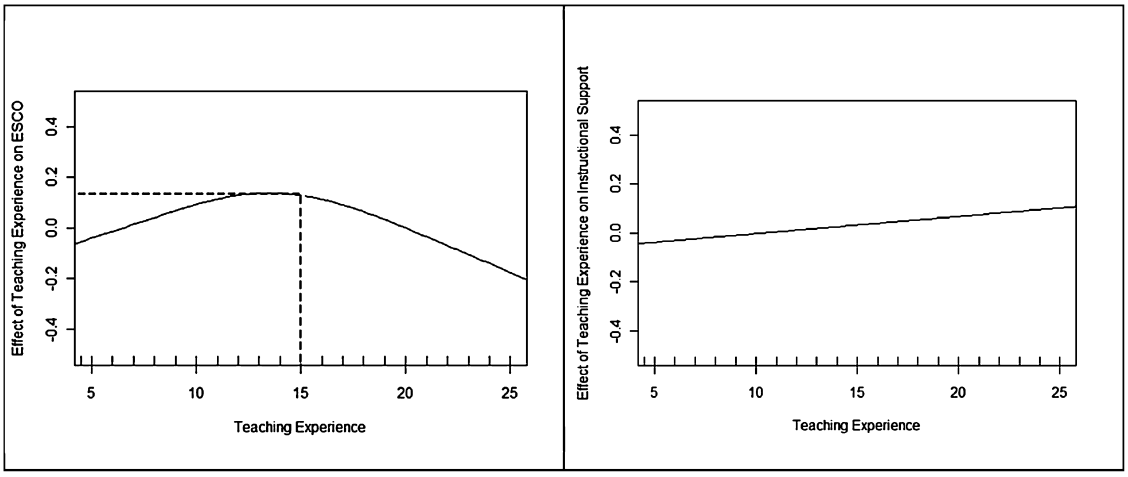

Fig. 1. GAM Plots: Using teaching experience to predict ESCO and instructional support.

fit than the linear model ($F = 4.355$, $p = 0.038$). Both adjusted R^2 and the AIC value indicated better fit of the piecewise regression model than the linear regression model.

3.1.3. Child-to-teacher ratio

Fig. 3 suggests that child-to-teacher ratio may have a ceiling threshold of 13 for ESCO outcome variable (left plot), while it shows a general downward trend with no obvious threshold for instructional support outcome variable. In our study sample, 41% ($n = 73$) of the teachers took care of more than 13 children. Although the GAM analysis suggested the threshold value of 13 for the child-to-teacher ratio variable in relation to ESCO outcome, this is not confirmed by the piecewise regression. Neither the linear regression model nor the piecewise regression model showed any statistically significant regres-

Table 3
Validation of Thresholds: Comparing Linear and Piecewise Regression Analyses for Predicting CLASS Scores (β and SE).

Structural quality Indicators	Thresholds	Classroom Interaction Quality Indicators				Adjusted R^2		AIC	
		ESCO		Instructional Support					
		M1	M2	M1	M2	M1	M2	M1	M2
Teaching experience	≤15	0.008(0.008)	0.033(0.013)*			0.242	0.263	338.4	333.4
	>15		−0.071(0.027)*						
Class size	≤35	0.001(0.006)	0.012(0.011)			0.237	0.257	339.6	337.1
	>35		−0.047(0.023)*						
Child-to-teacher ratio	≤13	0.020(0.012)	−0.020(0.025)			0.248	0.244	336.9	338.9
	>13		0.001(0.040)						
Teacher salary	≤7			0.051(.019)**	−0.022(0.049)	0.113	0.183	315.4	302.8
	7 < X ≤ 12.5				0.374(0.093)**				
	>12.5				−0.623(0.122)**				

Note: M1: Linear regression model; M2: Piecewise regression model.
All models controlled for teacher highest level of education, teacher certification, and school location (urban/rural) of the kindergarten.
ESCO: emotional support and classroom organization combined; AIC: Akaike information criterion.
For Teacher salary: 7 = 70,000 RMB; 12.5 = 125,000 RMB.
* $p < 0.05$.
** $p < 0.01$.

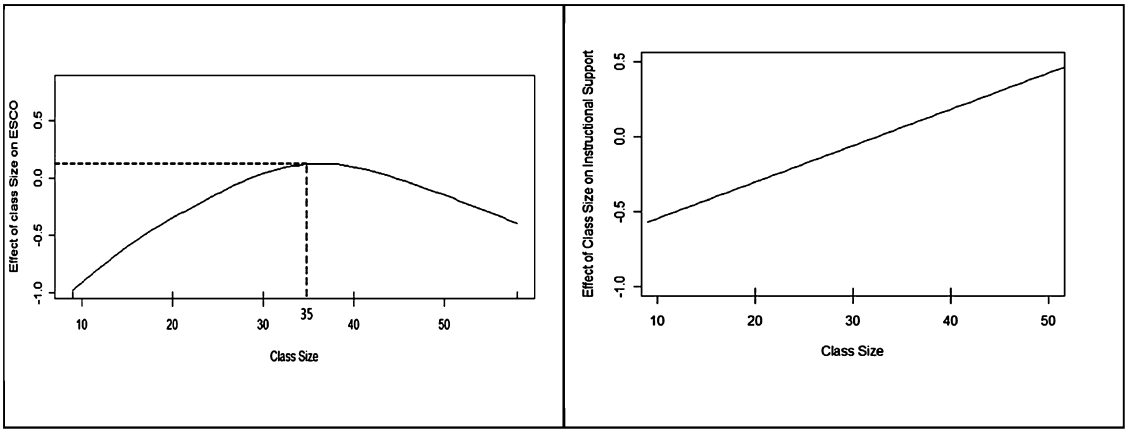

Fig. 2. GAM Plots: Using class size to predict ESCO and instructional support.

sion slope(s) when the child-to-teacher ratio variable was used for predicting ESCO.

3.1.4. Teacher salary

In Fig. 4, for ESCO variable, teacher salary has a general upward trend without obvious threshold (left plot). For instructional support (right plot), the plot shows a positive slope starting at approximately 70,000 RMB (equivalent to $10,165.67), and the upward trend changes direction at about 125,000 RMB (equivalent to $18,152.99). This suggests that relation between teacher salary and instructional support is positive in the range of 70,000–125,000 RMB, but no longer so beyond 125,000 RMB. In our study sample, most teachers, 88.1% (n = 156), had annual salary below 70,000 RMB, 10.2% (n = 18) between 70,000 RMB and 125,000 RMB, and 1.7% (n = 3) had more than 125,000 RMB. Linear regression analysis in Table 3 indicated that teacher salary was a statistically significant predictor for instructional support (β = 0.051, $p < 0.01$, effect size = 0.232). The piecewise regression model confirmed the thresholds at 70,000 and 125,000 respectively. More specifically, within the salary range of 70,000 and 125,000 RMB per year, the relation between teacher salary and classroom instructional support score was positive and strong (β = 0.374, $p < 0.01$, effect size = 0.940). Beyond 125,000, the relation turned negative ($\beta = -0.623$, $p < 0.01$, effect size = 0.725). Below 70,000, the relation is not statistically significant. The piecewise regression model had better fit than the linear model (F = 8.342, $p < 0.001$). Both the adjusted R^2 and AIC values indicated much better fit of the piecewise regression model than the linear regression model.

In summary, the piecewise regression models provided support for the thresholds derived from the GAM plots, with the exception for child-to-teacher ratio variable. Tentatively, we may consider that teaching experience has a ceiling threshold of 15 for ESCO, class size has a ceiling threshold of 35 for ESCO, and teacher salary has a baseline threshold of 70,000 RMB and a ceiling threshold of 125,000 RMB for instructional support.

4. Discussion

This study was exploratory in nature, and it examined whether class size, child-to-teacher ratio, teaching experience, and teacher salary might have threshold effects on two classroom interaction quality indicators (ESCO and instructional support of the CLASS). To our knowledge, this is the first study to examine such relations of these two types of variables, and it is also the first of its kind in the Chinese social/cultural context. Our findings suggest possible threshold effects on class size, teaching experience, and teacher salary, but not on child-to-teacher ratio. Also, for ESCO, thresholds were identified for teaching experience (ceiling threshold at 15 years) and class size (ceiling threshold at 35 students), but not for child-to-teacher ratio and expenditure on teacher salary. For IS, thresholds were identified for class size (a ceiling threshold at 35 students) and expenditure on teacher salary (a minimum salary threshold of 70,000 RMB), but not for child-to-teacher ratio and teaching experiences. Findings on different patterns of threshold effect for ESCO and instructional support is thought-provoking. Overall, such findings echoed previous literature in terms of either inconsistency on the associations between some of these structural variables (e.g., class size, teaching experience, and

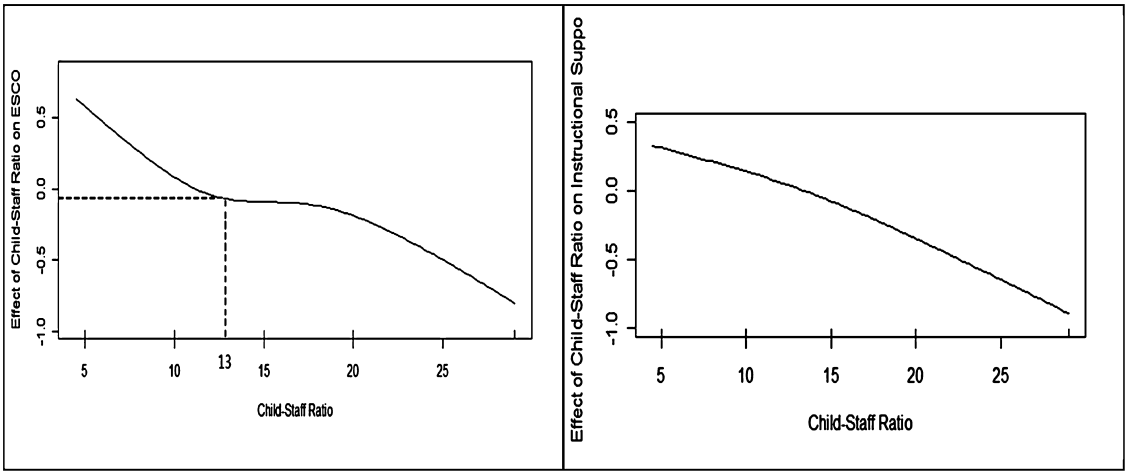

Fig. 3. GAM Plots: Using child-to-teacher ratio to predict ESCO and instructional support.

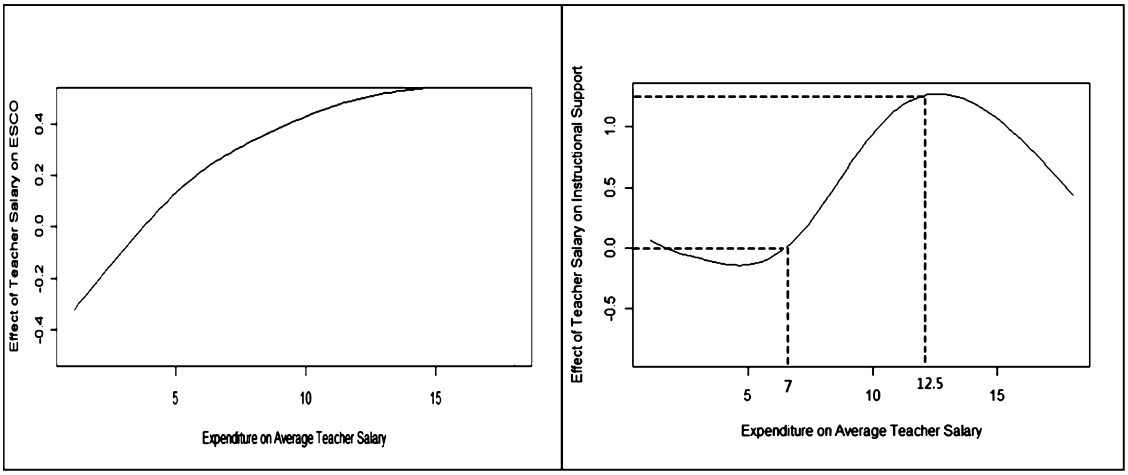

Fig. 4. GAM Plots: Using teacher salary (7.0 = 70,000 RMB; 12.5 = 125,000 RMB) to predict ESCO and instructional support.

child-to-teacher ratio) and the process quality (Early et al., 2006, 2007; Pianta et al., 2005; Slot et al., 2015), or the possibility of the existence non-linear relation.

Threshold effects of two structural variables are of most significance. First, we found class size being the only variable that showed a consistent threshold effect for both ESCO and instructional support, which confirms that, regardless of the type of interactions (emotional/social, behavioral, and academic), a manageable class size is the key to ensure minimum level of quality in Chinese preschool context. Second, regarding the threshold effect found on teacher salary that only existed for instructional support quality, but not for ESCO, the implications are multi-fold. On the one hand, it confirms the previous literature about the linear relations between expenditure on teacher salary and ECE process quality, suggesting that increasing teachers' salary is important for attracting and retaining ECE teachers who are likely to provide higher quality academic interaction (Hu, Fan, LoCasale-Crouch et al., 2016; Hu, Mak et al., 2016; Phillips et al., 2000; Phillipsen et al., 1997). On the other hand, it clarifies that, while better teacher pay is important for the overall process quality, it is particularly important for instructional support quality, which constitutes the most important aspect of teacher-child interaction quality linked to children's cognitive and academic gains (Burchinal et al., 2008).

Furthermore, the threshold effect on teacher salary helps to explain recent consistent findings on lower level of instructional support quality across cultures (Hu, Fan, Gu et al., 2016; Leyva et al., 2015; Pakarinen et al., 2010; Suchodoletz et al., 2014), which reinforces some scholars' concern that preschool teachers in China are underpaid (Hu, Mak et al., 2016; Li, Hu, Pan, Qin, & Fan, 2014). Therefore, the most important finding of the study is that schools that offer low teacher salary may not be able to attract and retain teachers who will provide higher-quality instructional support in the classroom. These findings on threshold effects have policy and programmatic implications.

For teaching experience, contrary to our expectation for a baseline threshold (e.g., 5–10 years) but not for a ceiling threshold, the findings revealed no baseline threshold, but a ceiling threshold at 15 years, above which the positive relation between teaching experience and ESCO disappeared. This was inconsistent with the finding of Le et al. (2015), who identified 15 years as the baseline threshold above which there is a positive relation between teaching experience and the ECERS-R. Although the ECERS-R measures global quality with the emphasis on structural aspect of ECE quality, whereas CLASS only assesses teacher-child interaction quality (Hu, Fan, Gu et al., 2016). Our finding here supports the very general belief in the Chinese cultural context that a teacher needs to accumulate a considerable number of years (e.g., 15) of teaching experience in order to become an "expert" teacher in classroom teaching. Yet, the finding does conditionally (i.e., only up to about 15 years) support the notion that the longer a teacher works in the classroom, the better he/she becomes at maintaining positive relationships with children and having better behavior management in classroom.

We suspect that the ceiling effect at 15 years could be the result of different teacher cohorts. Early childhood education has witnessed considerable development and change together with the fast economic development of China in the past two decades. Older teachers could have been trained before the *Guideline for Kindergarten Curriculum* (GKC) in 2001 (Ministry of Education of the People's Republic of China, 2001). The GKC represented a paradigm shift to bring child-centered teaching and learning concepts into teacher education and kindergarten practice in China (Hu, 2015), in contrast to the traditional dominance of teacher-centeredness in early childhood education practice. Therefore, we speculate that teachers with more than 15 years experience might have more adult-centered teaching, thus receiving lower scores on ESCO.

Neither a threshold effect nor a linear relation was shown between teaching experience and instructional support. This implies that more years of teaching has not helped teachers master instructional support strategies, which differs from what was previously found about the association between process quality and a moderate amount of experience (Phillipsen et al., 1997). However, the different result could be due to our seperate examination of ESCO and instructional support, whereas in the previous study, the total scores of process quality measurements were used. Nevertheless, this finding has implications for pre-service and in-service professional development, and it points to the necessity that training and practice of effective instructional support skills should be emphasized in teachers' professional development in China.

For the variable of class size, in relation to ESCO, we identified a threshold ceiling of 35 students beyond which classroom interaction quality (ESCO) started to deteriorate. This finding appeared to provide empirical support for the maximum class size of 35 recommended in both the *Chinese Kindergarten Bylaw* and the *National Guideline for Kindergarten Curriculum.* But for class size below 35, the larger class size was associated with better classroom interaction quality. This finding could be the result of confounding of class size with grade level in our sample data. As is typical in the Chinese social and cultural context, in the kindergarten classroom sample in the study, class size of 20–25 was typical for K1 classrooms, 25–30 for K2 classrooms, and 30–35 for K3 classrooms. Our data also showed a pattern that K3 classrooms had better scores in ESCO and instructional support than K2 classrooms, which in turn had better scores than K1 classrooms. Therefore, the positive relation between class size and classroom interaction quality could be mainly due to this grade "effect". For this grade "effect", we

suspect that older kids in Chinese kindergarten classrooms tended to be more responsive to teachers, which made it easier for the teachers to use more open-ended questions and other more interactive learning activities, thus resulting in more and better classroom interactions.

With respect to the child-to-teacher ratio, we found neither baseline nor ceiling threshold effects on the relations with the classroom interaction quality indicators of ESCO and instructional support. However, a linear relation between child-to-teacher ratio and instructional support suggested that a lower child-to-teacher ratio was associated with higher classroom interaction quality. Overall, the findings here are consistent with previous finding (Phillips et al., 2000) and the policy recommendation insisted by Chinese scholars that there should be two teachers and one assistant in each kindergarten classroom. In reality, the child-to-teacher ratio in Chinese kindergarten classrooms varies greatly. In our study sample, the classroom child-to-teacher ratio ranged from 6.33 to 31.00 ($M = 12.94$, $SD = 4.26$). Also, it is worth pointing out that in some classrooms, teachers have either morning or afternoon shifts which might cause confusing for calculating child-to-teacher ratio as they are not with children at the same time yet they are accounted as the number of staff. Future studies are hence needed to take into consideration the reality with regard to the child-to-teacher ratio in Chinese kindergartens when examining its threshold effects on classroom interaction quality.

For teacher salary, there was a general positive relation with both classroom interaction quality indicators (ESCO and instructional support). This is in line with our hypothesis that the schools offering higher salary could be more likely to recruit more competent teachers, and these teachers provide a better classroom interaction environment that would be conducive to children's learning. No threshold values on teacher salary was identified for ESCO. But for instructional support, a minimum salary threshold of 70,000 RMB (equivalent to $10,165.67) was identified, and so was a ceiling threshold of 125,000 RMB (equivalent to $18,152.99), beyond which the positive relation between teacher salary and instructional support disappeared. This finding of teacher salary's ceiling threshold was unexpected and appeared to be odd. But the validity of this ceiling threshold could be suspicious, and should be viewed with caution because the very limited number of teachers in the sample above this salary level ($n = 3$) could have made this estimation very unreliable and unstable. Taken together, the generally strong relation between teacher salary and classroom interaction quality (ESCO and instructional support) suggests that teacher salary could be an important factor in attracting more talented teaching professionals for kindergartens.

It is noteworthy that studies have shown that the instructional support domain of the CLASS is significantly related to children's learning and development outcomes both in the U.S. (Curby et al., 2009; Hamre, Hatfield, Pianta, & Jamil, 2014; Hamre & Pianta, 2005; Pianta & Hamre, 2009), and in the international contexts such as Finland (Pakarinen et al., 2014), Chile (Leyva et al., 2015), and Ecuador (Araujo, Carneiro, Cruz-Aguayo, & Schady, 2014). In our sample of kindergarten classrooms, the classroom interaction quality in the domain of instructional support was alarmingly low, with the highest score being 3.93 out of 7 and the average being 2.13. This suggests a glaring weakness or absence in the professional knowledge and expertise of many Chinese kindergarten teachers, and it points to the need for better professional training and development to overcome this weakness. The professionals in early childhood education need to explore effective ways for helping Chinese ECE teachers improve the quality of instructional support in the classroom.

4.1. Limitations and directions for future research

Our study has some limitations. First, as this was not an experimental design study, the relations, if any, between the structural indicators (i.e., kindergarten class size, child-to-teacher ratio, teaching experience, teacher salary) and classroom interaction quality indicators (ESCO and instructional support) should not be the basis for inferring any causal relation from the former to the latter. Given that teachers were not randomly assigned to the kindergarten students, the observed relationships may reflect existing differences among teachers.

The second limitation is the degree of generalizability of the findings in broader cultural and social contexts in China. Although our sample of 180 teachers and 60 preschools is considered fairly large and representative of the diverse ECE programs across the economic strata within Guangdong province, the degree of the sample's representativeness nationally remains uncertain. With this consideration, caution is highly warranted in generalizing the findings to the national context, especially considering the glaring differences in economic development across different regions of China. Finally, we used the classroom interaction sub-domains of the CLASS as measures of program quality in the Chinese kindergarten classrooms. Future studies may examine the relations between structural quality and process quality by using some other measures of program quality (e.g., some relevant sub-domains of the Chinese Early Childhood Environment Rating Scale; Li et al., 2014).

4.2. Implications for ECE policy and practice

As the first of its kind, these findings from this study may offer meaningful insights relevant to ECE policy and practice. The most important recommendation is that schools with reasonably competitive teacher salary (e.g., ≥70,000 RMB for the concerned localities in Guangdong province) could be more attractive for recruiting highly effective teachers, and the professional quality of the teachers will be reflected in better classroom interaction in kindergarten classrooms. Given the limited financial resources in many school districts, it is crucial to allocate financial resources to raising teacher salary above the baseline threshold (e.g., 70,000 RMB as the minimum in this region). In terms of government policy, this would ensure that the teaching profession in general could be financially attractive for talented professionals, and schools would be able to hire better teachers.

The unexpected ceiling threshold on teaching experience (i.e., 15 years of teaching experience) suggests a need for professional development training for teachers who entered the profession more than 15 years ago. More specifically, particular attention should be given to teachers who received their teacher education before the implementation of the GKC in 2001, before the switch to (and emphasis on) child-centeredness in kindergarten curricula and practice as reflected in the GKC (Ministry of Education of the People's Republic of China, 2001). These teachers might need targeted training to help them adopt child-centered practices to ensure that their teaching behaviors and practice contribute to positive classroom interaction quality of emotional and behavioral support. It is important for policy makers and school administrators to realize that one's teaching behavior and practice (e.g., teacher-centered classroom practice) would not change simply because of a new policy document. Instead, purposeful and effective professional development/training is needed to facilitate such teaching behavior change in the classrooms.

Finally, the findings showed that the classroom teacher-child interaction quality is generally low in the sampled Chinese kindergarten classrooms, especially how teachers effectively promote children's higher-order thinking skills and language abilities in classroom instructional activities. More and effective professional development training appears to be urgently needed for pre-service and in-service teachers to enhance their understanding of the concept and importance of classroom teacher-child interaction quality, to explore what teaching behaviors and practice constitute high-quality teacher-child interaction quality, and to enhance such teacher-child interaction quality in daily teaching activities and practice. In short, there is the need for ECE policy makers and professionals to take a fresh look at the importance of classroom teacher-child interaction quality (especially instructional support quality) and develop effective plans for both pre-service and in-

service training to help current and future teachers acquire the necessary knowledge and skills to enhance teaching quality and effectiveness in kindergarten classrooms.

Acknowledgements

The study and the preparation of the manuscript were supported by the following projects: (1) “A Longitudinal Study on the Effects of Preschool Program Quality on Children's Learning and Development Outcomes.” (University of Macau Multi-Year Research Grant; MYRG20l5-00156-FED); and (2) “Analyses of the Interactive Effects in Multilevel Structural Equation Models and Their Applications.” (The National Natural Science Foundation of China; 31400909). The authors of this paper deeply appreciate the support.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2014). *A helping hand? Teacher quality and learning outcomes in kindergarten*. Washington, DC: Banco Interamericano de Desarrollo [Inédito].

Arteaga, I., Humpage, S., Reynolds, A. J., & Temple, J. A. (2014). One year of preschool or two: Is it important for adult outcomes? *Economics of Education Review*, *40*, 221–237.

Baker-Henningham, H., & López Bóo, F. (2010). *Early childhood stimulation interventions in developing countries: A comprehensive literature review (IZA Discussions Paper No. 5282)*. [Retrieved from: ECONSTOR website: https://www.econstor.eu/bitstream/10419/46166/1/663155967.pdf].

Bryant, D. M., Zaslow, M., & Burchinal, M. (2010). Issues in measuring program quality. In P. W. Wesley, & V. Buysse (Eds.), *The quest for quality: Promising innovations for early childhood programs* (pp. 47–68). Baltimore, MD: Brookes Pub. Co.

Burchinal, M., Howes, C., & Kontos, S. (2002). Structural predictors of child care quality in child care homes. *Early Childhood Research Quarterly*, *17*(1), 87–105.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science*, *12*(3), 140–153.

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, *25*(2), 166–176.

Burchinal, M., Vernon-Feagans, L., Vitiello, V., Greenberg, M., & Family Life Project Key Investigators (2014). Thresholds in the association between child care quality and child outcomes in rural preschool children. *Early Childhood Research Quarterly*, *29*(1), 41–51.

Cadima, J., Leal, T., & Burchinal, M. (2010). The quality of teacher-student interactions: Associations with first graders’ academic and behavioral outcomes. *Journal of School Psychology*, *48*(6), 457–482.

Campbell, F., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, *6*(1), 42–57.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014). Early childhood investments substantially boost adult health. *Science*, *343*(6178), 1478–1485.

Central People’s Government of the People’s Republic of China. (2010). *国家中长期教育改革和发展规划纲要*(2010–2020*年*)) [Compendium for China’s Mid- and Long-term Education Reform and Development (2010–2020)]. Retrieved from: http://www.gov.cn/jrzg/2010-07/29/content_1667143.htm.

Cryer, D., Tietze, W., Burchinal, M., Leal, T., & Palacios, J. (1999). Predicting process quality from structural quality in preschool programs: A cross-county comparison. *Early Childhood Research Quarterly*, *14*(3), 339–361.

Curby, T. W., LoCasale-Crouch, J., Konold, T. R., Pianta, R. C., Howes, C., Burchinal, M., & Barbarin, O. (2009). The relations of observed pre-K classroom quality profiles to children’s achievement and social competence. *Early Education and Development*, *20*(2), 346–372.

Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., & Barbarin, O. (2006). Are teachers’ education, major, and credentials related to classroom quality and children’s academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, *21*(2), 174–195.

Early, D. M., Maxwell, K., Burchinal, M., Alva, D., Bender, R., Bryant, D., et al. (2007). Teachers’ education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development*, *78*(2), 558–580.

Fan, X., & Sun, S. (2014). Generalizability theory as a unifying framework of measurement reliability in adolescent research. *Journal of Early Adolescence*, *34*, 38–65.

Finch, J. E., Johnson, A. D., & Phillips, D. A. (2015). Is sensitive caregiving in child care associated with children’s effortful control skills? An exploration of linear and threshold effects. *Early Childhood Research Quarterly*, *31*, 125–134.

Freedman, D. A. (2006). On the so-called Huber sandwich estimator and robust standard errors? *The American Statistician*, *60*(4), 299–302.

Hamre, B., & Pianta, R. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, *76*(5), 949–967.

Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children’s development. *Child Development*, *85*(3), 1257–1274.

Harms, T., Clifford, R., & Cryer, D. (1998). *Early childhood environment rating scale-revised*. New York: Teachers College Press.

Hastie, T., & Tibshirani, R. (1990). *Generalized additive models*. New York, NY: Chapman & Hall.

Hu, B., & Li, K. (2012). The quality rating system of Chinese preschool education: Prospects and challenges? *Childhood Education*, *88*(1), 14–22.

Hu, B., & Szente, J. (2009). Exploring the quality of early childhood education in China: Implications for early childhood teacher education? *Journal of Early Childhood Teacher Education*, *30*(3), 247–262.

Hu, B., Zhou, Y., Li, K., & Roberts, S. K. (2014). Examining program quality disparities between urban and rural kindergartens in China: Evidence from Zhejiang. *Journal of Research in Childhood Education*, *28*(4), 461–483.

Hu, B. (2015). Comparing cultural differences in two quality measures in chinese kindergartens: The early childhood environment rating scale-Revised and the kindergarten quality rating system. *Compare: A Journal of Comparative and International Education*, *45*(1), 94–117.

Hu, B., Fan, X., Gu, C., & Yang, N. (2016). Applicability of the Classroom Assessment Scoring System in Chinese preschools based on psychometric evidence. *Early Education and Development*, *27*, 714–734.

Hu, B., Fan, X., LoCasale-Crouch, J., Chen, L., & Yang, N. (2016). Profiles of teacher-child interactions in Chinese kindergarten classrooms and the associated teacher and program features. *Early Childhood Research Quarterly*, *37*, 58–68.

Hu, B., Mak, M. C. K., Neitzel, J., Li, K., & Fan, X. (2016). Predictors of Chinese early childhood program quality: Implications for policies. *Children and Youth Services Review*, *70*, 152–162.

Le, V. N., Schaack, D. D., & Setodji, C. M. (2015). Identifying baseline and ceiling thresholds within the qualistar early learning quality rating and improvement system. *Early Childhood Research Quarterly*, *30*, 215–226.

Leyva, D., Weiland, C., Barata, M. C., Yoshikawa, H., Snow, C., Treviño, E., & Rolla, A. (2015). Teacher-child interactions in Chile and their associations with prekindergarten outcomes. *Child Development*, *86*(3), 781–799.

Li, K., Hu, B., Pan, Y., Qin, J., & Fan, X. (2014). Chinese early childhood environment rating scale (trial) (CECERS): A validity study. *Early Childhood Research Quarterly*, *29*(3), 268–282.

Marsh, L. C., & Cormier, D. (2002). *Spline regression models*. London, UK: Sage.

McCormick, M. C., Brooks-Gunn, J., Buka, S. L., Goldman, J., Yu, J., Salganik, M., et al. (2006). Early intervention in low birth weight premature infants: Results at 18 years of age for the Infant Health and Development Program. *Pediatrics*, *117*(3), 771–780.

Ministry of Education of People’s Republic of China. (2012). 3–6*岁儿童学习与发展指南* [Guideline for Learning and Development of 3–6 Aged Children]. Retrieved from http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s3327/201210/xxgk_143254.html.

Ministry of Education of People’s Republic of China. (2013). *幼儿园教职工配备标准*(*暂行*) [Standard for Staffing of Kindergarten (Temporary)]. Retrieved from: http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s7027/201301/147148.html.

Ministry of Education of the People’s Republic of China (2001). *幼儿园教育指导纲要*(*试行*) [The Guideline for Kindergarten Curriculum (Trial Version)]. Retrieved from: http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s7054/201403/xxgk_166067.html.

Moss, P., & Dahlberg, G. (2008). Beyond quality in early childhood education and care −Languages of evaluation? *New Zealand Journal of Teachers’ Work*, *5*(1), 3–12.

Pakarinen, E., Lerkkanen, M. K., Poikkeus, A. M., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., & Nurmi, J. E. (2010). A validation of the classroom assessment scoring system in finnish kindergartens? *Early Education and Development*, *21*(1), 95–124.

Pakarinen, E., Aunola, K., Kiuru, N., Lerkkanen, M. K., Poikkeus, A. M., Siekkinen, M., & Nurmi, J. E. (2014). The cross-lagged associations between classroom interactions and children’s achievement behaviors. *Contemporary Educational Psychology*, *39*(3), 248–261.

Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott-Shim, M. (2000). Within and beyond the classroom door: Assessing quality in child care centers? *Early Childhood Research Quarterly*, *15*(4), 475–496.

Phillipsen, L. C., Burchinal, M. R., Howes, C., & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly*, *12*, 281–303.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*, 109–119.

Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144–159.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system*. Baltimore: Paul H. Brookes.

Rao, N., Sun, J., Zhou, J., & Zhang, L. (2012). Early achievement in rural China: The role of preschool experience. *Early Childhood Research Quarterly*, *27*(1), 66–76.

Reynolds, A. J., Temple, J. A., Ou, S. R., Arteaga, I. A., & White, B. A. (2011). School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science*, *333*(6040), 360–364.

Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children’s behavior in

kindergarten classrooms. *The Elementary School Journal*, *105*(4), 377–394.
Schaack, D., Tarrant, T., Boller, K., & Tout, K. (2012). Quality rating and improvement systems: Alternative approaches to understanding their impact on the early learning system. In S. L. Kagan, & K. Kaurez (Eds.), *Early childhood systems: Transforming early learning* (pp. 71–86). New York, NY: Teacher's College Press.
Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The High/Scope perry preschool study through age 40*. Ypsilanti, MI: High/Scope Educational Research Foundation.
Setodji, C. M., Le, V. N., & Schaack, D. (2013). Using generalized additive modeling to empirically identify thresholds within the ITERS in relation to toddlers' cognitive development. *Developmental Psychology*, *49*(4), 632–645.
Slot, P. L., Leseman, P. P. M., Verhagenm, J., & Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early childhood education and care settings. *Early Childhood Research Quarterly*, *33*, 64–76.
Suchodoletz, A., Fäschea, A., Gunzenhausera, C., & Hamre, B. K. (2014). A typical morning in preschool: Observations of teacher-child interactions in German preschools. *Early Childhood Research Quarterly*, *29*, 509–519.
Tobin, J. (2005). Quality in early childhood education: An anthropologist's perspective. *Early Education and Development*, *16*(4), 421–434.
Vitiello, V. E., Moas, O., Henderson, H. A., Greenfield, D. B., & Munis, P. M. (2012). Goodness of fit between children and classrooms: Effects of child temperament and preschool classroom quality on achievement trajectories. *Early Education and Development*, *23*(3), 302–322.
Weiland, C., Ulvestad, K., Sachs, J., & Yoshikawa, H. (2013). Associations between classroom quality and children's vocabulary and executive function skills in an urban public prekindergarten program? *Early Childhood Research Quarterly*, *28*(2), 199–209.
Williford, A. P., Maier, M. F., Downer, J. T., Pianta, R. C., & Howes, C. (2013). Understanding how children's engagement and teachers' interactions combine to predict school readiness. *Journal of Applied Developmental Psychology*, *34*(6), 299–309.
Wong, H. L., Luo, R., Zhang, L., & Rozelle, S. (2013). The impact of vouchers on preschool attendance and elementary school readiness: A randomized controlled trial in rural China. *Economics of Education Review*, *35*, 53–65.
Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Tarullo, L., & Burchinal, M. (2010). *Quality dosage, thresholds, and features in early childhood settings: a review of the literature (OPRE 2011-5)*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
Zellman, G. L., & Perlman, M. (2008). *Child care quality rating improvement systems in five pioneer states: Implementation issues and lessons learned*. Santa Monica, CA: RAND Corporation.