Language development as a mechanism linking socioeconomic status to executive functioning development in preschool

Rachel R. Romeo1,2 | John C. Flournoy1 | Katie A. McLaughlin1 | Liliana J. Lengua3

1 Department of Psychology, Harvard University, Cambridge, Massachusetts, USA
2 Department of Human Development and Quantitative Methodology, University of Maryland College Park, College Park, Maryland, USA
3 Department of Psychology, University of Washington, Seattle, Washington, USA

Correspondence
Rachel R. Romeo, Department of Psychology, Harvard University, Cambridge, MA, USA. Email: rachelromeo@fas.harvard.edu

Funding information
the Eunice Kennedy Shriver National Institute for Child Health and Human Development, Grant/Award Numbers: K99-HD103873, R01-HD054465; the National Institute for Mental Health, Grant/Award Number: R01-MH106482

Abstract
Childhood socioeconomic status (SES) is related to disparities in the development of both language and executive functioning (EF) skills. Emerging evidence suggests that language development may precede and provide necessary scaffolding for EF development in early childhood. The present preregistered study investigates how these skills co-develop longitudinally in early childhood and whether language development explains the relationship between SES and EF development. A socioeconomically diverse sample of 305 children completed repeated assessments of language (sentence comprehension) and EF (cognitive flexibility, behavioral inhibition, and cognitive inhibition) at four waves spaced 9 months apart from ages 3 to 5 years. Bivariate latent curve models with structured residuals were estimated to disaggregate between-person and within-person components of stability and change. Results revealed bidirectional relationships between language and EF across all waves. However, at 3 years, language comprehension more strongly predicted EF than the reverse; yet by 5 years, the bidirectional effects across domains did not significantly differ. Children from higher-SES backgrounds exhibited higher initial language and EF skills than children from lower-SES families, though SES was not associated with either rate of growth. Finally, early language-mediated the association between SES and early EF skills, and this model outperformed a reverse direction mediation. Together, results suggest that EF development is driven by early language development, and that SES disparities in EF are explained, at least in part, by early differences in language comprehension. These findings have implications for early interventions to support children’s language skills as a potential pathway to improving early EF development.

KEYWORDS
cognitive development, executive functioning, language, longitudinal modeling, socioeconomic status

1 INTRODUCTION

Childhood socioeconomic status (SES) is broadly associated with multiple domains of neurocognitive development, with children from higher SES backgrounds on average exhibiting more advanced cognitive skills than children from lower SES families (Bradley & Corwyn, 2002; Farah, 2017). Two of the cognitive domains most strongly associated with SES are language skills (for review, see Pace et al., 2017) and executive functioning (EF) skills (for review, see Lawson et al., 2018). Both sets of skills are critical for academic learning, and differences in language and EF each contribute to SES-related disparities in academic achievement (e.g., Burchinal et al., 2002; Durham et al., 2007; Fitzpatrick et al., 2014; Lawson & Farah, 2017; Rosen et al., 2018; Waters et al., 2021).
EF refer to a set of higher-level cognitive skills that support goal-directed action, and typically include inhibitory control, working memory, and cognitive flexibility (Diamond, 2013; Miyake & Friedman, 2012), which develop rapidly throughout early childhood, and especially the preschool years (i.e., ages 3–5 years; Best & Miller, 2010; Zelazo et al., 2003). Similarly, children’s language skills, including both their ability to understand others and to linguistically express themselves, also develop rapidly in the first years of life, with demonstrable lexical and syntactic growth in the toddler and preschool years (Hoff, 2013). In addition to overlapping developmental trajectories, language and EF skills are supported by the development of adjacent and overlapping prefrontal brain systems during childhood (Fiske & Holmboe, 2019; Friederici, 2006; Merz et al., 2019).

On average, children from lower SES backgrounds exhibit reduced performance on both language and EF assessments than their more advantaged peers (Noble et al., 2007). While there are likely many environmental factors contributing to these disparities, one mechanism with strong empirical support is cognitive stimulation (Sheridan & McLaughlin, 2016). Children from higher SES homes are more likely to have access to developmentally appropriate learning materials such as books and toys, experience enriching activities outside the home (e.g., museums), and engage with caregivers in interactive linguistic exchanges and scaffolded learning activities (Rosen et al., 2019). These cognitively stimulating activities, and especially language, rich caregiver-child interaction, independently explain SES differences in both language and EF development (Daneri et al., 2019; Romeo, 2019). Importantly, structural inequities in financial, educational, and social resources combined with the day-to-day stressors of socioeconomic disadvantage systematically constrain caregivers’ capacity to provide cognitively stimulating interactions (Ellwood-Lowe et al., 2021), which may in turn drive SES disparities across multiple cognitive domains, including language and EF development.

Additionally, language and EF skills are conceptually related and developmentally intertwined, such that skill acquisition in one domain scaffolds and supports the emergence of improved skills in the other. For example, language involves a command of abstract verbal labels (vocabulary) that are combined in a hierarchical rule-based system (syntax); use of abstract representations and rule structures are key components of many EF tasks (Marcovitch & Zelazo, 2009). Additionally, children often use “self-talk” to scaffold their performance on challenging tasks (Vygotsky, 1962), and even adults may use verbalizing strategies or subvocalization to support difficult cognitive tasks, including working memory (e.g., the “phonological loop,” Badeley & Hitch, 1974). On the other hand, skills such as attentional control and error monitoring may support vocabulary learning by honing focus to relevant word-concept mappings (Kapa & Erikson, 2020; Weiland et al., 2014), and inhibition may support conversational turn-taking skills by inhibiting responses until the prior speaker is finished (Levinson & Torreira, 2015). Indeed, several longitudinal studies have found bidirectional relationships between these cognitive domains, such that language and EF skills build upon each other in preschool and elementary school (Bohlmann et al., 2015; Daneri & Blair, 2017; Slot & von Suchodoletz, 2018). However, when comparing the directional relationships, most studies find that language development more strongly scaffolds subsequent EF development than the reverse, especially in preschool (Fuhs & Day, 2011; Jones et al., 2020; Kuhn et al., 2016; Slot & von Suchodoletz, 2018; Vallotton & Ayoub, 2011; White et al., 2017).

These findings suggest that variation in children’s language skills could contribute to the well-established relationship between SES and EF development. A few studies provide preliminary evidence for this hypothesis, finding that SES-EF relationships can be explained by receptive language skills (and specifically vocabulary skills) in school-aged children (Catale et al., 2012; Noble et al., 2005, 2007; Pluck et al., 2020), but not by spontaneous expressive language usage (Sarsour et al., 2011). However, these studies are all cross-sectional, which limits the conclusions that can be drawn regarding the mediating role of language development in EF development, and provide little evidence for how these relationships may arise and change across early childhood.

The present study aims to overcome these limitations by investigating the relationships between SES and the development of language and EF skills in a multi-wave longitudinal study spanning ages 3–5 years. We hypothesize that language will more strongly predict EF development than vice versa, and that language skills will mediate the relationship between SES and EF. If confirmed, findings would have implications for supporting language development as a pathway for reducing SES disparities in EF skills.

2 MATERIALS AND METHODS

2.1 Participants

Participants were 306 children (152 male) who were aged 36–40 months (M = 36.6, SD = 0.9) at the first assessment (T1). Three follow-up assessments were spaced 9 months apart: T2 at 44–49 months of age (M = 46.1, SD = 1.1, n = 290), T3 at 53–59 months of age (T3, M = 55.0, SD = 1.1, n = 288), and T4 at 62–69 months of age (M = 63.9, SD = 1.2, n = 288). Of those who participated in the first session, 278
(91%) attended all four sessions, 13 (4%) attended three sessions, six (2%) attended two sessions, and nine (3%) attended only the first session. All study procedures were approved by the institutional review board at the University of Washington, and active parental informed consent and child assent were secured prior to data collection.

Participants were recruited as a part of a longitudinal study of associations between family income and the development of effortful control (Lengua et al., 2015). Families were recruited from various publicly and privately funded sources, including birth registries, daycares, preschools, libraries, health clinics, and charitable agencies and organizations serving low-income families in the urban and suburban areas surrounding Seattle. Recruitment aimed to achieve balanced representation across income levels, with one-third of families at/near poverty, one-third with lower-income, and one-third with middle-to-upper-income, based on the federal poverty guidelines in place at the time of their first visit (2008–2010). Families were required to be proficient in English (self-reported), and children diagnosed with a developmental disability were excluded.

2.2 Socioeconomic status and demographics

Demographic measures were collected through maternal self-report. The SES measures of interest were income-to-needs (ITN) and maternal education (MEd) level, measured at the first wave. Total family income (inclusive of wages, investments, child support/alimony, and state/federal aid) was reported in 14 bins calibrated to correspond to federal poverty guidelines, ranging from less than $14,570 to more than $150,000. Parents also reported the number of family members depending on that income (inclusive of all adults and children), which was used to determine the corresponding federal poverty level. ITN ratios were calculated as the ratio of total income to the federal poverty level, such that values less than one indicate income below the poverty threshold, and values greater than one indicate income above the poverty threshold. Sample ITN ranged from 0.22 to 8.19 (M = 2.2, SD = 1.88). Mothers also reported their highest education level in 10 bins, ranging from less than 8th grade to receiving a doctoral degree or equivalent; the sample mean was a technical/associate’s degree (14–15 years of education). Attrition (i.e., not attending all four sessions) was marginally associated with ITN (t = 1.97, p = 0.06) and MEd (t = 2.04, p = 0.05), such that children from lower-SES families were less likely to attend all sessions. Distributions of the SES measures are in Figure 1. The racial and ethnic make-up of the sample are reported for descriptive purposes. The distribution roughly matched that of the geographical region sampled: 64% White/Caucasian, 9% Black/African American, 3% Asian, 10% Latino/Hispanic, 2% Native American/American Indian, and 12% multiple racial and ethnic backgrounds.

2.3 Language assessment

Children’s language comprehension skills were measured with the Comprehension of Instructions subtest of the NEPSY-II (Korkman et al., 2007), which assesses the ability to receive, process, and execute oral instructions of increasing semantic and syntactic complexity. For each item, the child points to appropriate stimuli in response to oral instruction, and assessment stops when the child gets seven correct consecutively. This measure is strongly associated with composite language indices on comprehensive assessments of receptive and expressive language in children from diverse backgrounds (Korkman et al., 2007). This assessment was administered at all four timepoints. The outcome score was the proportion of correct items out of 33 total items.

2.4 Executive functioning assessment

Children completed an identical battery of EF assessments at each of the four timepoints. Children’s EF skills were measured with five assessments that tapped three different EF components: one measure of cognitive flexibility/switching, two measures of behavioral inhibition, and two measures of cognitive inhibition. Measures were selected to be of varying difficulty for children of varying ages so that identical measures could be used over time. For both behavioral and cognitive inhibition, one measure was better suited for the younger time points (i.e., older children were at ceiling) while the other was better suited for the older time points (i.e., younger children were at floor). The flexibility measure was well suited for children at all four time points. See Supplementary Figure S1 and Supplementary Table S1 for task proficiency across waves.

Cognitive flexibility was assessed with the Dimensional Change Card Sort (DCCS; Zelazo, 2006). In this task, children sort varying in shape and color. Children were shown two boxes, each with an attached target card. In the first part, the target cards included a black figure on a colored background (star on blue, truck on red), and children were instructed to sort a series of bivalent image cards first according to shape (six trials) and then according to color (six trials). In the second part, the sorting properties were integrated, such that target cards consisted of a colored figure on a white background (blue star, red truck), and children were instructed to first sort according to shape (six trials), and then by color (six trials). In the third part, the target cards remained the same, but children were instructed to sort by color if the card had a black border on it, and by shape if the card lacked the border (12 trials). The outcome score was the proportion of correct trials out of the total 36 trials.

Behavioral inhibition was assessed by Monkey/Dragon and Head, Toes, Knees, and Shoulders tasks. Monkey/Dragon is a version of the Bear-Dragon task (Kochanska et al., 1996), except a monkey puppet is substituted in place of the bear. It requires the child to perform actions when the directive is given by the monkey puppet (five trials), but not when given by a dragon puppet (five trials). Children receive 0 points for no movement, 1 point for a wrong movement, 2 points for a partially correct movement, or 3 points for a completely correct movement. The outcome score was total points divided by the total possible score (30 points). This was the easier behavioral inhibition measure, so it is included as an indicator of EF at T1 and T2. In Head, Toes, Knees,
Shoulders (Ponitz et al., 2009), children follow a series of paired behavioral rules. First, children are asked to touch their head when told to touch their toes and vice versa (10 trials). Then, two additional commands are added, and children are asked to touch their knees when told to touch their shoulders and vice versa (10 trials). Finally, the rules are changed, such that head goes with knees and toes go with shoulders (10 trials). Children received 0 points for an incorrect response, 1 point for a self-correct, and 2 points for a correct response. The outcome score was total points divided by the total possible score (60 points). This was the more difficult behavioral inhibition measure, so it is included as an indicator of EF at T3 and T4.

Cognitive inhibition was assessed by Day/Night and NEPSY-II Shape Inhibition tasks. Day/Night (Gerstadt et al., 1994) requires the child to say “day” when shown a picture of moon/stars and “night” when shown a picture of the sun. Responses were scored as correct or incorrect, and the outcome score was the proportion correct out of the total 16 trials. This was the easier cognitive inhibition measure, so it is included as an indicator of EF at T1 and T2. In the shape inhibition subtest of the NEPSY-II, children were shown an array of 40 circles and squares and asked to say the opposite shape name for each item (i.e., circle for square and vice versa) as quickly as possible. Children received 0 points for an incorrect response, 1 point for a self-correct, and 2 points for a correct response. The outcome score was total points divided by the total possible score (80 points). This was the more difficult cognitive inhibition measure, so it is included as an indicator of EF at T3 and T4.

All cognitive assessment scores were logit-transformed prior to analyses to transform their scale from [0, 1] to [-Inf, Inf], which provided continuous, normally-distributed inputs for SEM. Proportion scores of 0 and 1 were substituted with half of the next lowest and highest possible scores, respectively (i.e., for an assessment with 36 items, a score of 0 became \(1/36)/2 = 0.014\), and a score of 1 became \(1-(1/36)/2 = 0.986\). Children’s scores on individual assessments were excluded for child non-compliance (e.g., refused to participate, quit the task early, etc.), for task inappropriateness (e.g., did not understand the instructions, did not know shape names necessary for the inhibition test, colorblindness for the language task that requires identifying colored shapes, etc.), and for experimenter error (e.g., early discontinue, did not administer enough practice trials, etc.). Additionally, assessments with <80% valid responses (i.e., not skipped because of child or experimenter error) were excluded. Of the 16 assessment measures (one language and three EF at each of four time points), 93 participants (30.4%) had complete data, 195 (63.7%) had 10–15 valid assessments, and 18 (5.7%) had fewer than 10 valid assessments. One participant did not have any valid assessments and thus was excluded. Even after accounting for attrition, SES was associated with missingness (ITN: \(r = 0.14, p = 0.02\); MEd: \(r = 0.16, p = 0.01\)). Thus, if SES is related to both missingness and our outcomes of interest, then its inclusion in the model leads to unbiased estimates (Matta et al., 2018).

2.5 Analyses

All analyses were carried out in Mplus version 8.6 (Muthén & Muthén, 2017). Full information maximum likelihood (FIML) was used to deal with missing data. First, latent EF factors were estimated using confirmatory factor analysis with three indicators at each time point, including one measure of cognitive flexibility (DCCS, all waves), one measure of behavioral inhibition (Monkey/Dragon at T1 and T2; HTKS at T3 and T4), and one measure of cognitive inhibition (Day/Night at T1 and T2; Shape Inhibition at T3 and T4). This factor structure was determined prior to pre-registration to limit metric invariance due to ceiling effects for Monkey/Dragon and Day/Night at the later waves, and floor effects of HTKS and NEPSY Inhibition at the earlier waves. Covariances of repeated measures were allowed to vary, and longitudinal measurement invariance was evaluated by constraining loadings of identical measures across waves (Widaman et al., 2010). The resulting EF factors served as indicators for all remaining models.
Longitudinal associations between language and EF were evaluated by estimating a bivariate latent curve model with structured residuals (LCMSR, Curran et al., 2014). This method estimates time-specific regressions on residuals rather than observed variables, which allows for a more rigorous evaluation of bidirectional within-person relationships. Thus, this method is optimal for providing simultaneous estimates of individualized between-subject growth processes and time-specific within-subject processes of the relation between language and EF over four repeated measures. Upon the introduction of each new parameter, increasingly strict levels of invariance were evaluated by imposing equality constraints and conducting model comparisons of with likelihood ratio tests. If model fit decreased significantly, the added constraint was relaxed.

For language and EF separately, univariate unconditional latent curve models were estimated to determine growth trajectories. Second, the residuals were structured such that later residuals were regressed on earlier residuals through time-adjacent autoregressions. Then, the two univariate models were combined into a bivariate LCMSR to assess across-construct relations at the level of the latent factors and of the time-specific residuals. Intercepts and slopes were allowed to covary with one another. The across-construct time-specific residuals were also allowed to covary and constrained to equality for T2, T3, and T4. Unidirectional cross-lagged regression parameters were first freely estimated and then constrained to equality across time to assess parsimonious fit. If fit was significantly reduced, we evaluated whether the regression magnitudes increased/decreased over time with the asymptotic parameter comparison approach (i.e., assessing the significance of the difference term). We also estimated a new parameter that serves to increment the value of the lagged effect with each unit-passage of time to determine whether cross-lag effects changed systematically (Bauer & Hussong, 2009; Curran et al., 2014). In the final bivariate model with the best parsimonious fit, the asymptotic parameter comparison approach was again used to compare the cross-lag regression magnitudes to determine if the effects were stronger in one direction than the other.

Next, the bivariate model was conditioned on participant sex assigned at birth and SES (at T1) by regressing the four latent curve factors (intercept and slope of language and EF) on sex and SES as exogenous, time-invariant predictors. The two SES measures—MEd and ITN—were added to separate parallel models because they were highly colinear ($r = 0.47$, $p < 1^{-17}$), yet may theoretically exert different influences on cognitive development (Duncan & Magnuson, 2003, 2012). Both measures of SES were remarkably stable across the four measurement waves (MEd Cronbach's $\alpha = 0.98$; ITN $\alpha = 0.96$), supporting the inclusion of only T1 measures as predictors. The effects of each covariate on each of the four latent curve factors were evaluated by hypothesis tests and change in model fit was assessed to determine whether non-significant paths were functionally equivalent to zero.

Finally, in each of the parallel models, we evaluated the indirect paths from the SES measure to the EF intercept through the language intercept (and vice versa). The between-intercept covariances were replaced with direct paths and change in model fit was assessed. The significance of the indirect effects was evaluated using bias-corrected bootstrap confidence intervals based on 10,000 replications. Because four indirect effects were assessed (2 SES measures $\times$ 2 potential directions of effects), the 99% confidence interval was used.

These methods differed from the preregistered analysis methods in several minor ways, which were all necessitated by the findings during the model fitting process. First, as described below, several latent factors were largely skewed, and were thus fit better by skew t distributions. While standard structural equation models fit the means and covariances of data, skewed SEM models are a more flexible parametric family of distributions that fit the skewness and kurtosis in addition to the means and covariances, which allows for fitting the entire distributions of the data (Asparouhov & Muthén, 2015). The normal distribution is nested within the skew t distribution, allowing likelihood ratio tests to test the appropriateness of the added parameters, and restricted skew t distributions allow explicit maximum likelihood estimation for structural equation models (Asparouhov & Muthén, 2015). This approach also accounts for skew in the predictor variables, so there was no need to log transform ITN. However, these models do not provide absolute fit indices, so model comparison by the change in the McDonald Noncentrality Index was not possible, and instead models were compared by the change in the Akaike Information Criterion (AIC) and loglikelihood ratio tests (LRT). Additionally, regression coefficients in a model estimated with skew t distributions are not comparable to those in a normal regression (they are usually smaller), because much of the relationship is channeled through a shared skewness factor (Asparouhov & Muthén, 2015). This prevented the selection of equivalence-bounds based on traditional effect sizes, so equivalence tests of regression parameters were replaced with comparison to nested models with non-significant parameters set to zero. Second, linear growth models were hypothesized, but these fit the data poorly, so models were estimated using free slopes to determine the shape of growth. This prevented the use of TSCORES, which assign the participant's exact age at testing in place of assessment wave; however, this ultimately resulted in a more parsimonious model. Finally, we had intended to evaluate the growth unidirectional effects alone (i.e., with the reverse effects set to zero) before combining them into a bidirectional model; however, because of the strong bidirectional effects, these models did not converge. Thus, equivalence constraints on cross-lag regression parameters were evaluated with opposite direction parameters freely varying.

3 RESULTS

Longitudinal trajectories of raw scores on the six cognitive assessments are depicted in Supplementary Figure S1. The measurement model for creating the time-specific EF latent factors exhibited excellent fit (CFI = 1.00, SRMR = 0.027, RMSEA = 0.00, AIC = 10560). Constraining the loadings of repeated measures marginally reduced model fit ($\Delta$AIC = +1.73, LRT $p = 0.05$), though the change was not large enough to raise concerns about lack of metric invariance. DCCS was the only indicator with more than one covariance, and model fit was not reduced by constraining DCCS repeated measures
covariances ($\Delta$AIC = −7.69, LRT $p = 0.99$), so the more parsimonious model was retained. However, the EF latent factor at T1 was strongly right-skewed (skew = 1.149), and thus was fit better by a skew t distribution ($\Delta$AIC = −102.39). Factor structure and loadings are shown in Supplementary Figure S2. Thus, the remainder of models are estimated using skew t distributions; however, because these do not allow for absolute fit indices, when new non-nested models are introduced, we first report absolute fit indices assuming a normal distribution are reported first.

### 3.1 Growth in language comprehension

A latent curve model for language with free slopes fit the data well (CFI = 1.00, SRMR = 0.02, RMSEA = 0.03, AIC = 1188) and revealed a pattern of nonlinear growth with larger increases between earlier waves and smaller increases between later waves (Figure 2). Both latent factors were highly skewed (intercept skew = −1.16, slope skew = 1.15), and a skew t distribution improved model fit ($\Delta$AIC = −39.46). The means for both the latent intercept ($\mu = -0.20$, $SE = 0.04, p < 0.001$) and latent slope ($\mu = 0.852, SE = 0.05, p < 0.001$) were significant, yet neither variance was significant, nor was the covariance between the latent intercept or slope. The addition of freely varying residual variances with time-adjacent autoregression significantly improved fit ($\Delta$AIC = −3.12, LRT $p = 0.02$), and the autoregressive parameter was significant ($B = 0.17, SE = 0.076, p = 0.03$).

### 3.2 Growth in executive function

A latent curve of factors model for EF with free slopes also fit the data well (CFI = 0.98, SRMR = 0.07, RMSEA = 0.03, AIC = 10571) and was further improved with skew t distributions ($\Delta$AIC = −91.62).

The mean of latent intercept was not significant, but the mean of the latent slope was ($\mu = 2.28, SE = 0.11, p < 0.001$). The variance of each latent growth factor was significant (intercept $\psi = 0.22, SE = 0.05, p < 0.001$; slope $\psi = 1.9, SE = 0.05, p < 0.001$). The covariance between the latent intercept and slope was significant and negative ($\psi = -0.20, SE = 0.04, p < 0.001$). As with language, a pattern of nonlinear growth indicated larger increases between earlier waves and smaller increases between later waves (Figure 2). The addition of freely varying residual variances with time-adjacent autoregression did not significantly alter fit ($\Delta$AIC = +3.71, LRT $p > 0.1$), so these parameters were retained in accordance with the hypotheses. The autoregressive parameter was not significant, but it was not set to zero so that it could be evaluated in the bivariate model.

### 3.3 Joint language and executive function growth

The bivariate LCMR model fit the data well (CFI = 0.98, SRMR = 0.06, RMSEA = 0.03, AIC = 11505) and was further improved with skew t distributions ($\Delta$AIC = −33.92). Constraining the cross-lag regressions of EF on language over time to equality significantly reduced model fit ($\Delta$AIC = +73.18, LRT $p < 0.001$), and thus was not retained. Comparison of regression parameters revealed that the effect of language on EF decreased over time, such that the effect of wave 2 language on wave 3 EF was significantly smaller than the effect of wave 1 language on wave 2 EF ($\DeltaB = -0.63, p < 0.001$). There was no significant difference between the effect of wave 2 language on wave 3 EF and the effect of wave 3 language on wave 4 EF ($\DeltaB = -0.07, p = 0.50$). Because the reduction in regression strength did not systematically decrease at a constant rate, the regression parameters remained freely estimated. Constraining the cross-lag regressions of language on EF to be equal over time did not significantly alter model fit (LRT $p = 0.98$) and improved parsimonious fit ($\Delta$AIC = −3.95) so this restraint was retained. Thus, there was only one regression parameter indexing the influence of prior EF on later language. All four cross-lag regression parameters were significant (all $p < 0.005$), indicating bidirectional effects. However, from wave 1 to wave 2, the effect of language on EF was significantly stronger than the effect of EF on language ($\DeltaB = 0.80, p < 0.001$). This was marginally true from wave 2 to wave 3 ($\DeltaB = 0.17, p = 0.09$) as the effect of language on EF was reduced ($\DeltaB = -0.63, p < 0.001$). Finally, at the last transition from wave 3 to wave 4, the bidirectional effects did not significantly differ. These relationships are displayed in Figure 3.

### 3.4 Socioeconomic associations with language and executive functioning

To characterize predictors of these relationships, the bivariate model was simultaneously conditioned on participant sex and SES, which were allowed to covary. SES measures of MEd and ITN assessed in parallel models (Figure 4, upper). Both models fit the data well (MEd: CFI = 0.97, SRMR = 0.06, RMSEA = 0.04, AIC = 13306; ITN: CFI = 0.98,
FIGURE 3  Diagram representing the bivariate latent curve model with structured residuals (LCMSR) of the relationships between language (lang) and executive functioning (EF) development. Solid lines represent significant paths with associated unstandardized point estimates, and dotted lines indicate covariances that were estimated but not significant. The average age of participants (in months) is represented for each manifest variable. The measurement model for the EF latent factors is visualized in Supplementary Figure S2. Latent structured residual factors (SR) have loadings fixed to 1, and autoregressive paths are fixed to equality between waves. Cross-lagged paths are fixed to equality for language SRs regressed on EF SRs, but freely estimated for EF SRs regressed on language SRs, which decrease over time. The within-wave covariances of SR residuals are fixed to equality for the last three waves.

SRMR = 0.06, RMSEA = 0.03, AIC = 13163), and both were improved with a skew t distribution (MEd: ΔAIC = −191.57; ITN: ΔAIC = −22.60). Sex was significantly associated with the language intercept in both models (MEd: B = −0.11, p = 0.02; ITN: B = −0.17, p = 0.001), such that boys had lower initial language scores than girls. In the model with ITN, sex was also associated with the language slope (B = 0.14, p = 0.02), indicating that, controlling for ITN, boys also showed steeper language growth over time. Sex was not associated with the EF intercept or slope in either model. SES measures were significantly and positively associated with both the latent intercept for language (MEd: B = 0.08, p < 0.001; ITN: B = 0.06, p < 0.001) and the latent intercept for EF (MEd: B = 0.10, p < 0.001; ITN: B = 0.10, p < 0.001), such that higher SES was associated with higher initial scores. Neither SES measure was significantly associated with either latent slope.
FIGURE 4  Diagrams representing the relationships between language (lang) and executive functioning (EF) development conditioned on (a, c) participant sex and maternal education (MEd), or (b, c) participant sex and family income-to-needs ratio (ITN). Solid lines represent significant paths with associated unstandardized point estimates, and dotted lines indicate paths that were estimated but not significant. Models in (a) and (c) allow latent intercepts and slopes to covary, while models in (c) and (d) estimate direct paths from the language intercept to the EF intercept. Bolded lines represent the indirect paths from each SES measure through the language intercept to the EF intercept (solid bold lines) and the direct (c’) paths that are no longer significant after accounting for the indirect paths (dotted bold lines)

(all |B| < 0.03, all p > 0.14); however, removing these paths reduced model fit (MEd: ΔAIC = +33.87, LRT p < 0.001; ITN: ΔAIC = +8.18, LRT p = 0.002), suggesting that the effects of SES on both latent slopes are neither statistically different from nor statistically equivalent to zero.

3.5  Mediation of socioeconomic associations

Mediation models were then estimated to determine whether the language intercept mediated the relationship between SES measures and the EF intercept (Figure 4, lower). Even though the intercepts are derived using longitudinal information, they are both between-person measures, so the reverse mediations (i.e., whether the EF intercept mediated relationships between SES and language) were also evaluated in competing models.

For MEd, the model with an indirect path from MEd to EF through language (Figure 4c) fit significantly better than the model with undirected paths (Figure 4a, ΔAIC = −11.75), and also fit better than the model with the reverse indirect path (ΔAIC = −23.48). The directional path from the language intercept to the EF intercept (b-path) was significant (B = 1.41, p = 0.04, 99%CI = [0.41, 4.32]). Also, the indirect effect of MEd on EF through language was significant (B = 0.10, 99%CI = [0.02, 0.32]), while the path from MEd to the EF intercept (c’-path) was no longer significant (B = −0.01, p = 0.89, 99%CI = [−0.27, 0.08]).

The same pattern was observed for ITN. Specifically, the model with an indirect path from ITN to EF through language (Figure 4d) fit significantly better than the model with undirected paths (Figure 4b, ΔAIC = −3.38), and also fit better than the model with the reverse indirect path (ΔAIC = −22.97). The direct path from the language intercept to the EF intercept (b-path) was significant (B = 1.48, p = 0.01, 99%CI = [0.40, 3.30]). Also, the indirect effect of ITN on EF through language was also significant (B = 0.09, 99%CI = [0.02, 0.24]), while the path from ITN to the EF intercept (c’-path) was no longer significant (B = 0.01, p = 0.91, 99%CI = [−0.15, 0.08]).
This study investigated relationships between SES and the development of language and EF skills across preschool, from ages 3 to 5 years. Growth models revealed significant bidirectional effects across the entire age range studied. However, at the earlier ages (3 years), language comprehension more strongly predicted later EF skills than the reverse, yet at later ages (4–5 years), the bidirectional effects across domains did not significantly differ. Both MEd and ITN were associated with language and EF skills at age 3 but were not associated with growth in either cognitive domain, suggesting that children from higher-SES backgrounds exhibit higher skills than their peers from lower-SES households by the age of 3 years, but do not differ in the rates of growth from 3 to 5 years. Finally, language comprehension at the earliest timepoint mediated the associations of SES measures with initial EF skills, suggesting that SES differences in language skills may contribute to disparities in EF in early childhood. These results have both empirical and translational implications.

The finding of bidirectional influences between language and EF development is consistent with a developmental systems framework (Fischer & Bidell, 2006; Ford & Lerner, 1992), in which these skills dynamically co-develop through interaction between biological and experiential influences. Similarly, these findings fit with a biological explanation of dynamic neural plasticity, in which the neural architecture supporting language and other domains of cognitive develop flexibly, such that the functional and structural organization of cognition are dynamic across developmental time and in response to experiences (Blumstein & Amso, 2013). While these domains appear to be neurobiologically distinct in adults (Fedorenko et al., 2011), it remains an open question whether language and EF rely on distinct neural processes in childhood, or if these become differentiated throughout development.

Furthermore, the finding that language more strongly influences EF than vice versa is consistent with several prior studies of children in this age range (Fuhs & Day, 2011; Kuhn et al., 2016; Slot & von Suchodoletz, 2018; White et al., 2017). Importantly though, this difference in effect magnitude was only found at the earliest ages (3 years) before leveling off. This suggests that language skills, and specifically language comprehension, are particularly important for scaffolding early EF development. One mechanism that might explain this pattern is that language provides a hierarchical rule-based system for processing information and helps to support abstract representations of concepts or goals through the use of labels (Kuhn et al., 2016; Slot & von Suchodoletz, 2018). Notably, children undergo rapid lexical development in toddlerhood (i.e., the “word spurt” or “language explosion”) (Goldfield & Reznick, 1990) alongside robust morphological and syntactic development (Brown, 1973), both of which may serve as critical foundations for early EF development. As EF skills begin to grow more rapidly across the preschool period, the feedback loop between language and EF appears to balance out, with language continuing to support EF development, as well as varied EF skills supporting linguistic skills such as word learning from context and social communication skills.

Alternatively, it is also possible that language appears to influence EF more strongly than the reverse because the EF tasks used in the current study were administered verbally. By this account, better language skills could enhance understanding of task rules and goals, which could lead to higher performance, no matter whether the child used linguistic strategies (e.g., verbal repetition of the rules) to complete the task. However, we find this explanation unlikely since children were given multiple practice trials with feedback on each task to ensure understanding. Yet to fully rule out such an account, future research should examine relationships with EF tasks requiring little or no language to explain, such as tablet-based tasks that use visual demonstrations instead of verbal instructions.

A wealth of prior studies have observed strong associations between SES and measures of language and EF development (for reviews, see Lawson et al., 2018; Pace et al., 2017). The present finding that SES predicts the level of both language and EF skills at the beginning of the study (age three years) but not growth in these skills over time suggests that whatever influence SES is having on these aspects of cognitive development unfolds very early in life. This is consistent with other studies of preschool-aged children that have found that SES associations with both language (Farkas & Beron, 2004; Levine et al., 2020) and EF (Hackman et al., 2015; Hughes et al., 2010) emerge early, and these disparities remain relatively stable through the transition to school. Together this suggests that intervention efforts to reduce SES disparities in cognitive development may have the greatest benefit early in life when differences are first emerging.

Finally, the finding that the association of SES with early EF were mediated by early language comprehension indicates that individual differences in receptive language development in the first years of life serve as one potential mechanism contributing to SES differences in early EF. This is consistent with findings in cross-sectional studies of children ranging from 5 to 17 years of age (Catale et al., 2012; Noble et al., 2005, 2007; Pluck et al., 2020), and suggests that the mediating role of language occurs even earlier, by age 3. Given that language is often found to be the neurocognitive domain with the greatest SES gradients (Farah et al., 2006; Jednorog et al., 2012; Noble et al., 2007), it is logical that this influence may have downstream influences on other cognitive skills supported early language development.

These findings also have important practical implications. Given the significant influence of EF on academic and other life outcomes, many intervention programs have been designed to improve young children’s EF through direct cognitive training as well as modifications to their environments (for reviews, see Diamond & Lee, 2011; Scionti et al., 2019). These interventions are commonly targeted at children from lower SES backgrounds with the goal of reducing disparities. The present findings suggest that improving children’s early language skills may serve as an indirect pathway to improving EF development. A targeted way of accomplishing this would be through direct, family-based intervention programs; indeed, increases in children’s early home language environments result in commensurate improvements in both language and EF skills (Romeo et al., 2021). However, further research is needed to determine whether language experience has direct influence on EF development, or whether it indirectly scaffolds EF through
language development. Additionally, it is unclear whether direct interventions on language skills would result in long-term EF improvement, as well as the effects of multi-pronged approaches targeting the co-development of language and EF. Alternatively, interventions targeting factors that are further upstream to reduce the structural inequities that shape children's early cognitive environments (e.g., a universal basic income, expanded access to high-quality early childhood education programs) may have cascading effects on both language and EF development. Such approaches have the potential for more widespread reduction of disparities in cognitive development by targeting the source.

Despite its advances, this study has several limitations. Regarding the measures, only receptive language was assessed. However, language domains tend to be more strongly intercorrelated than EF domains (Korkman et al., 2007), and language samples at Time 1 reveal strong correlations between the included receptive measure and multiple expressive measures of lexical and morphosyntactic development ($p < 0.001$). Further, because the language task in the present study involved escalating complexity of both vocabulary and syntax, it is impossible to discern which precise language skills most facilitate early EF development. Future research should encompass more comprehensive assessments of language subdomains to reveal more precise pathways of influence. Additionally, as mentioned above, the EF tasks used were all explained verbally, and future research should aim to fully disassociate language and EF tasks to the extent possible. Regarding the methods, the mediation models suggest the possibility of causality, but these do not rule out third variables or reciprocal relationships; experimental manipulation would be needed to have greater confidence in the proposed causal pathways. Finally, SES is an imprecise index of children's proximal experiences, such as cognitive stimulation through language rich caregiver-child interaction. Thus, while this study provides evidence about the temporal relationship between language and EF development, we can only speculate on the mechanisms by which early language skills support EF development. Future studies examining children's cognitive and linguistic experience is needed to reveal the precise nature of the language-EF relationship across development.

In conclusion, this study revealed dynamic relationships between SES, language, and EF development in the preschool years. Findings suggest that SES-related differences in receptive language development, instantiated before age 3, may contribute to the well-established findings on EF development. We can only speculate on the mechanisms by which early language skills support EF development. Future studies examining children's proximal experiences, such as cognitive stimulation through language rich caregiver-child interaction. Thus, while this study provides evidence about the temporal relationship between language and EF development, we can only speculate on the mechanisms by which early language skills support EF development. Future studies examining children's cognitive and linguistic experience is needed to reveal the precise nature of the language-EF relationship across development.

ACKNOWLEDGMENTS
This research was supported by the Eunice Kennedy Shriver National Institute for Child Health and Human Development (K99-HD103873 to RRR, and R01-HD054465 to LJL) and the National Institute for Mental Health (R01-MH106482 to KAM). We thank the families who participated in this study and members of the Longitudinal Modeling Club for feedback on analytical design.

CONFLICT OF INTEREST STATEMENT
The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT
The pre-registered hypotheses/procedures, code, and analytical output for this study can be found at https://osf.io/34hf6/. Raw data is available by request to the final author.

ORCID
Rachel R. Romeo https://orcid.org/0000-0002-0315-4385

REFERENCES


