Early Childhood Research Quarterly 60 (2022) 226–236

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

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

Self-regulation development among young Spanish-English dual language learners

Margaret O'Brien Caughy[a,*], Dawn Y. Brinkley[a], Daniel Pacheco[b], Raul Rojas[b], Alicia Miao[c], Mariah M. Contreras[d], Margaret Tresch Owen[b], M. Ann Easterbrooks[d], Megan McClelland[c]

[a] *Department of Human Development and Family Science, University of Georgia, 405 Sanford Drive, Athens, GA, 30602, USA*
[b] *University of Texas at Dallas, Richardson, TX, USA*
[c] *Oregon State University, Corvallis, OR, USA*
[d] *Tufts University, Medford, MA, USA*

ARTICLE INFO

Article history:
Received 24 May 2021
Revised 7 February 2022
Accepted 14 February 2022

Keywords:
Self-regulation
Trajectory analysis
Spanish-English dual-language learners

ABSTRACT

Despite strong evidence self-regulation skills are critical for school readiness, there remains a dearth of longitudinal studies that describe developmental trajectories of self-regulation, particularly among low-resource and underrepresented populations such as Spanish-English dual-language learners. The present study examined individual differences in trajectories of self-regulation among 459 Spanish-English dual-language learners who were Hispanic from 4 different samples and 3 geographic locations in the United States. Self-regulation was assessed in all samples using repeated administration of the Head-Toes-Knees-Shoulders (HTKS) task from early childhood through early elementary school. Results of latent growth curve analyses revealed that growth was best represented by quadratic trajectories. Latent class growth analyses captured significant individual differences in self-regulation trajectories. One group of children (41%) started with higher HTKS scores and displayed rapid early growth in performance. A similar percentage of children (41%) displayed intermediate growth in self-regulation, starting with lower HTKS scores but displaying rapid growth commencing arrange 4.5 years. Finally, about 18% of the sample did not display growth in HTKS performance until after entry to elementary school, around age 6 years. Girls were half as likely as boys to be in this later developing group. Likewise, children from families at the upper end of the socioeconomic distribution in this low-income sample were significantly less likely to be in the later developing group relative to children from families with lower SES. Study findings indicate the importance of monitoring growth rates in self-regulation as a means of identifying children at risk for entering school without the requisite self-regulation skills.

Self-Regulation development among young Spanish-English dual language learners

Hispanic children represent the fastest growing and the largest minority population of school-age children in the United States, of which many are Spanish-English dual language learners, a population for which there is a well-documented academic achievement gap (Francis et al., 2019; Rojas, Hiebert, Gusewski, & Francis, 2019). The term "dual language learners" (DLLs) describes children without a specific profile of relative proficiency in two languages (e.g., whether a child's use of both languages is balanced versus if a child uses 1 language predominantly and the other minimally). For DLL children in the United States, the 2 languages spoken are overwhelmingly Spanish and English, and these children have a higher likelihood of living in environments that increase risk for lower academic achievement. Specifically, Spanish-English DLLs in the United States (U.S.) who are Hispanic have a higher prevalence of living in poverty (25%) than non-Hispanic, White children (11%) and are also more likely to have at least 1 parent who did not complete high school (Federal Interagency Forum on Child & Family Statistics, 2019).

As a result of greater exposure to environmental risk factors, Spanish-English DLLs in the U.S. are at risk for early academic struggles (Jackson, Schatschneider, & Leacox, 2014; McFarland et al., 2018). Further, Spanish-English DLLs in the U.S. are at risk for early academic struggles due to other contextual factors such as being more likely to attend large, highly segregated, under-resourced urban public schools (Fry, 2008). Self-regulation skills have been identified as an important component of school readiness, with children who enter school with better self-regulation skills performing better academically

* Corresponding author.
E-mail address: Margaret.Caughy@uga.edu (M.O. Caughy).

https://doi.org/10.1016/j.ecresq.2022.02.004
0885-2006/

(McClelland, Cameron Ponitz, Messersmith, & Tominey, 2010; Robson, Allen, & Howard, 2020). However, research on the normative development of self-regulation development among Spanish-English DLLs is virtually non-existent.

Importance of self-regulation in early development

Self-regulation is defined as a complex, multidimensional set of skills including aspects of emotion, cognition, and behavior. Self-regulation development is a major milestone of early childhood, and individual differences are a strong predictor of school readiness and early academic achievement (McClelland, Cameron Ponitz, Messersmith, & Tominey, 2010). A recent meta-analysis by Robson, Allen, & Howard, 2020 found positive associations between better self-regulation and academic achievement. Effect sizes for the longitudinal associations between self-regulation in preschool and early academic achievement 0.24 for vocabulary/literacy and 0.31 for mathematics.

The current study focuses on the aspects of self-regulation most relevant for early learning and development, which are based on 3 executive function (EF) processes: working memory, attentional or cognitive flexibility, and inhibitory control (Cameron Ponitz, McClelland, Matthews, & Morrison, 2009). Working memory refers to the maintenance and manipulation of information (Gathercole, 2008) and is important for children to be able to remember and follow directions. Attentional or cognitive flexibility is the ability to maintain focus and adapt to changing goals (Rothbart & Posner, 2005) and is important for children to be able to persist on difficult tasks or switch attention when needed. Finally, inhibitory control is the ability to stop a dominant response in favor of a more adaptive one (Blair, 2003) and helps young children learn self-control, for example, not impulsively hitting a child to get a toy they want. Although each aspect of EF contributes to learning outcomes, evidence indicates it is the integration of the 3 that is particularly important for success in learning and development (McClelland & Cameron, 2012). This also suggests that in young children, measures that capture these aspects of EF are useful for understanding the development of self-regulation, especially in the context of early learning (Caughy, Owen, Mills, & Hurst, 2013; McClelland et al., 2014). Thus, research has focused on improving the measurement of EF and self-regulation in young children (McClelland et al., 2014). In the present study, we use a measure of self-regulation, the Head-Toes-Knees-Shoulders (HTKS) task, that captures behavioral aspects of EF in Spanish-English DLLs from 4 sites located in 3 different geographic locations in the U.S.

Normative development of self-regulation during early childhood and its determinants

Children's self-regulation emerges gradually over the first few years of life and is influenced by child, family, and sociodemographic factors. Although individual differences exist, the development of self-regulation typically moves from external to internal regulation in young children (Kopp, 1982). For example, children often learn to self-regulate through a process by which parents and caregivers initially provide external regulation (e.g., "Don't touch that, it's hot!") followed by a gradual internalization of regulation (e.g., a child avoids hot things on their own). A child's temperament is one source of individual differences in self-regulation. For example, children who rate higher on measures of temperamental effortful control, which traditionally measure aspects of attentional focusing, inhibitory control, perceptual sensitivity and low-intensity pleasure (Rothbart & Bates, 2006), are more likely to demonstrate stronger self-regulation. In addition, gender plays a role with research girls displaying small but significantly better self-regulation relative to boys (Kia-Keating, Nylund-Gibson, Kia-Keating, Schock, & Grimm, 2018; Matthews, Cameron Ponitz, & Morrison, 2009), although research has also been mixed (e.g., Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016; Wanless et al., 2016).

Sociodemographic factors are consistent and substantive predictors of children's self-regulation with children from low socioeconomic status (SES) households routinely performing well below their peers from middle class and upper middle-class households (McClelland et al., 2014). Children from low SES households are more likely to experience household disorganization and inconsistent parenting, which are associated with poorer self-regulation development. For example, Vernon-Feagans, Willoughby, & Garrett-Peters, 2016 collected self-regulation data on a large sample of low-income White and Black children living in rural North Carolina and Pennsylvania and found that higher levels of household chaos were associated with poorer self-regulation development, and this association was mediated by negative impacts of household chaos on maternal sensitivity and responsiveness.

Individual differences in growth in self-regulation across early childhood

Due to a relative lack of longitudinal data, there are limited empirical data documenting developmental trajectories of self-regulation during early childhood and how trajectories of growth may differ between children. One exception to this is the analysis reported by Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 who used latent growth curve analysis to examine growth trajectories in a measure of behavioral self-regulation in a multi-study sample of nearly 1400 children primarily from middle-SES and mostly English-speaking communities Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016. included data from 3 different studies with multiple measurements of self-regulation between ages 4 and 7.5 years. Results indicated that self-regulation skills accelerated during preschool into early elementary school, with some children evidencing a slowing of growth during the latter part of the observation period. Furthermore, Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 identified 3 different groups of self-regulation growth: children who developed skills early, those who developed them later, and those who developed them intermediate to the other two groups. Predictors of these different groups were examined, and findings indicated that higher levels of maternal education were associated with increased probability of being in the early or intermediately developing self-regulation groups.

Another study from Taiwan also examined self-regulation growth between ages 3.5 and 6 years and found two groups of growth in young children (Wanless et al., 2016). One group was an "increasing regulators" group that showed steady development between 4 and 5.5 years of age. The other group was a "steady-then-increasing" group that showed slower initial growth which increased by the end of the study period. Group membership was related to age but not child gender or maternal education level, which highlights possible cultural differences in the development of children's self-regulation.

Self-regulation development among dual language learners

In general, children learn to self-regulate through dynamic, reciprocal interactions with parents and caregivers that continue over time. These interactions contribute to significant plasticity and variability in the development of these skills (McClelland, Geldhof, Cameron, & Wanless, 2015). It remains unclear, however, if children who are Spanish-English DLLs experience similar development in self-regulation compared to monolingual Hispanic or non-Hispanic English-speaking children. One possibility is that Spanish-English DLLs experience faster growth in self-regulation. A body of research suggesting that bilingualism may have an advantage in self-regulation development specifically related to enhanced executive function would lend support to this possibility (Barac, Bialystok, Castro, & Sanchez, 2014). Specifically, it has been theorized that bilingualism enhances set shifting, one of the core executive functions, which in turn enhances self-regulation development.

However, other researchers have failed to replicate a bilingual advantage using large national samples and meta-analytic approaches (Dick et al., 2019). Critical reviews of the research supporting a bilingual advantage in executive function development have identified a number of methodological weaknesses of this research such as publication bias favoring empirical findings supporting a bilingual advantage and failure to appropriately control for confounding variables (Paap, Johnson, & Sawi, 2015). Furthermore, a significant limitation of this research is that it is based solely on cross-sectional analyses. An unanswered question is whether being bilingual is associated with faster growth in self-regulation.

The current study

The present study extends the limited research on longitudinal trajectories of self-regulation development to examine growth of these skills in a large, multi-site sample of Spanish-English DLLs employing an integrated data analysis (Curran & Hussong, 2009). As noted, Spanish-English DLLs are uniquely at risk for early academic failure, and as such, a better understanding of developmental trajectories of self-regulation in this population would facilitate the promotion of interventions to support school readiness. We address 3 research questions. Our first research question addresses the overall shape (form) of the growth trajectory of self-regulation in the combined sample of Spanish-English DLL children who were Hispanic. Consistent with the work of Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016, we expect Spanish-English DLLs will demonstrate curvilinear, quadratic growth in self-regulation skills from early childhood to early elementary school, with accelerated growth followed by an attenuation of growth rates in early elementary school.

Our second research question addresses whether there are meaningful subgroups of Spanish-English DLL children who demonstrate unique patterns of self-regulation growth during this developmental period. As noted, previous research has identified either 2 (Wanless et al., 2016) or 3 (Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016) patterns of growth in self-regulation across early childhood and the transition to formal schooling. Comparing the profiles in self-regulation between these 2 studies indicate that the two found by Wanless et al. (2016) were similar to the late and intermediate growth groups found by Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016. The third group reported by Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 was an early growth group that showed higher initial levels of self-regulation, increased skills overtime and slowing growth at an earlier point that the other 2 groups. In the present study, we predict we will find two groups that correspond to the intermediate and late developing groups identified by both Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 and Wanless et al. (2016). However, we do not make an a priori prediction regarding whether we will identify the third, early development group similar to that reported by Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016.

Finally, our third research question examines whether there are individual differences in self-regulation growth related to child gender, family socioeconomic status, and language use. Relative to boys, we predict girls will be more likely to display higher levels of self-regulation (Kia-Keating, Nylund-Gibson, Kia-Keating, Schock, & Grimm, 2018). Likewise, we predict children from higher SES homes will display higher levels of self-regulation compared to children from low SES homes (Evans & Rosenbaum, 2008). Finally, in this sample of DLLs, we hypothesize children who speak both English and Spanish will display faster growth of self-regulation relative to children who speak only English or only Spanish.

Method

Participants

The study sample included a total of 459 participants comprised of 4 different samples from 3 different geographic locations. At each site, participants were administered the *Head-Toes-Knees-Shoulders* (HTKS) task (Cameron Ponitz et al., 2008; McCabe, Rebello-Britto, Hernandez, & Brooks-Gunn, 2004; McClelland et al., 2014) between the ages of 3 and 7.5 years. In the following sections, we describe the samples for each individual site. All study protocols were reviewed and approved by the institutional review boards of the respective institutions: University of Texas Health Science Center at Houston (HSC-SPH-08–0016); University of Texas at Dallas (17–86 and 13–11), Tufts University (#0,705,005), and Oregon State University (IRB 4766).

Dallas bilingual language and self-regulation sample

Participants from this sample were 75 Spanish-English DLL children (42 girls, 52%) enrolled in a laboratory school serving a predominantly low-income Hispanic population. The children were academically instructed in English starting in preschool, although Spanish support was available in preschool when requested by the children. The majority of the children in the study sample were receiving free or reduced price-lunch (n = 69, 92%). The children were part of a larger investigation of language development and self-regulation skills among Spanish-English speaking DLLs between 2013 and 2015 (Gusewski & Rojas, 2017) and had a mean age of 3.79 years (*SD* = 0.46 years) at onset (fall of prekindergarten-3). The majority of parents (81%) reported that the language use at home ranged from primarily Spanish to using Spanish and English relatively equally. Study children were assessed twice a year (fall and spring) with the HTKS task starting from prekindergarten-3 to Grade 1 (4 academic years) for a total of 8 consecutive semesters. For the current study, data from the HTKS task administered in English were used as the children were academically instructed in English.

Dallas preschool readiness project (DPREP)

Participants included 156 Spanish-English DLL children (47% girls) drawn from a larger longitudinal study of self-regulation development and school readiness among low-income African American and Hispanic children recruited between November 2009 and February 2011 through a variety of community venues serving low-income families and by word of mouth (author reference). Children were first assessed with the HTKS task during the second wave of data collection when they were 3½ years of age, and children were re-assessed twice more in kindergarten and first grade. Of the 224 Hispanic children enrolled in the study, 5 was excluded due to subsequent diagnosis of significant disability, and 40 were excluded because they were monolingual English-speaking at the time of enrollment. Of the remaining 179 children, 156 (87%) completed at least one HTKS assessment between 37 and 87 months of age. Assessments were administered in the child's preferred language.

Tufts University, Tufts Interdisciplinary Evaluation Research (TIER)

Participants from the TIER sample were recruited from communities across Massachusetts between February 2008 and October 2009 and consisted of 164 children (39% girls) who were followed across 6 waves starting prenatally through second grade as part of a randomized control trial (RCT) evaluation of the Massachusetts Healthy Families home visiting program for young parents (Tufts Interdisciplinary Evaluation Research, 2017).

All Hispanic children from the larger evaluation study were considered for the current integrated study as well as any non-Hispanic children whose primary caregiver used Spanish as their preferred language in the data collection. Data from Waves 4 and 5, collected during a phone interview and research home visit between 2013 and 2015, were used for analyses. The analytic sample represents 23% of the total baseline evaluation sample. At Wave 4, children were on average 5 years of age (M = 4.99, SD = 0.47) and entering or enrolled in kindergarten. Although most children were administered the protocol in English, 16% of children used Spanish or a mix of Spanish and English at the time of interview. At Wave 5, children were on average 6 years of age (M = 6.12, SD = 0.54), and, at this point, only two children used Spanish or a mix of Spanish and English for the interview.

Oregon state university sample (OSU)

The OSU sample included 64 Spanish-English DLLs (41% girls) enrolled in Head Start preschool programs enrolled between 2011 and 2014 as part of a larger study on the development of self-regulation (McClelland et al., 2014). At the first wave of the study, children were 4.8 years old on average (SD = 0.31, range = 4.3 – 5.3 years old). Children were assessed on a battery of self-regulation and academic measures in the fall and spring of both preschool and kindergarten at their preschools. Children were identified as Spanish-English DLLs in the fall of preschool by teachers, and all assessments were administered in Spanish by a bilingual research assessor.

Measures

Self-regulation

In all 4 samples, self-regulation was assessed through the administration of the HTKS task (McClelland et al., 2014), a measure that has demonstrated strong validity and reliability in both English-speaking and Spanish-speaking samples (Lonigan, Lerner, Goodrich, Farrington, & Allan, 2016; McClelland et al., 2007). Although the HTKS has gone through a number of revisions, it was originally adapted from a measure called the “Head and Feet” task described by McCabe, Cunnington, and Brooks-Gunn, which asked children to remember two rules (McCabe, Rebello-Britto, Hernandez, & Brooks-Gunn, 2004). The HTKS task begins with a verification exercise asking the child to mirror the examiner first touching their own head followed by touching their own toes. The child is then instructed to touch their toes when asked to touch their head and to touch their head when asked to touch toes. During this first phase of the task, 6 practice trials are followed by 10 test trials.

The second phase of the task introduces two additional body parts: knees and shoulders. Five practice trials precede 10 test trials in a similar manner as the first phase. Finally, in the third phase of the task, “head” is paired with “knees,” and “toes” is paired with “shoulders.” The third phase consists of 6 practice trials and 10 test trials.

Each trial is scored on a scale from 0 to 2 (0 = *Incorrect*, 1 = *Self-Correct*, 2 = *Correct*). As recommended to reduce floor effects on the HTKS at younger ages, a total score was computed by summing the scores for both test trials as well as practice trials (Fuhs, Nesbitt, Farran, & Dong, 2014). The maximum number of points for the entire task was 94, 34 points for practice trials and 60 points for test trials. Of the 459 children in the study sample, 73% had two or more HTKS assessments over the period of observation. The internal reliability of the HTKS measure ranged from 0.89 to 0.97 across the 4 sites.

There were differences between the sites with regard to where the HTKS task was administered. For the DPREP and TIER samples, the task was administered in the child’s home as part of a larger data collection protocol. For both the Dallas Bilingual and OSU samples, the task was administered in a private testing space in the child’s school.

Child expressive language use

The availability of data regarding the child’s use of Spanish and English differed across studies. For TIER and OSU, only language of assessment administration was available. However, for the Dallas Bilingual and DPREP samples, more detailed child language data were available. At each time point in both studies, parents reported the language their child used to speak to parents, siblings and others outside the family. For the Dallas Bilingual sample, parents responded to 3 expressive language items: language spoken to mother, language spoken to father, and language spoken to other children (Francis et al., 2005). Each item used a 5-point Likert scale (1 = *Only Spanish*; 2 = *Mostly Spanish but Some English*; 3 = *Both Equally*; 4 = *Mostly English but Some Spanish*; 5 = *Only English*). Each item was recoded such that a 1 was assigned if the child spoke both English and Spanish equally, a 0 was assigned if the child spoke either English or Spanish, and a 0.5 was assigned if the child spoke mainly one language but some of the other language. These recoded variables were averaged, resulting in a composite expressive language use variable, with values closer to 1 indicating the child spoke both languages equally more of the time.

For DPREP, child language use was assessed using questions from the Receptive One Word Picture Vocabulary Test – Spanish Bilingual Edition (ROWPVT-SBE, Brownell, 2001). Parents reported on 5 expressive language items: language spoken with mother, language spoken with father, language spoken with siblings, language spoken with friends in school, and language spoken with friends outside school. Responses used a 3-point scale (1 = *Spanish*, 2 = *Both*, 3 = *English*). Each of these 5 variables were recoded such that a 1 was assigned if the child spoke both languages and a 0 if the child only spoke one language with that conversation partner. These 5 variables were averaged, creating a composite variable, and values closer to 1 indicated the child spoke both languages equally more of the time.

Socioeconomic status

A composite index of family socioeconomic status (SES) was created by using a variety of variables available across the 4 sites. All 4 sites had collected data regarding maternal education, dichotomized as less than high school versus high school or greater. Family enrollment in the Supplemental Feeding Program for Women, Infants, and Children (WIC) was available for 3 of the 4 sites (DPREP, TIER, and OSU).

Finally, household physical disorder data were available for DPREP. The household physical disorder composite was comprised of 3 items rated on 4- point Likert scales and 10 yes/no items that the research assistant conducting the home visit used to rate the physical condition of the home, both inside and outside. The measure was developed specifically for this study based on input from research assistants’ observations in the homes and is similar to the household disorganization scale reported by Vernon-Feagans et al. (2012). The three 4-point items included the cleanliness of the home’s inside (0 = *Very Clean*, 3 = *Very Dirty*), organization of the inside of the home (0 = *Very Organized/Uncluttered*, 3 = *Very Disorganized/Cluttered*), and the condition of the lawn (0 = *Very Well-Maintained*, 3 = *Not Cared For at All/Very Overgrown*). The 7 yes/no items included evidence of pests in the home, safety issues, unpleasant smells in the home, peeling paint on the outside of the home, and the presence of discarded appliances, trash, or vehicles in the yard. The 10 items were summed to create a composite index of household physical disorder, with higher values reflecting higher levels of household physical disorder. The internal reliability of the composite was 0.75.

To create an indicator of family SES, a confirmatory factor model was fit in M*plus* (Muthén & Muthén, 1998–2017) using maternal education, WIC enrollment, and household physical disorder as indicators. Maternal education was dichotomized as less than high school or more than high school, WIC receipt was dichotomized as yes or no, and household physical disorder were averaged across waves. We used maximum likelihood robust (MLR) estimation to accommodate missing data and categorical variables as indicators. MPlus does not provide fit indices when using MLR with categorical variables. Instead, univariate and bivariate standardized *z*-scores reported in the TECH10 output are used to assess fit. Standardized residual scores below 1.96 indicate a good model fit. For the SES factor, univariate standardized residual scores were below |.09|, and bivariate standardized residual scores were below |1.0|. Finally, we obtained factor scores for the latent variables to use as a predictor in the latent profile analysis.

Covariates

Covariates included child gender and study site.

Analytic approach

To facilitate a more nuanced examination of patterns of growth in HTKS performance, data were divided into 3-month child age bands ranging from 37 months to 87 months of age. Due to floor effects at the youngest ages, the first 4 age bands (37 months to 48 months) were combined into a single age band. We used latent growth curve models (LGCMs) to estimate individual trajectories of development of self-regulation. The first step was to fit an unconditional growth curve plotting the change in self-regulation across time for each child in the study. The next step was to identify an average intercept (i.e., initial level) of self-regulation scores across all children, an average slope or rate of change in scores, and whether this change was linear, curvilinear, or some other form. We estimated 4 trajectory models: linear, quadratic, latent basis, and piecewise. To evaluate goodness-of-fit between nested models: linear, quadratic, and latent basis, we utilized the nested chi-square difference test (χ^2_{DIFF}). A significant χ^2_{DIFF} value is evidence that the fuller model (i.e., model with more parameters and fewer degrees of freedom) is the better fitting model. Because the piecewise model is a non-nested model, we also evaluated optimal model fit by utilizing the comparative fit index (CFI, Hu & Bentler, 1999) and the Akaike Information Criteria (AIC, Brown, 2015).

We used a latent class growth analysis (LCGA) approach to identify trajectory classes of growth in HTKS performance and the optimal number of growth profiles. Within the LCGA models, we restricted trajectory shape to the shape indicated by the LGCM, and within class, variance was fixed to zero. In these analyses, we compared several models with varying numbers of latent classes by evaluating the following model fit indices: Bayesian Information Criteria (BIC), Sample-Size Adjusted Bayesian Information Criterion (SSABIC), Lo-Mendell-Rubin Likelihood Ratio Test (LMR-LRT, Lo, Mendell, & Rubin, 2001), and the Bootstrapped Likelihood ratio test (BLRT, Peel & McLachlan, 2000). Best practices in evaluating LCGA model fit suggest comparing fit indices simultaneously, as no one model fit index is considered most appropriate (Nylund, Asparouhov, & Muthen, 2007). The BIC statistics (BIC and SSABIC) account for model fit, number of parameters, and sample size. Lower BIC and SSABIC values suggest better model fit. The LMR-LRT and the BLRT are used to compare nested models (Lo, Mendell, & Rubin, 2001; Nylund, Asparouhov, & Muthen, 2007). Statistically significant *p*-values for the LMR-LRT and the BLRT indicate the given model has a significantly better fit to the data than a model with 1 less class (Nylund, Asparouhov, & Muthen, 2007). Additionally, *entropy*, a measure of classification precision, is a standardized estimate with values ranging between 0 and 1 with higher values indicating better class separation (Nagin, 2005). We also considered expectations based on prior research and the conceptual interpretability of each class solution taking into account the need for parsimony in model selection (Nylund, Asparouhov, & Muthen, 2007).

Following the determination of classes, we added child gender and household SES to the model as predictors of class membership. Finally, because expressive language use was only available for a subset of the sample, we exported the class membership variable and used analysis of variance to examine whether language use differed by HTKS growth class. Except for this analysis of variance, all analyses were completed utilizing using M*plus* 8.1 (Muthén & Muthén, 1998–2017), including full information maximum likelihood (FIML) to account for missing data.

Results

Characteristics of the study sample are displayed in Table 1. Overall, there were slightly more boys (56%) in the sample. Level of maternal education differed significantly across sites, with a higher proportion of DPREP and OSU mothers having less than a high school education than the Dallas Bilingual and TIER samples, 51% and 52% vs 12% and 23%, respectively.

Descriptive statistics for HTKS scores by child age and study site are displayed in Table 2. As noted above, the HTKS has a potential range of 0–94. Significant floor effects were noted at the youngest age, particularly for the DPREP and TIER samples, each of which had an average HTKS score of less than 3 for the 37–48- month age band. Although the Dallas Bilingual sample was approaching a ceiling score on average by the oldest age range (e.g., 85–87 months), the other 3 study sites did not appear to be approaching the maximum HTKS score.

Latent growth curve analysis

Due to issues of non-convergence, models were examined in an iterative fashion that dropped individual age bands from the analysis one at a time. As a result of this process, the first 3 age bands (37–54 months) were excluded from the analysis, leaving a total of 418 children for the latent growth curve analysis. Fit indices for the 4 different unconditional growth models are displayed in Table 3. The quadratic model fit significantly better than the linear model as evidenced by a significant improvement in model fit, χ^2 (4) = 29.50, $P = 0.000$. The latent basis model did not fit the data better than the linear model. The fit indices for the quadratic versus the piecewise model were very similar, although the AIC for the quadratic model was slightly lower than the piecewise model, $\Delta\mathrm{AIC}_{\mathrm{quadratic:\ piecewise}} = -7.8$, indicating a better fit.

Latent class growth analysis

Model fit indices and distributions of cases across classes are displayed in Table 4. Although the 2-class model had the highest entropy, the lower BIC and SSABIC values for the 3, 4, and 5 class models indicated the 2- class model was not the best fit. Likewise,

Table 1
Dual language learner (DLL) sample characteristics (N = 459).

	Full Sample (N = 459)	Dallas Bilingual (n = 75)	DPREP (n = 156)	TIER (n = 164)	OSU(n = 64)	χ^2
	N (%)	N (%)	N (%)	N (%)	N (%)	
Child sex						
Boy	256 (55.8)	35 (46.7)	83 (53.2)	100 (61.0)	38 (59.4)	5.08
Girl	203 (44.2)	40 (53.3)	73 (46.8)	64 (39.0)	26 (40.6)	
Maternal education						
Less than high school	158 (34.3)	9 (12.0)	79 (50.6)	37 (22.6)	33 (51.6)	77.76***
High school	161 (35.1)	31 (41.3)	53 (34.0)	62 (37.8)	15 (23.4)	
More than high school	127 (27.7)	35 (46.7)	24 (15.4)	65 (39.6)	3 (4.7)	
Missing	13 (2.8)	0 (0.0)	0 (0.0)	0 (0.0)	13 (20.3)	
Child language use w/parents (initial)[a]						27.00***
DLL, Spanish-dominant	168 (36.6)	36 (48.0)	132 (84.5)	–	–	
DLL, balanced	41 (8.9)	25 (33.3)	16 (10.3)	–	–	
DLL, English-dominant	14 (3.1)	7 (9.3)	7 (4.5)	–	–	
Not available	236 (51.4)	7 (9.3)	1 (0.6)	–	–	
Child language use w/others (initial)[a]						11.56**
DLL, Spanish-dominant	141 (30.7)	32 (42.7)	109 (69.9)	–	–	
DLL, balanced	49 (10.7)	18 (24.0)	31 (19.9)	–	–	
DLL, English-dominant	31 (6.8)	16 (21.3)	15 (9.6)	–	–	
Not available	238 (51.9)	9 (12.0)	1 (0.6)	–	–	

Note. DPREP = Dallas Preschool Readiness Project; TIER = Tufts Interdisciplinary Evaluation Research; OSU = Oregon State University.
** $P < 0.01$.
*** $P < 0.001$.
[a] Child language dominance data were only available for the two Dallas samples

Table 2
Descriptive statistics for HTKS by sample and age.

Age in months	Full sample n	M	SD	Dallas bilingual n	M	SD	DPREP n	M	SD	TIER n	M	SD	OSU n	M	SD
37–48	119	5.21	10.66	24	14.21	19.24	91	2.96	5.21	4	2.50	1.92	–	–	–
49–51	31	20.16	24.30	22	23.41	26.71	–	–	–	7	15.43	16.41	2	1.00	1.41
52–54	63	23.59	26.13	28	30.57	28.04	–	–	–	23	24.52	26.25	12	5.50	7.59
55–57	73	26.78	28.57	21	48.10	32.46	–	–	–	37	21.22	21.74	15	10.67	20.92
58–60	80	26.01	29.34	25	48.80	31.19	–	–	–	29	18.76	24.52	26	12.19	18.29
61–63	88	35.93	34.08	26	50.54	34.60	–	–	–	29	40.72	31.50	33	20.21	30.08
64–66	84	45.55	30.39	29	60.76	27.43	1	69.00	–	26	46.19	26.82	28	28.36	28.60
67–69	66	42.21	31.52	16	67.00	23.50	3	52.67	34.96	22	41.64	30.03	25	25.60	27.26
70–72	79	57.46	29.46	19	71.95	26.72	20	58.65	26.73	19	56.58	27.38	21	44.00	31.58
73–75	108	60.59	27.23	15	74.80	22.21	44	63.66	19.43	31	65.74	24.27	18	32.39	34.26
76–78	93	66.96	24.61	13	83.77	9.94	50	65.24	22.43	15	72.60	21.39	15	52.47	33.83
79–81	54	69.31	23.98	9	83.78	7.38	24	72.21	19.04	10	57.70	29.57	11	61.73	31.26
82–84	34	71.88	20.19	5	82.40	10.43	23	69.57	23.16	6	72.00	10.37	–	–	–
85–87	55	78.35	12.95	7	83.14	8.40	46	78.89	11.56	2	49.00	26.87	–	–	–

Table 3
Comparison of unconditional latent growth curve models.

Model No.	Type	Model fit indices χ^2 (df)	*P*	Model comparison Models	$\Delta\chi^2$(df)	*P*-value	CFI / TLI	SRMR	RMSEA	AIC
1	Linear	48.78 (40)	0.02	–	–	–	.953/0.958	0.219	0.023	6769.91
2	Quadratic	19.28 (36)	0.93	M1 vs M2	29.19 (4)	< 0.001	1.00/1.09	0.141	0.000	6748.42
3	Latent Basis	44.62 (33)	0.01	M1 vs M3	4.16 (7)	0.761	.938/0.932	0.224	0.029	6779.78
4	Piecewise	17.08 (31)	0.90	–	–	–	1.00/1.08	0.148	0.000	6756.22

Table 4
Latent class fit indices of growth model parameters.

Fit statistics	2 classes	3 classes	4 classes	5 classes
−2LL (No. of parameters)	−3257.16 (18)	−3183.96 (24)	−3159.04 (30)	−3138.97 (36)
BIC	6622.96	6512.78	6499.15	6495.22
SSABIC	6565.84	6436.62	6403.95	6380.98
Entropy	0.834	0.655	0.668	0.611
LMR LR test *P*-value	288.00 (0.000)	142.46 (0.000)	49.84 (0.216)	40.15 (0.479)
BLRT test *P*-value	<0.01	<0.01	<0.01	<0.01
Group size (%)				
Class 1	77 (18.42)	170 (40.67)	50 (11.96)	39 (9.33)
Class 2	341 (81.58)	171 (40.91)	163 (39.00)	154 (36.84)
Class 3		77 (18.42)	36 (8.61)	118 (28.23)
Class 4			169 (40.43)	48 (11.48)
Class 5				59 (14.12)

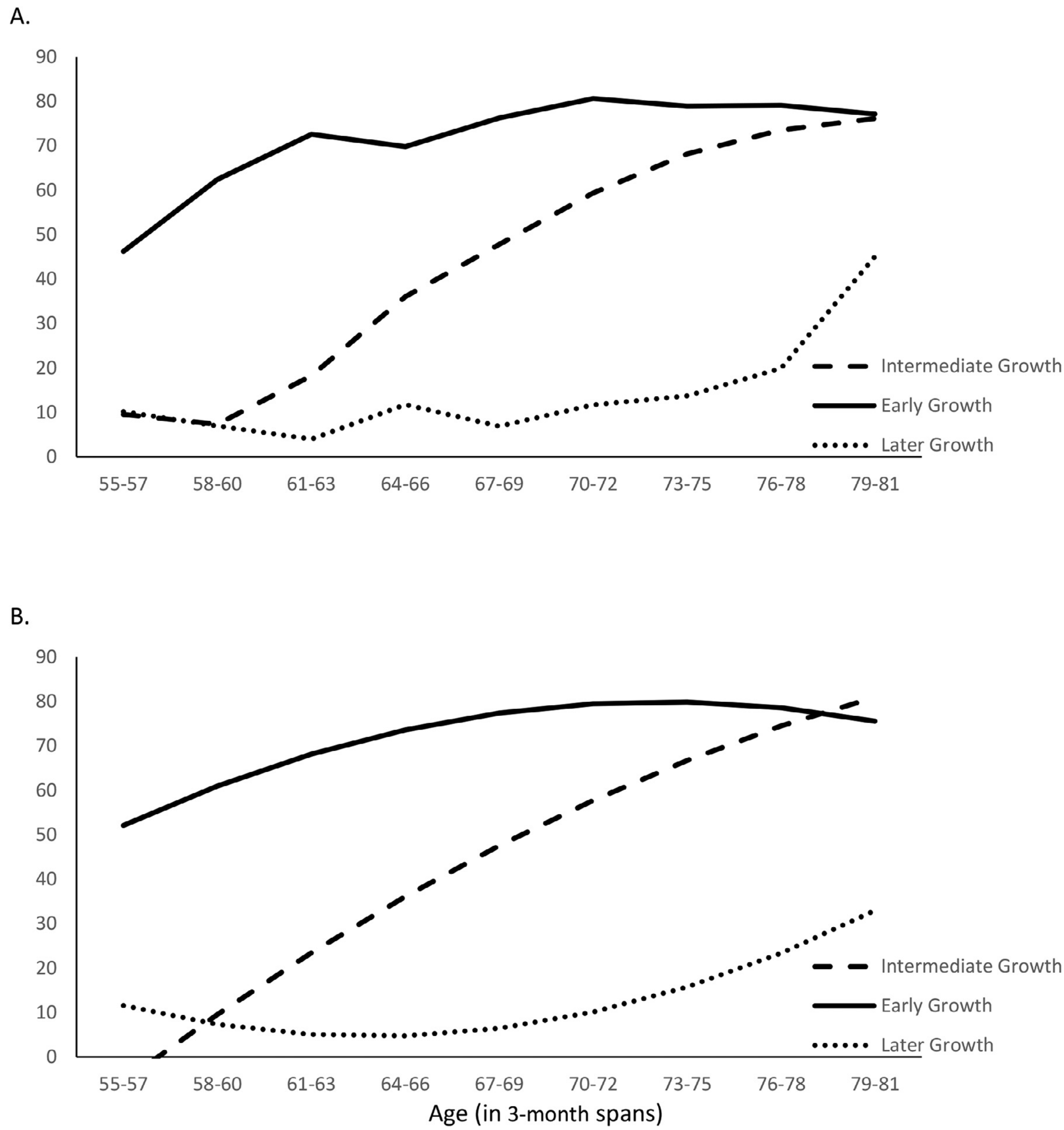

Fig. 1. Observed (A) and estimated (B) trajectories for self-regulation.

although BLRT is significant for all 4 models, the LMR LR was only significant for the 2 and 3 class models. The non-significant LMR LR for the 4-class model suggested that the 3-class model was a better fit. In addition, the 4- and 5-class models had one group that was very small (less than 10% of the cases), and the log-likelihood value of the k-1 class model was non-replicated. Taking this into consideration, the 4- and 5-class models were rejected in favor of the more parsimonious 3-class model (Feldman, Masyn, & Conger, 2009).

Both the observed and predicted growth curves for the 3-class model is displayed in Fig. 1, and descriptive data for the growth parameters for each group are displayed in Table 5. All pairwise comparisons, after a Bonferonni correction, of the growth parameters differed significantly for all 3 groups. The Early group, 40% of the sample, started the observation period with HTKS scores that were at least 4 times as high as the other 2 groups. Children in this group gained more than 10 points in the HTKS at each observation point with growth, slowing and leveling out in the high 70 s. Children in the Intermediate group, about 41% of the sample, started with the lowest scores but displayed the fastest growth starting when children were about 5 years of age. The HTKS scores of children in the Intermediate group caught up with the Early group by the end of the observation period at approximately age 6.5 years. In contrast, children in the Later group (19%) started with HTKS scores that were higher than the Intermediate group but evidenced declining scores early followed by increasing scores later in the observation period. However, the HTKS performance of children in the Later group did not catch up with the other 2 groups by the end of the observation period.

Predictors of pattern of growth in HTKS performance

Table 6 displays results of regressing the 3 latent growth classes on child gender and family socioeconomic status. For child gender, only the Later Group vs. Early Group comparison was statistically significant. The odds ratio was 0.51 [95% CI 0.27, 0.95], $d = 0.28$,

Table 5
Growth parameters for 3 different patterns of HTKS growth.

Group	n (%)	Intercept M (SE)	Growth rate M (SE)	Change in growth rate M (SE)
Early	171 (40.7)	41.07 (0.85)	10.61 (0.09)	−0.78 (0.01)
Intermediate	170 (40.9)	6.20 (0.78)	13.41 (0.17)	−0.57 (0.01)
Later	77 (18.4)	10.52 (0.27)	−3.60 (0.36)	.86 (0.03)

Table 6
Multinomial logistic regression of HTKS growth group on child gender and household socioeconomic status.

	Early vs intermediate		Later vs intermediate		Later vs early	
	OR	95% CI	OR	95% CI	OR	95% CI
Child gender						
Girls vs boys	1.37	[0.72, 2.62]	0.70	[0.35, 1.39]	0.51	[0.27, 0.95]
Family socioeconomic status						
Lower (SES = −0.5 SD)*	1.06	[0.47, 2.41]	1.88	[0.79, 4.47]	1.76	[0.82, 3.80]
Middle/Median (SES = 0.1 SD)	0.99	[0.44, 2.69]	0.88	[0.37, 2.42]	0.89	[0.50, 2.55]
Mid-Upper (SES = 0.5 SD)**	0.94	[0.42, 2.13]	0.53	[0.22, 1.27]	0.57	[0.25, 1.22]
Upper (SES = 1.00 SD)	0.88	[.39, 2.00]	0.28	[0.12, 0.68]	0.32	[0.15, 0.69]

Note. All models are adjusted for study site.
* $P < 0.10$.
** $P < 0.05$.

indicating that girls were half as likely as boys to be in the Later growth group relative to the Early growth group.

Family SES was associated with the probability of being in the Later Group vs. the Intermediate Group as well as the probability of being in the Later Group vs the Early Group. Because SES was included as a continuous predictor, we estimated odds ratios at 4 levels of SES: 0.5 *SD* below the mean, 0.1 *SD* above the mean, 0.5 *SD* above the mean, and 1.0 *SD* above the mean. Odds ratios and 95% confidence intervals are displayed in Table 6. Only odds ratios for children at the upper end of the SES distribution were significant for the Later versus Intermediate comparison, $OR = 0.28$, $d = 0.15$ and the Later versus Early comparison $OR = 0.32$, $d = 0.18$. Children at the upper end of the SES distribution were 72% less likely to be in the Later Group compared to the Intermediate Group, and 68% less likely to be in the Later Group compared to the Early Group.

Finally, analysis of variance (ANOVA) was used to examine whether language use differed between the HTKS growth groups. The average language use score was 0.39 (SD = 0.26) for the Early Group, 0.35 (SD = 0.27) for the Intermediate Group, and 0.33 (SD = 34) for the Later group. These means were not statistically different, $F(2, 177) = 0.63$, $P = 0.54$.

Discussion

Individual differences in child self-regulation skills have been identified as a strong predictor of school readiness and early academic achievement. Children who enter formal schooling with better self-regulation skills display more optimal outcomes in both academic and social domains (Kia-Keating, Nylund-Gibson, Kia-Keating, Schock, & Grimm, 2018; Mills et al., 2019; Schmitt, Geldof, Purpura, Duncan, & McClelland, 2017). Therefore, improving our understanding of the determinants of individual differences in the development of self-regulation during early childhood has the potential for supporting the development of interventions to improve school readiness. However, examinations of individual differences in self-regulation trajectories are limited in general and non-existent for Spanish-English DLLs who are Hispanic, a population uniquely at risk for early academic failure. The present study extends the literature by employing an integrated data analysis of 4 samples to model latent group trajectories and examine individual differences in self-regulation development in a large sample of Spanish-English DLLs.

Individual differences in self-regulation development

One of the goals of this study was to examine characteristics of growth in self-regulation among Spanish-English DLLs. Consistent with findings for monolingual children (Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016), the growth of self-regulation in this sample of Spanish-English DLLs followed a negative quadratic pattern, with faster growth in early childhood followed by a slowing of growth during early elementary school. Also consistent with findings for monolingual children, there was significant variability in both levels of self-regulation skill as well as in patterns of growth during early childhood. The findings of the present study aligned with those of Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 rather than those of Wanless et al. (2016) by identifying 3 groups rather than 2. Children were classified into 3 groups based on their patterns of self-regulation growth: a group that acquired self-regulation skills early, a group that acquired them later, and a group that acquired self-regulation at a rate intermediate to the other 2 groups. It should also be noted that the Wanless et al. (2016) study was conducted in a different country/cultural context (Taiwan), while the Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 study was conducted in the U.S. More research on trajectories of self-regulation development in diverse cultural contexts is needed.

While the early and intermediate groups displayed rapid growth in their self-regulation skills during early childhood, children in the latter group showed virtually no growth until about age 6 years. Approximately 19% of the children in this sample of Spanish-English DLLs fell into the later growth group, which is almost identical to the 20% of children in Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 sample of monolingual English-speaking children who fell into the later self-regulation growth group. In our sample of Spanish-English DLLs as well as in some, but not all, of the children in the Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 analysis, children in the later growth group lagged significantly behind other children in early elementary school in terms of their self-regulation skills. As has been well-documented, deficits in self-regulation are associated with poorer academic achievement and social competence (Schmitt, Geldof, Purpura, Duncan, & McClelland, 2017).

The results of our analysis as well as those reported by Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 indicate that as many as one in 5 young children may be at risk for experiencing deficits in self-regulation development. These findings have significant implications for policy makers and programming to improve school readiness. Early identification of children at risk would provide the opportunity for early intervention to support self-regulation development of in these children. The HTKS measure used in this study has the potential to serve as a no-cost screening tool that could be used to identify children at risk. Early in childhood, children at risk for delayed self-regulation development are indistinguishable from children in the intermediate growth group, those whose initial skills are negligible but who subsequently display rapid growth as they move into formal schooling. As such, a single assessment of self-regulation skill during early childhood is likely insufficient to identify those at risk. Repeated screening during preschool using the HTKS could identify children whose self-regulation skills are stagnant for targeted support. Because the HTKS requires no special equipment, little training, and very little time to administer, routine screening systems for young children in pediatrician's offices and/or preschools could integrate the HTKS as a way of identify children who are lagging behind in their self-regulation development.

Determinants of individual differences in self-regulation development

The similarities in patterns of self-regulation growth between the Spanish-English DLLs in the present analysis and the monolingual English-speaking sample reported by Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 suggest these variabilities represent normative developmental processes. In addition, there were similarities in the predictors of self-regulation development between the 2 samples. In the Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 study, level of maternal education, as a proxy for socioeconomic status, was a significant correlate of individual differences in patterns of growth in self-regulation. In our sample of Spanish-English DLLs, we used a composite measure of socioeconomic status that combined maternal education, receipt of WIC social service benefits, and household physical disorder. Children at the upper end of the socioeconomic distribution in the present sample were significantly less likely to be in the later growth group. A similar positive relation was found between maternal education and growth of self-regulation skills in the Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 sample.

There were some differences, however, between the current findings and those of Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016. Most notably, the accelerating growth of self-regulation skills appeared to commence at a later age among the Spanish-English DLLs in the present sample relative to the monolingual English speakers of the Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 sample. For 2 of the subsamples in the Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 analysis, acceleration began at approximately 3½ years of age, whereas in the current Spanish-English DLL sample, the acceleration began about 18 months later, around age 5 years. This delay in self-regulation development may be a function of socioeconomic status differences between the 2 studies, as most of the participants in the Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 study was drawn from samples that were middle- or high-SES. In contrast, families in the current study that were at the upper end of the distribution of SES were still relatively low-income. For example, average household income available for the families in DPREP sample indicates that the upper end of income distribution was below 150% of the federal poverty level. Our findings indicate that, even within a relatively low-income sample, variability in family economic status is a powerful predictor of self-regulation skill trajectories, as has been confirmed in studies of correlates of self-regulation skill at a single point in time (Raver, Blair, & Willoughby, 2013; Sektnan, McClelland, Acock, & Morrison, 2010).

Consistent with research indicating girls often display higher levels self-regulation compared to boys (Kia-Keating, Nylund-Gibson, Kia-Keating, Schock, & Grimm, 2018), we found that girls were significantly less likely to be in the later growth group relative to the early group. Findings regarding gender differences in self-regulation growth trajectories have been inconsistent Wanless et al. (2016). did not find gender differences in their trajectory analysis of self-regulation development in a sample of Taiwanese children. In contrast, Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016 found girls were more likely to be early developers but only in one of the 3 samples they studied. As the majority of research on gender differences in self-regulation has been cross-sectional, more studies are needed that examine gender differences in self-regulation growth trajectories.

Language use and growth in self-regulation in early childhood

As noted in the introduction, speaking 2 languages have been identified as conveying advantages in the executive function skills underlying self-regulation development. Specifically, children who speak 2 languages appear to have better executive function skills related to cognitive flexibility and attention shifting (Barac, Bialystok, Castro, & Sanchez, 2014) although this "bilingual advantage" is not universally supported in the literature (Dick et al., 2019). However, research on this bilingual advantage has been limited to cross-sectional comparisons of bilingual and monolingual children. Likewise, the inclusion of Spanish-English DLLs in the literature on the bilingual advantage in executive function has been limited. Based on this literature, we hypothesized that the Spanish-English DLLs in our sample who spoke both English and Spanish would display faster growth of self-regulation skills. Our study findings did not support our hypothesis.

There are several reasons why we may not have identified an advantage for Spanish-English DLLs children in self-regulation in the present sample. First, there are methodological considerations. In previous studies, monolingual children were compared to DLLs, whereas the present investigation utilized a measure of expressive language use was based on parental report to examine how individual differences were related to self-regulation growth. As such, our analysis approach is less likely to be confounded by unmeasured differences between monolingual children and DLLs. However, because our measure of expressive language use was based on parent report of child expressive language preference with parents and other individuals, it did not provide a nuanced assessment of the child's English language skills in this sample of Spanish-English DLLs. Future longitudinal research on self-regulation development should consider including direct assessment of child language use along with repeated measures of child self-regulation to examine whether bilingualism is related to individual differences in growth of self-regulation skills.

Limitations

There are several limitations of the current investigation that should be kept in mind when interpreting the results. Although combining samples from multiple sites increased statistical power as well as generalizability, we were limited in the range of covariates common across the different studies from which the samples were drawn. For example, household income was not available for all samples, limiting our measures of household socioeconomic status. Likewise, parental report measures of language use were only available for 2 of the subsamples, and those measures were not ideal. More nuanced measures of child language use as well as repeated assessments to capture changes in bilingualism that may have implications for self-regulation development. For example, one of the co-authors has used biannual narrative language sampling by having Spanish-English language learners retell a story from a wordless picture book in both English and Spanish. These data allowed for growth curve modeling of language use over the first 3 years of formal schooling (Rojas & Iglesias, 2013). Such nuanced data on language use may have demonstrated a different relation with self-regulation development.

The present study would have benefited from other measures of the household context that are important for self-regulation. For example, differences in other family factors such as acculturation stress and parental sensitivity have been identified as important correlates of child self-regulation skills both at a single point in time (Vernon-Feagans, Willoughby, & Garrett-Peters, 2016; von Suchodoletz, Trommsdorff, & Heikamp, 2011) as well as of individual differences in self-regulation growth trajectories (author reference; NICHD ECCRN, 2008). The covariates for the present study were limited to those that were available for all 4 study sites.

Conclusions

The growth of self-regulation for Spanish-English DLL children living in the United States follows a trajectory of development largely similar, albeit somewhat protracted or delayed, to the trajectory of growth which has been found in prior studies of monolingual English-speaking peers. Meaningfully distinct trajectories of self-regulation growth exist for Spanish-English DLL children, and associations for these trajectories with socioeconomic status and child gender, support the implementation of interventions targeting self-regulation skills for children experiencing economic disadvantage. Specifically, with children in the late group showing little to no growth until about age 6 years (the age by which children are typically starting formal schooling), recommendations to policymakers regarding the implementation of programs or interventions to promote self-regulation prior to the transition to kindergarten are critical to providing disadvantaged DLL children with the skills necessary to learn in a school setting.

Author contributions

Margaret O'Brien Caughy: Conceptualization, Methodology, Formal analysis, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition, Dawn Y. Brinkley: Conceptualization; Methodology; Formal analysis, Writing – original draft; Writing – review & editing; Visualization, Daniel Pacheco: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Raul Rojas: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Funding acquisition, Alicia Miao: Conceptualization; Methodology; Formal analysis, Writing – review & editing, Mariah M. Contreras: Conceptualization; Methodology; Formal analysis, Writing – original draft; Writing – review & editing, Margaret Tresch Owen: Writing – review & editing, M. Ann Easterbrooks: Writing – review & editing, Megan McClelland: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Funding acquisition.

Funding

This research was supported by the Eunice Kennedy Shriver National Institute of Child Health and Development (Grant numbers 1R01HD058643, 2R01HD075311), the Jerry M. Lewis, M.D. Mental Health Research Foundation (Grant GA-2013–016), an Anonymous Donor (Grant 13180), the Children's Trust of Massachusetts (MA 5014), U.S. Department of Health and Human Services under the Maternal, Infant and Early Childhood Home Visiting Grant Program (Grant X10MC29474), and the U.S. Department of Education Institute for Education Sciences (Grant R305A100566).

References

Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly, 29*(4), 699–714. https://doi.org/10.1016/j.ecresq.2014.02.003.

Blair, C. B. (2003). Behavioral inhibition and behavioral activation in young children: Relations with self-regulation and adaptation to preschool in children attending Head Start. *Developmental Psychobiology, 42*(3), 301–311. https://doi.org/10.1002/dev.10103.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford Publications.

Brownell, R. (2001). *Receptive one-word picture vocabulary tests*. Harcourt Assessment, Inc.

Cameron Ponitz, C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., Morrison, F. J., et al. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly, 23*(2), 141–158. https://doi.org/10.1016/j.ecresq.2007.01.004.

Cameron Ponitz, C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*(3), 605–619. https://doi.org/10.1037/a0015365.

Caughy, M. O., Owen, M. T., Mills, B. A., & Hurst, J. R. (2013). Emergent self-regulation skills among very young ethnic minority children: A confirmatory factor model. *Journal of Experimental Child Psychology, 116*(4), 839–855. https://doi.org/10.1016/j.jecp.2013.07.017.

Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods, 14*(2), 81–100. https://doi.org/10.1037/a0015914.

Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., … Gonzalez, R., et al. (2019). No evidence for a bilingual executive function advantage in the ABCD study. *Nature Human Behaviour, 3*(7), 692–701. https://doi.org/10.1038/s41562-019-0609-3.

Evans, G. W., & Rosenbaum, J. (2008). In *Self-regulation and the income-achievement gap: 23* (pp. 504–514). Early Childhood Research Quarterly. https://doi.org/10.1016/j.ecresq.2008.07.002.

Federal Interagency Forum on Child and Family Statistics. (2019). *America's Children: Key National Indicators of Well-Being, 2019*. Retrieved 01/31/2020 from https://www.childstats.gov/pdf/ac2019/ac_19.pdf .

Feldman, B. J., Masyn, K. E., & Conger, R. D. (2009). New approaches to studying problem behaviors: A comparison of methods for modeling longitudinal, categorical adolescent drinking data. *Developmental Psychology, 45*(3), 652–676. https://doi.org/10.1037/a0014851.

Francis, D. J., Carlson, C. J., Fletcher, J. M., Foorman, B. R., Goldenberg, C. R., Vaughn, S., Papanicolaous, A., et al. (2005). Oracy/literacy development of Spanish-speaking children: A multi-level program of research on language minority children and the instruction, school and community contexts, and interventions that influence their academic outcomes. *Perspectives (Gerontological Nursing Association (Canada)), 31*(2), 8–12.

Francis, D. J., Rojas, R., Gusewski, S., Santi, K. L., Khalaf, S., Hiebert, L., Bunta, F., et al. (2019). Speaking and reading in two languages: On the identification of reading and language disabilities in Spanish-speaking English learners. *New Directions for Child and Adolescent Development, 2019*(166), 15–41. https://doi.org/10.1002/cad.20306.

Fry, R. (2008). *The role of schools in the English language learner achievement gap*. Pew Hispanic Center Retrieved 12/20/21 from https://files.eric.ed.gov/fulltext/ED502050.pdf .

Fuhs, M. W., Nesbitt, K. T., Farran, D. C., & Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology, 50*(6), 1698–1709. https://doi.org/10.1037/a0036633.

Gathercole, S. E. (2008). Working memory in the classroom. *The Psychologist, 21*(5), 382–385.

Gusewski, S., & Rojas, R. (2017). Tense marking in the English narrative retells of dual language preschoolers. *Language, Speech, and Hearing Services in Schools, 48*(3), 183–196. https://doi.org/10.1044/2017_LSHSS-16-0093.

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118.

Jackson, C. W., Schatschneider, C., & Leacox, L. (2014). Longitudinal analysis of receptive vocabulary growth in young Spanish English–speaking children from migrant families. *Language, Speech, and Hearing Services in Schools, 45*(1), 40–51. https://doi.org/10.1044/2013_LSHSS-12-0104.

Kia-Keating, M., Nylund-Gibson, K., Kia-Keating, B. M., Schock, C., Grimm, R. P., et al. (2018). Longitudinal patterns of self-regulation among ethnic minority children facing poverty. *Journal of Child and Family Studies, 27*(2), 398–411. https://doi.org/10.1007/s10826-017-0883-5.

Kopp, C. B. (1982). Antecedents of self regulation: A developmental perspective. *Developmental Psychology, 18*(2), 199–214. https://doi.org/10.1037/0012-1649.18.2.199.

Lo, Y., Mendell, N. R., & Rubin, B. D. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*(3), 767–778. https://doi.org/10.1093/biomet/88.3.767.

Lonigan, C. J., Lerner, M. D., Goodrich, J. M., Farrington, A. L., & Allan, D. M. (2016). Executive function of Spanish-speaking language-minority preschoolers: Structure and relations with early literacy skills and behavioral outcomes. *Journal of Experimental Child Psychology, 144*, 46–65. https://doi.org/10.1016/j.jecp.2015.11.003.

Matthews, J. S., Cameron Ponitz, C., Morrison, F. J., et al. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology, 101*(3), 689–704. https://doi.org/10.1037/a0014240.
McCabe, L. A., Rebello-Britto, P., Hernandez, M., & Brooks-Gunn, J. (2004). Games children play: Observing young children's self-regulation across laboratory, home, and school settings. In R. DelCarmen-Wiggins, & A. Carter (Eds.), *Handbook of infant, toddler, and preschool mental health assessment* (pp. 491–521). Oxford University Press.
McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., Morrison, F. J., et al. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43*(4), 947–959. https://doi.org/10.1037/0012-1649.43.4.947.
McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., Pratt, M. E., et al. (2014). Predictors of early growth in academic achievement: The head-toes-knees-shoulders task. *Frontiers in Psychology, 5*(599), 1–14. https://doi.org/10.3389/fpsyg.2014.00599.
McClelland, M. M., & Cameron, C. E. (2012). Self-regulation in early childhood: Improving conceptual clarity and developing ecologically valid measures. *Child Development Perspectives, 6*(2), 136–142. https://doi.org/10.1111/j.1750-8606.2011.00191.x.
McClelland, M. M., Cameron Ponitz, C., Messersmith, E. E., & Tominey, S. (2010). Self--regulation: Integration of cognition and emotion. In R. M. Lerner, & B. Overton (Eds.), *Handbook of life-span development*. Wiley & Sons.
McClelland, M. M., Geldhof, G. J., Cameron, C. E., & Wanless, S. B. (2015). Development and self regulation. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (pp. 1–43). John Wiley & Sons, Inc.
McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathbun, A., … Bullock Mann, F., et al. (2018). *The Condition of Education 2018* NCES 2018-144). *National Center for Education Statistics*. Retrieved 7/6/18 from https://nces.ed.gov/pubsearch/pubsinfoasp?pubid=2018144.
Mills, B., Dyer, N., Pacheco, D., Brinkley, D., Owen, M. T., Caughy, M. O., et al. (2019). Developmental transactions between self-regulation and academic achievement among low income African American and Latino children. *Child Development, 90*(5), 1614–1631. https://doi.org/10.1111/cdev.13091.
Montroy, J. J., Bowles, R. P., Skibbe, L. E., McClelland, M. M., & Morrison, F. J. (2016). The development of self-regulation across early childhood. *Developmental Psychology, 52*(11), 1744–1762. https://doi.org/10.1037/dev0000159.
Muthén, B., & Muthén, L. (1998-2017). *Mplus user's guide* (8th ed.). Muthén and Muthén http://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf.
Nagin, D. S. (2005). *Group-based modeling of development*. Harvard University Press.
NICHD ECCRN. (2008). Mothers' and fathers' support for child autonomy and early school achievement. *Developmental Psychology, 44*(4), 895–907. https://doi.org/10.1037/0012-1649.44.4.895.
Nylund, K. L., Asparouhov, T., & Muthen, B. (2007). Deciding on the number of classes in latent class analysis and mixture growth modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535–569. https://doi.org/10.1080/10705510701575396.
Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex; a journal devoted to the study of the nervous system and behavior, 69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014.
Peel, D., & McLachlan, G. J. (2000). Robust mixture modelling using the t distribution. *Statistics and Computing, 10*(4), 339–348. https://doi.org/10.1023/a:1008981510081.
Raver, C. C., Blair, C. B., & Willoughby, M. (2013). Poverty as a predictor of 4-year-olds' executive function: New perspectives on models of differential susceptibility. *Developmental Psychology, 49*(2), 292–304. https://doi.org/10.1037/a0028343.
Robson, D. A., Allen, M. S., & Howard, S. J. (2020). Self-regulation in childhood as a predictor of future outcomes: A meta-analytic review. *Psychological Bulletin, 146*(4), 324–354. https://doi.org/10.1037/bul0000227.
Rojas, R., Hiebert, L., Gusewski, S., & Francis, D. J. (2019). Moving forward by looking back: Understanding why some Spanish-speaking English learners fall behind. *New Directions for Child and Adolescent Development, 166*, 43–77. https://doi.org/10.1002/cad.20305.
Rojas, R., & Iglesias, A. (2013). The language growth of Spanish-speaking English language learners. *Child Development, 84*(2), 630–646. https://doi.org/10.1111/j.1467-8624.2012.01871.x.
Rothbart, M. K., & Bates, J. E. (2006). Temperament. In W. Damon, & R. M. Lerner (Eds.), *Social, emotional, and personality development* (pp. 99–166). Wiley.
Rothbart, M. K., & Posner, M. I. (2005). Genes and experience in the development of executive attention and effortful control. *New Directions for Child and Adolescent Development, 2005*(109), 101–108. https://doi.org/10.1002/cd.142.
Schmitt, S. A., Geldof, G. J., Purpura, D. J., Duncan, R., & McClelland, C. M. (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology, 109*(8), 1120–1140. https://doi.org/10.1037/edu0000193.
Sektnan, M., McClelland, M. M., Acock, A., & Morrison, F. J. (2010). Relations between early family risk, children's behavioral regulation and academic achievement. *Early Childhood Research Quarterly, 25*(4), 464–479. https://doi.org/10.1016/j.ecresq.2010.02.005.
Tufts Interdisciplinary Evaluation Research. (2017). The Massachusetts Healthy Families Evaluation-2 Early Childhood (MHFE-2EC): Follow-up study of a randomized, controlled trial of a statewide home visiting program for young parents. *Final report to Massachusetts Department of Public Health, Children's Trust of Massachusetts.*
Vernon-Feagans, L., Willoughby, M., & Garrett-Peters, P. (2016). Predictors of behavioral regulation in kindergarten: Household chaos, parenting, and early executive functions. *Developmental Psychology, 52*(3), 430–441. https://doi.org/10.1037/dev0000087.
von Suchodoletz, A., Trommsdorff, G., & Heikamp, T. (2011). Linking maternal warmth and responsiveness to children's self-regulation. *Social Development, 20*(3), 486–503. https://doi.org/10.1111/j.1467-9507.2010.00588.x.
Wanless, S. B., Kim, K. H., Zhang, C., Degol, J. L., Chen, J. L., Chen, F. M., et al. (2016). Trajectories of behavioral regulation for Taiwanese children from 3.5 to 6 years and relations to math and vocabulary outcomes. *Early Childhood Research Quarterly, 34*, 104–114. https://doi.org/10.1016/j.ecresq.2015.10.001.