The effects of socioeconomic status on working memory in childhood are partially mediated by intersensory processing of audiovisual events in infancy

Elizabeth V. Edgar a,*, Bret Eschman b, James Torrence Todd c, Kaitlyn Testa c, Bethany Ramirez c, Lorraine E. Bahrick c,*

a Yale Child Study Center, Yale University School of Medicine, United States
b Department of Psychology, University of Tennessee at Chattanooga, United States
c Department of Psychology, Florida International University, United States

ARTICLE INFO

Keywords:
Intersensory processing
Socioeconomic status
Working memory
Executive functions
Individual differences

ABSTRACT

Socioeconomic status (SES) is a well-established predictor of individual differences in childhood language and cognitive functioning, including executive functions such as working memory. In infancy, intersensory processing—selectively attending to properties of events that are redundantly specified across the senses at the expense of non-redundant, irrelevant properties—also predicts language development. Our recent research demonstrates that individual differences in intersensory processing in infancy predict a variety of language outcomes in childhood, even after controlling for SES. However, relations among intersensory processing and cognitive outcomes such as working memory have not yet been investigated. Thus, the present study examines relations between intersensory processing in infancy and working memory in early childhood, and the role of SES in this relation. Children (N = 101) received the Multisensory Attention Assessment Protocol at 12-months to assess intersensory processing (face-voice and object-sound matching) and received the WPPSI at 36-months to assess working memory. SES was indexed by maternal education, paternal education, and income. A variety of novel findings emerged. 1) Individual differences in intersensory processing at 12-months predicted working memory at 36-months of age even after controlling for SES. 2) Individual differences in SES predicted intersensory processing at 12-months of age. 3) The well-established relation between SES and working memory was partially mediated by intersensory processing. Children from families of higher-SES have better intersensory processing skills at 12-months and this combination of factors predicts greater working memory two years later at 36-months. Together these findings reveal the role of intersensory processing in cognitive functioning.

* We thank the children and families who participated in this research. This research was supported by NIMH grant T32 MH18268 awarded to the first author and NICHD grants R01HD094803 and R01HD053776 awarded to the sixth author. This study was not preregistered. Data and study materials are not currently available online but will be available in the future.

* Correspondence to: Department of Psychology, Florida International University, Miami, FL 33199, United States.
E-mail addresses: elizabeth.edgar@yale.edu (E.V. Edgar), bahrick@fiu.edu (L.E. Bahrick).

https://doi.org/10.1016/j.infbeh.2023.101844
Received 15 December 2022; Received in revised form 26 April 2023; Accepted 19 May 2023
Available online 2 June 2023
0163-6383/Published by Elsevier Inc.
1. Introduction

Socioeconomic status (SES) is a well-established predictor of language and cognitive functioning in childhood (Hart & Risley, 1995; Mistry et al., 2008), including executive functions (EFs) such as working memory (Hackman et al., 2014; Noble et al., 2007). In turn, SES and EFs predict important outcomes such as academic achievement and socioemotional development (Barnes et al., 2022; Bradley & Corwyn, 2002; Lawson et al., 2017; Riggs et al., 2006). Intersensory processing—detecting and selectively attending to synchronously co-occurring information across two or more senses—is also thought to be a foundation for language, social, and cognitive functioning (Bahrick et al., 2020; Bahrick & Lickliter, 2012; Bremner et al., 2012). For example, recent research demonstrates that individual differences in intersensory processing in infancy predict a variety of language outcomes even after controlling for SES (Edgar et al., 2022, 2023). However, relations between intersensory processing and EFs, such as working memory, have not yet been investigated. Nor has the role of SES in this potential relation been assessed. We expect to find links between intersensory processing and working memory, given that both rely on attention control (e.g., Bahrick & Lickliter, 2014; Buss et al., 2018; for a review, see Soto-Faraco et al., 2019). For example, intersensory processing and working memory both require selective attention—selectively focusing on target information at the expense of other information. Thus, we predict that early intersensory processing skills will predict working memory later in development. We also explore the role of SES in this relation given its importance in predicting working memory.

1.1. Socioeconomic status

A large body of research demonstrates that SES predicts individual differences in a variety of EF skills (Last et al., 2018; Noble et al., 2005), particularly working memory (Farah et al., 2006; Hackman et al., 2014; Noble et al., 2007). For example, children from lower-SES families, indexed by parental education, score lower on a variety of EF tasks (e.g., go/no-go, spatial working memory, dimensional card sort) than children from higher-SES families (Noble et al., 2005). Further, children from lower-SES families also score lower on tasks assessing cognitive control and working memory compared to children from higher-SES families (Farah et al., 2006; Noble et al., 2007).

SES may also influence infant and child attention patterns. For example, children from lower-SES families showed deficits in maintaining attention during performance on a computerized behavioral task in the presence of distractors (Howse et al., 2003). Another study found that infants from higher-SES families engaged in more instances of focused attention, and maintained focused attention longer than infants from lower-SES families (Clearfield & Jedd, 2013). Thus, although it has been established that SES is predictive of working memory (for a meta-analysis, see Lawson et al., 2017) and of attention patterns, these findings come from separate studies. To our knowledge, little or no research has explored the links among SES, early attention skills to multisensory events, and EFs, particularly working memory. Assessing the links among these constructs can reveal important developmental pathways and provide a basis for identifying children at risk for impairments in working memory and the later achievements that rely on this foundation.

SES has been defined and measured in a variety of ways. The most widely used definition of SES refers to an individual’s access to financial, educational, and social resources, as well as the social positioning and privileges that are derived from access to these resources (Bradley & Corwyn, 2002; Entwisle & Astone, 1994; Pace et al., 2017). Popular indices of SES include household income, parental education, parental occupation, or some combination of the three (Bradley & Corwyn, 2002; Pace et al., 2017). In the present study, we explored three indicators of SES, maternal education, paternal education, and income, given that each component has been shown to uniquely influence developmental outcomes (e.g., Mistry et al., 2008).

1.2. Executive functions

EFs encompass a set of interrelated cognitive skills that support goal-directed behavior, (Best & Miller, 2010; Devine et al., 2019; Garon et al., 2008; Vrantsidis et al., 2019; Zelazo & Müller, 2010) including inhibition, set shifting, and working memory (Carlson, 2005; Diamond, 2013; Friedman & Miyake, 2017; Lehto et al., 2003; Miyake et al., 2000; Welsh et al., 1991; Zelazo et al., 1997). Inhibition refers to the ability to override prepotent impulses or habits (Diamond, 2013). Set shifting refers to the capacity to flexibly switch between mental states, rule sets, and tasks (Miyake et al., 2000). Working memory involves the capacity to selectively attend and briefly retain information in an accessible state, as well as monitor, manipulate, and update that information in order to engage in mental tasks (Baddeley & Hitch, 1974; Cowan, 1998; Engle et al., 1999).

Research demonstrates that inhibition, set shifting, and working memory emerge in infancy and continue to develop through early childhood (Diamond & Goldman-Rakic, 1989; Fiske & Holmboe, 2019; Garon et al., 2008; Vrantsidis et al., 2019), showing moderate stability by early childhood (Carlson et al., 2004). In early childhood, EFs predict later developmental outcomes such as socioemotional development (Blair et al., 2004; Devine & Hughes, 2014; Jahromi & Stifter, 2008; Riggs et al., 2006; Schoemaker et al., 2013), and academic achievement (Bayliss et al., 2003; Blair et al., 2015; Blair & Razza, 2007; Mishel et al., 1989). Of the three EF skills, working memory has most consistently predicted academic achievement, whereas findings regarding relations with the other skills (inhibition, set shifting) are mixed (e.g., Ahmed et al., 2019; Rose et al., 2011; though see Blair & Razza, 2007). Here, we focus on working memory, given its link with early attention skills and SES, and its important role in academic achievement (Ahmed et al., 2019; Barnes et al., 2022; Bayliss et al., 2003; Waters et al., 2021).

Researchers have emphasized the role of selective attention in the functioning of working memory (e.g., Baddeley & Hitch, 1974; for reviews, see Gazzaley & Nobre, 2012; van Ede & Nobre, 2022). Working memory requires selective attention to target information
while inhibiting attention to other competing information (Cowan, 1988; Cowan et al., 2005). For example, adults with greater selective attention skills score better on working memory tasks such as an auditory selective listening task (Conway et al., 2001) and a visual filtering task (Plebanek & Sloutsky, 2019). Relations between attention and working memory are present in infancy (for a review, see Reynolds & Romano, 2016), with at least one study demonstrating that attentional orienting and selective attention facilitate working memory (Ross-Sheehy et al., 2011). Similarly, in the case of intersensory processing, infants must selectively focus on synchronous and unitary audiovisual events while filtering out simultaneously occurring competing auditory and visual stimulation. Therefore, we expect that early selective attention skills, such as intersensory processing, should be predictive of working memory.

1.3. Intersensory processing

Intersensory processing provides a fundamental basis for guiding selective attention and perceptual development (Bahrick et al., 2020; Bahrick & Lickliter, 2012). It involves selectively attending to properties of events that are redundantly specified across the senses while ignoring non-redundantly specified, irrelevant properties. Selective attention provides the basis for what is perceived, learned, and remembered. In turn, what is perceived, learned, and remembered influences what is selectively attended to at later points in time (Bahrick & Lickliter, 2014). Intersensory redundancy (the synchronous co-occurrence of stimulation across two or more senses) is provided by most naturalistic events and is highly salient to infants. Thus, intersensory processing (the detection of intersensory redundancy) ensures that inexperienced perceivers selectively attend to patterns of stimulation that belong together and constitute unitary events (e.g., the face and voice of a person speaking) while simultaneously ignoring stimulation from unrelated events (e.g., the television; Bahrick & Hollich, 2017).

Research has just begun to demonstrate the importance of intersensory processing as a foundation for later developmental outcomes at the individual-level. Until now, there were no commonly accepted individual difference measures of intersensory processing. Without sufficiently fine-grained individual difference measures of intersensory processing, it has not been possible to identify pathways from early intersensory processing skills to later developmental outcomes or to assess predictive relations between early

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Demographic Information for the Sample (N = 101).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>49</td>
</tr>
<tr>
<td>Female</td>
<td>52</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>65</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>35</td>
</tr>
<tr>
<td>Did not disclose</td>
<td>1</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White/European-American</td>
<td>71</td>
</tr>
<tr>
<td>Black/African-American</td>
<td>16</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>4</td>
</tr>
<tr>
<td>More than 1 race</td>
<td>7</td>
</tr>
<tr>
<td>Did not disclose</td>
<td>3</td>
</tr>
<tr>
<td>SES Variables</td>
<td></td>
</tr>
<tr>
<td>Maternal Education</td>
<td></td>
</tr>
<tr>
<td>High school or equivalent</td>
<td>14</td>
</tr>
<tr>
<td>Some college</td>
<td>16</td>
</tr>
<tr>
<td>Associate’s degree</td>
<td>15</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>26</td>
</tr>
<tr>
<td>Master’s degree or higher</td>
<td>26</td>
</tr>
<tr>
<td>Did not disclose</td>
<td>4</td>
</tr>
<tr>
<td>Paternal Education</td>
<td></td>
</tr>
<tr>
<td>High school or equivalent</td>
<td>19</td>
</tr>
<tr>
<td>Some college</td>
<td>21</td>
</tr>
<tr>
<td>Associate’s degree</td>
<td>10</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>20</td>
</tr>
<tr>
<td>Master’s degree or higher</td>
<td>21</td>
</tr>
<tr>
<td>Did not disclose</td>
<td>10</td>
</tr>
<tr>
<td>Income</td>
<td></td>
</tr>
<tr>
<td>$0 – $23,494</td>
<td>10</td>
</tr>
<tr>
<td>$23,493 – $47,297</td>
<td>18</td>
</tr>
<tr>
<td>$47,298 – $75,000</td>
<td>6</td>
</tr>
<tr>
<td>$75,000 – $100,000</td>
<td>24</td>
</tr>
<tr>
<td>$100,000 +</td>
<td>34</td>
</tr>
<tr>
<td>Did not disclose</td>
<td>9</td>
</tr>
<tr>
<td>Age</td>
<td>M</td>
</tr>
<tr>
<td>12-month visit</td>
<td>12.22</td>
</tr>
<tr>
<td>36-month visit</td>
<td>36.77</td>
</tr>
</tbody>
</table>
developing skills and later outcomes. To address this need, Bahrick and colleagues developed two new individual difference measures of intersensory processing, the Multisensory Attention Assessment Protocol (MAAP; Bahrick, Todd, et al., 2018) and the Intersensory Processing Efficiency Protocol (IPEP; Bahrick, Soksa, et al., 2018). Recent research using the MAAP and IPEP has demonstrated that individual differences in intersensory processing across the first year of life predict individual differences in language development at 18, 24, and 36 months, even after controlling for other well-established predictors including parent language input and SES (Edgar et al., 2022, 2023).

In contrast, less is known about the role of intersensory processing as a predictor of individual differences in cognitive development. However, findings with group-level data have demonstrated that perception of intersensory redundancy facilitates operant learning, numerical discrimination (Jordan et al., 2008), sequence detection (Lewkowicz, 2004), and abstract rule learning (Frank et al., 2009) in human infants, and enhances memory for the maternal call in bobwhite quail chicks (Lickliter et al., 2004). Assessing this relation at the individual-level can address a host of important questions that group difference approaches are not designed to address, including relations and developmental pathways between intersensory processing and later working memory skills.

1.4. The present study

The present study examines links among intersensory processing, working memory, and SES. We hypothesized that intersensory processing at 12 months of age would predict working memory at 36 months of age. Given our prior findings that intersensory processing at 12 months predicted language even when SES was controlled, we expected it would also predict cognitive functioning, such as working memory. Further, we explored the role of SES in the relation between intersensory processing and working memory and potential relations among these three variables. By understanding the relations among intersensory processing, working memory skills, and SES, we improve our ability to identify early markers that predict challenges in later EFs.

2. Method

2.1. Participants

One-hundred and four children participated as part of a longitudinal study assessing the development of intersensory processing and cognitive, social, and language outcomes. The study, “Development of Intermodal Perception of Social and Nonsocial Events”, received IRB approval from the Social and Behavioral Review Board of Florida International University (IRB-13-0448-CR06). The present study includes data from the 12- and 36-month visits. Eighty-one children contributed data at 12 months (78%), and 80 children contributed data at 36 months (77%; for a detailed breakdown of available and missing data, see Table S1). The final sample consisted of 101 children who had data for at least one variable (SES, intersensory processing, working memory). Demographic information for the sample, including the SES variables (maternal education, paternal education, and income) can be found in Table 1. Data for the present study are currently available online at https://nyu.databrary.org/volume/1581.

2.2. Intersensory processing: MAAP

2.2.1. Apparatus and equipment

The MAAP was presented on a 116.84-cm. widescreen monitor (NEC Multisync PV61) with children seated on their caregiver’s lap approximately 101.6 cm. away. Caregivers wore black-out glasses to ensure that they were blind to the side of the screen that depicted the sound-synchronous event. An experimenter was seated behind the caregiver and child, and presented the stimuli to the widescreen monitor using a custom MatLab-based program from a second computer (Mac Pro Computer with 16 GB of RAM, a 3.33-GHz processor, and a 400-MHz graphics card).

2.2.2. Stimulus events and procedure

The MAAP (Bahrick, Todd, et al., 2018) is a three-screen video procedure (depicting three events side-by-side) that assesses multisensory attention skills to audiovisual social and nonsocial events. Social events depict women telling stories in child-directed speech, and nonsocial events depict small wooden objects being dropped into a container in an erratic temporal pattern (see Fig. 1). The MAAP consists of 12 social and 12 nonsocial trials, grouped into 4 blocks of 6 trials each. The order of presentation (social, nonsocial, social, nonsocial, or vice versa) is counterbalanced across participants. Each trial begins with a 3-second central stimulus (silent visual event depicting moving geometric shapes) followed by the onset of two 12-second lateral events (right and left sides of the three-screen display). Depending on the event condition, the lateral events portray two social (two different women each telling a different story) or two nonsocial events (two different object sets being dropped into a container). The movements and sounds of one lateral event are synchronous with the soundtrack, whereas the movements of the other lateral event are asynchronous with the soundtrack. For half of the trials, the central stimulus event (the morphing geometric shapes) is presented for the duration of the 12-second trial to provide an additional source of competing stimulation (high competition trials). For the other half of the trials, the central distractor event disappears at the onset of the lateral events (low competition trials). For an example video, visit: https://nyu.databrary.org/volume/326. Additional details regarding the MAAP procedures can be found in Bahrick, Todd, et al. (2018; pp. 2216–2217).

Trials were included in analyses if one of the two lateral events was fixated for a minimum of 250-ms. Out of the 24 trials, children met this criterion on an average of 22.59 trials (SD = 3.47; 94% of trials). However, children were fixating the central distractor event...
Infant Behavior and Development 72 (2023) 101844

2.2.3. MAAP measures

The MAAP provides measures of three multisensory attention skills: intersensory processing (indexed by the accuracy of intersensory matching), speed of shifting/disengaging, and duration of sustained attention. Speed of shifting/disengaging is the infant’s reaction time to look to a lateral event and was calculated for each trial as the latency to shift attention (in seconds) to either of the two lateral events. Sustained attention is the proportion of available looking time spent fixating the two lateral events and was calculated for each trial by dividing the total looking time to both lateral events (the synchronous and asynchronous events) by the length of the trial. The present study focused on intersensory matching on high competition trials, given that it has been shown to be a particularly strong predictor of developmental outcomes (Edgar et al., 2022). Intersensory matching is the proportion of total looking time to the sound-synchronous event (PTLT) and was calculated for each trial by dividing the looking time to the sound-synchronous event by the total looking time to the synchronous and asynchronous events (both lateral events). PTLT reflects matching based on synchrony detection, and a value greater than .50 reflects a significant preference for the sound-synchronous display. Trained observers, hidden behind the widescreen monitor with a black curtain, also viewed the child through the front facing camera. The observers coded looking time to the right, center, and left sides of the screen on a game pad in real time. A second trained observer coded visual looking time to the three screens for 54% (n = 45) of the sample at 12 months. Results of Pearson’s correlations between estimates of the primary and secondary observers indicate that interobserver reliabilities for the MAAP measures (averaged across trials) were quite high: intersensory matching: r = 0.92; speed of shifting/disengaging: r = 0.94; sustained attention: r = 0.96.

2.3. Working memory: WPPSI-IV

At 36 months of age, children received the short version of the WPPSI-IV (Weschler, 2012) appropriate for use with children 30–47 months of age. The WPPSI-IV provides a measure of Full-Scale IQ (FSIQ), and also provides three primary indices (Working Memory Index, Visual-Spatial Index, and Verbal Comprehension Index of cognitive performance). The three indices are calculated on the basis of seven subtests. The Working Memory Index is comprised of the Picture Memory and Zoo Locations subtests. The Visual-Spatial Index is comprised of the Block Design and Object Assembly subtests. The Verbal Comprehension Index is comprised of the Information and Receptive Vocabulary subtests. Given that Visual-Spatial and Verbal Comprehension indices make up the overall FSIQ measure along with the Working Memory index and they were correlated with Working Memory, we focus specifically on the Working Memory Index, but include the Visual-Spatial and Verbal Comprehension indices as covariates in the present study.

2.4. Data analysis plan

All analyses were conducted in MPlus with the Full Information Maximum Likelihood (FIML) estimator to account for missing data (see Table 2 for percent of available data for each variable). We first conducted bivariate correlations among 12-month MAAP measures (intersensory matching, speed of shifting/disengaging, sustained attention), the SES variables (maternal education, paternal education, and income), and the 36-month WPPSI-IV indices (Working Memory, Verbal Comprehension, Visual-Spatial). We then assessed whether a latent variable for SES (comprised of maternal education, paternal education, and income) could be derived, given the correlational structure of the data. Finally, we tested the relations among 12-month intersensory matching, the latent SES variable, and...
Table 2
Means (M), Standard Deviations (SD), Sample Sizes (N), and Percentages of Available Data for Intersensory Matching at 12 Months of Age and the Three Primary Indices of the WPPSI-IV at 36 Months of Age.

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>n</th>
<th>%Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersensory Matching at 12 Months</td>
<td>52.00</td>
<td>0.08</td>
<td>80</td>
<td>79%</td>
</tr>
<tr>
<td>WPPSI at 36 Months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working Memory Index</td>
<td>99.12</td>
<td>18.33</td>
<td>78</td>
<td>77%</td>
</tr>
<tr>
<td>Verbal Comprehension Index</td>
<td>102.96</td>
<td>17.09</td>
<td>80</td>
<td>79%</td>
</tr>
</tbody>
</table>

and 36-month working memory with a mediation model using structural equation modeling (SEM; with and without controlling for the other two WPPSI-IV indices – Visual-Spatial and Verbal Comprehension).

3. Results

3.1. Correlations

Descriptive statistics for all variables are located in Table 2. To assess the relations among 12-month MAAP measures (intersensory matching, speed of shifting/disenengaging, sustained attention), SES (maternal education, paternal education, and income), and 36-month WPPSI-IV indices (Working Memory, Verbal Comprehension, Visual-Spatial), we conducted Pearson’s r correlations using FIML. A matrix of these correlations can be found in Table 3. We controlled for family-wise error using a modified Bonferroni procedure (Holm, 1979). Several significant correlations among the measures were evident. First, results indicated that intersensory matching (but not sustained attention or speed of shifting/disenengaging) at 12 months, and one of the SES variables, maternal education, were significantly correlated with the Working Memory Index at 36 months, rs > 0.21 ps < 0.05. Second, intersensory matching at 12 months, and two of the three SES variables (maternal and paternal education, but not income), predicted the Verbal Comprehension Index at 12 months, and higher SES (maternal education, paternal education and/or income) predicted greater working memory and verbal comprehension at 36 months, and higher SES predicted greater intersensory matching at 12 months.

3.2. Latent SES Variable

Correlation analyses indicated that all three SES variables (maternal education, paternal education, income) were significantly correlated with one another and that the size of the correlations was large, rs > 0.50, ps < 0.001. Thus, using confirmatory factor analysis (CFA), we tested a latent SES variable comprised of the three SES variables (see Fig. 2). The CFA for the latent SES variable showed excellent model fit, $\chi^2(1) = 0.10, p = .76$ (see Table 4). The excellent fit for the latent SES variable indicates that a single latent factor, SES, accounts for the inter-related variability across the three SES variables.

3.3. A structural equation model assessing main research questions

Results of the bivariate correlation analyses indicated that greater intersensory matching at 12 months and higher SES measures were significant predictors of greater Working Memory Index scores at 36 months. Further, higher SES significantly predicted greater intersensory matching at 12 months. We next constructed and tested a structural equation model (SEM) to address our two main research questions. First, does intersensory matching at 12 months predict working memory at 36 months, holding constant SES?

---

3 We also conducted the same correlations without FIML using traditional Pearson correlations. Both approaches yielded estimates similar in magnitude.

4 We created a family of three for each WPPSI-IV outcome (Working Memory, Verbal Comprehension, Visual-Spatial) given there are three MAAP measures (intersensory matching, speed of shifting/disenengaging, sustained attention) and three indices of SES (maternal education, paternal education, income; see Table 3). Thus, the correlation with the smallest p value is compared against a critical value of $p < .017$ (0.05 / 3). If the correlation with the smallest p value is less than.017, it is declared significant. Then, the correlation with the next smallest p value is compared against a critical value of $p < .025$ (0.05 / 2), and so on.

5 The CFI was 1, indicating that our hypothesized model reduces 100% of the approximation error of the baseline model. The RMSEA was.00, 90% CI [0.00 – 0.18], indicating a.00 increase in standardized covariance residual per degree of freedom due to approximation error. The lower value of the confidence interval was ideally at the value of 0, and the interval contained the value of.00. Although the range of the confidence interval for the RMSEA is large, indicating less-than-excellent fit, all other fit indices (e.g., CFI, SRMR, actual value of the RMSEA) show excellent fit to the data. Finally, the SRMR demonstrated that the average residual correlation was 0.02, indicating good fit.
Second, we explored pathways among SES, intersensory matching, and working memory, in particular whether intersensory processing mediated the relation between SES and working memory. Our model was constructed as follows: path a: the latent SES variable (predictor) predicts intersensory matching at 12 months (mediator), path b: intersensory matching at 12 months predicts the Working Memory Index at 36 months (outcome), and path c: the latent SES variable predicts the Working Memory Index at 36 months (see Fig. 3). The model was tested in MPlus using FIML. Global fit indices are located in Table 4 and model parameters can be found in Table 5. The mediation model showed very good fit to the data, $\chi^2(4) = 4.89, p = .30$.6

First, results indicated that intersensory matching of audiovisual events at 12 months predicted working memory at 36 months, holding constant SES. Greater intersensory matching at 12 months predicted greater Working Memory Index scores at 36 months, $\beta = 0.21, SE = 0.10, p = .03$, controlling for the individual differences in the latent SES variable. However, higher SES was a marginal

---

### Table 3
Correlations Among Intersensory Matching at 12 Months of Age, the SES Variables (Maternal Education, Paternal Education, Income) the Three Primary Indices of the WPPSI-IV at 36 Months of Age (Working Memory, Visual-Spatial, and Verbal Comprehension), and the Other Two Multisensory Attention Skills Assessed by the MAAP at 12 Months of Age (Sustained Attention, Speed of Shifting/Disengaging).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intersensory Matching</th>
<th>Sustained Attention</th>
<th>Speed</th>
<th>Maternal Education</th>
<th>Paternal Education</th>
<th>Income</th>
<th>Working Memory</th>
<th>Visual Spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES Maternal Education</td>
<td>.26***</td>
<td>.09</td>
<td>-.09</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Paternal Education</td>
<td>.27***</td>
<td>.29***</td>
<td>-.33***</td>
<td>.53***</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Income</td>
<td>.04</td>
<td>.23*</td>
<td>-.26***</td>
<td>.52***</td>
<td>.50***</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WPPSI Primary Indices</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working Memory</td>
<td>.30**</td>
<td>.05</td>
<td>-.03</td>
<td>.21*</td>
<td>.15</td>
<td>.22†</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Visual-Spatial</td>
<td>.08</td>
<td>-.04</td>
<td>.21†</td>
<td>.12</td>
<td>.08</td>
<td>.14</td>
<td>.34***</td>
<td>-</td>
</tr>
<tr>
<td>Verbal Comprehension</td>
<td>.25*</td>
<td>.10</td>
<td>-.15</td>
<td>.32*</td>
<td>.22†</td>
<td>.07</td>
<td>.57***</td>
<td>.35***</td>
</tr>
</tbody>
</table>

Note: *** $p < .001$, ** $p < .01$, * $p < .05$; † did not meet significance cutoff (critical $p$ value) when controlling for familywise error.

---

Fig. 2. A Confirmatory Factor Model Depicting the Latent Variable for SES Comprised of Maternal Education, Paternal Education, and Income. Standardized Loadings are Presented Outside Parentheses, and Unstandardized Loadings are Presented Within Parentheses.

---

6 The CFI was .99, the RMSEA was .05, 90% CI [0.00 – 0.16], and the SRMR was .04.
predictor of Working Memory Index scores, $\beta = 0.24$, $SE = 0.14$, $p = .08$, controlling for individual differences in intersensory matching. Thus, given equal levels of SES, intersensory processing of audiovisual events in infancy predicts individual differences in working memory in childhood.

Second, our results indicated that intersensory matching at 12 months significantly and partially mediated the relation between SES and working memory. Higher SES significantly predicted greater intersensory matching at 12 months, $\beta = 0.27$, $SE = 0.11$, $p < .001$, which in turn, significantly predicted greater Working Memory Index scores at 36 months, $\beta = 0.21$, $SE = 0.10$, $p = .03$. All findings remained significant and similar in magnitude when controlling for the Visual-Spatial and Verbal Comprehension Indices of the WPPSI-IV (for details, see Supplement pp. 1–2 and Tables S2-S3). Inspection of path coefficients showed that a 1-unit increase in SES was related to a .27-unit increase in intersensory matching on average, and that a 1-unit increase in intersensory matching was associated with a .21-unit increase in working memory skills. The total effect of SES and intersensory processing on working memory was significant, $b = 5.36$, $SE = 2.24$, $p = .02$. Though the pathways from SES to intersensory processing (a path) and from intersensory matching to working memory (b path) were significant, the indirect effect (a times b path, bootstrapped) failed to reach significance, $b = 1.05$, $SE = 0.71$, $p = .14$. Finally, the ratio of the indirect effect to the total effects was .20, indicating that 20% of the total effect on
working memory was due to the indirect pathway from SES through intersensory matching.

4. Discussion

In this study, we examined links among SES, intersensory processing in infancy (at 12 months), and working memory skills in early childhood (at 36 months). Our prior findings demonstrated that intersensory processing in infancy predicted language outcomes in childhood (18, 24, and 36 months), even when holding constant well-established predictors including SES (Edgar et al., 2022, 2023). Here, we extend our prior findings to a new childhood cognitive outcome, working memory, an important EF, which prior research has demonstrated predicts later childhood outcomes such as academic achievement (Ahmed et al., 2019; Rose et al., 2011). Further, we explored potential relations among SES, intersensory processing, and childhood working memory skills. Results were consistent with previous research in that SES predicted working memory (Farah et al., 2006; Hackman et al., 2014; Noble et al., 2007). We also found a variety of novel relations revealing that 1) intersensory processing in infancy predicts working memory in childhood even after controlling for SES, 2) the well-established relation between SES and working memory is partially mediated by intersensory processing, and 3) individual differences in SES are related to intersensory processing.

4.1. Intersensory processing in infancy predicts working memory in childhood

Findings from our analyses demonstrate that intersensory processing (but not speed of shifting/disengaging or sustained attention) of social and nonsocial events at 12 months in the presence of competing stimulation was a significant predictor of working memory at 36 months of age, even when holding constant individual differences in SES. Thus, at 12 months, given equal levels of SES, children who showed higher levels of intersensory matching (i.e., greater detection of synchrony across auditory and visual stimulation) showed higher scores on the Working Memory Index of the WPPSI-IV at 36 months. These novel findings build on our prior research demonstrating that children who showed higher levels of intersensory processing in infancy (6 and 12 months) showed better language outcomes in childhood, including larger receptive and expressive vocabularies, as well as more frequent and diverse speech production, at 18, 24, and 36 months (Edgar et al., 2022, 2023). However, results of our prior studies were based on a somewhat different index of intersensory processing than the present study. Our prior studies demonstrated that intersensory processing of social (matching faces and voices), but not nonsocial events (matching moving objects and sounds), predicted language outcomes. In contrast, in the present study, we found that a combined index of intersensory processing of both social and nonsocial events best predicted working memory outcomes. This raises the possibility that although intersensory processing of faces and voices plays a central role in the development of language (see Edgar et al., 2022, 2023), general intersensory processing skills (i.e., across domains) may play a more important role in the development of cognitive outcomes, such as working memory.

Together, our prior and present findings demonstrate that intersensory processing in infancy predicts individual differences in both language and working memory. These findings highlight the importance of assessing intersensory processing in infancy as predictors of language and EFs alongside well-established predictors such as SES. Findings are the first to demonstrate that individual differences in intersensory processing predict individual differences in later cognitive functioning. Findings extend prior research focused on group differences which revealed that detection of intersensory redundancy facilitates performance on a number of cognitive tasks (Frank et al., 2009; Jordan et al., 2008; Kraebel, 2012; Lewkowicz, 2004).

Why might intersensory processing predict individual differences in working memory? Intersensory processing of sights and sounds requires selectively attending to properties of events that are common across visual and auditory stimulation, such as face-voice or object-sound synchrony, while at the same time ignoring irrelevant stimulation (see Bahrick et al., 1981; Bahrick & Lickliter, 2014; Stein, 2012; Talsma et al., 2010). Further, psychologists have long appreciated the role of selective attention in working memory (e.g., Baddeley & Hitch, 1974). Working memory requires selectively attending to target information while ignoring irrelevant information, holding the target information in memory, and then using that information to complete necessary tasks. For example, working memory plays an important role in performing novel sequences such as imitation play, games involving changing instructions (e.g., Simon Says), or following a story line to interpret character goals and actions. Despite their common foundation of selective attention, there has been little research investigating relations between intersensory processing and working memory. Here, we demonstrate that infants with greater intersensory processing show better working memory in early childhood.

4.2. Intersensory processing in infancy partially mediates relations between SES and working memory in childhood

Our findings demonstrate that the well-established link between SES and working memory is partially mediated by intersensory processing in infancy. Findings from our mediation model demonstrate that infants from higher-SES families showed higher levels of intersensory matching of auditory and visual stimulation (both social and nonsocial) at 12 months, and, in turn, showed greater working memory at 36 months. These findings illustrate an important cascade among SES background, intersensory processing in infancy, and later working memory skills. Our findings suggest that though children from higher-SES families may show greater working memory (relative to children from lower-SES families), those with greater intersensory processing skills show even greater working memory skills (relative to their same age higher-SES peers who have lower levels of intersensory processing). Further, it may be that enhanced intersensory processing skills in children from lower-SES families function as a “buffer” against the potentially negative impact of lower-SES backgrounds on working memory outcomes. The present findings raise the possibility that targeted interventions for improving intersensory processing in early development may lead to improved working memory outcomes.

The developmental cascade from SES to intersensory processing and later working memory highlights the important role of
intersensory processing in working memory outcomes. However, given that intersensory processing partially mediated the relation between SES and working memory, it is likely that other skills also account for the relation between SES and working memory. Some of these skills identified in prior research include infant processing speed (Rose et al., 2011), duration of orienting (Bosquet Enlow et al., 2019), and speech processing efficiency (Marchman & Fernald, 2008). These skills could also be investigated as potential mediators in this cascade.

Prior research has also demonstrated that children from higher-SES families show greater working memory skills compared to children from lower-SES families (Farah et al., 2006; Hackman et al., 2014; Noble et al., 2007). Here, however, we demonstrate that SES was a marginal predictor of working memory when controlling for individual differences in intersensory processing in infancy. Further, our supplementary analyses indicated that SES was no longer a unique predictor of working memory when holding constant intersensory processing at 12 months, as well as holding constant Verbal Comprehension and Visual-Spatial Skills from the WPPSI-IV at 36 months (for details, see Supplement Material, p. 1). This raises the possibility that SES may not be a strong predictor of EFs, such as working memory, when accounting for individual differences in other child skills, such as language and general cognitive functioning.

4.3. Relations between SES and intersensory processing in infancy

Finally, our findings are also the first to demonstrate that individual differences in SES are related to infant intersensory processing. Specifically, children from higher-SES families showed higher levels of intersensory processing of audiovisual social and nonsocial events in the presence of competing stimulation. Findings are consistent with research demonstrating that other indices of attention (e.g., sustained attention) are associated with SES (Clearfield & Jedd, 2013; Howse et al., 2003), and are the first to extend this link to SES and intersensory processing in infancy.

Why might children from higher-SES families show better intersensory processing? It may be that children from higher-SES families receive a greater amount of high-quality cognitive stimulation in their home environment. Research has shown that infants from higher-SES families have greater opportunities for face-to-face interactions with caregivers than infants from lower-SES families (Bradley & Corwyn, 2002; Shonkoff & Phillips, 2000). Face-to-face interactions may provide greater quality and quantity of language input and higher quality interactions (e.g., contingent responses from caregivers, instances of joint attention, etc.; e.g., Hirsh-Pasek et al., 2015; Vernon-Feagans et al., 2013). This may in turn foster or allow for more practice selectively attending to and perceiving audiovisual speech provided by caregivers. Also, infants from higher-SES families likely have greater resources, including a greater variety of educational toys (Bradley & Corwyn, 2002) that promote exploration across auditory, visual, and spatial modalities than children from lower-SES families. Thus, compared to infants from lower-SES families, infants from higher-SES families may have more opportunities to match auditory and visual information for both social (e.g., caregivers) and nonsocial (e.g., toys) events.

4.4. Limitations

The present study has several limitations. First, the earliest age at which working memory skills were incorporated into the present study was 36 months. Future research should assess working memory at earlier ages and test for bidirectional effects between intersensory processing and working memory. Second, the present study assessed just one EF skill, working memory, given its important role in predicting later academic achievement. It is possible that intersensory processing may predict other EF skills such as inhibition and set shifting. Our longitudinal dataset includes additional measures of EFs at older ages (e.g., 48, 60, 72 months), including inhibitory control and set shifting. Thus, in future research we will examine SES and intersensory processing as foundations of EF by including other components of EFs, relations with other cognitive outcomes (e.g., processing speed, fluid reasoning, IQ), and extending to older children.

4.5. General conclusion

In conclusion, the present study is the first to examine relations among SES, intersensory processing, and a component of EF skills, working memory. It reveals novel relations between intersensory processing in infancy and working memory in childhood. Findings demonstrate that the well-established link between SES and working memory is partially mediated by intersensory processing in infancy. Thus, better sight-sound matching in infancy is fostered by higher-SES backgrounds, and in turn, promotes better working memory skills in childhood.

Author statement

This research was conducted in accordance with APA ethical standards in the treatment of the study sample. The authors have no conflicts of interest to declare.

Data Availability

Data for the present study are currently available online at https://nyu.databrary.org/volume/1581.
Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.infbeh.2023.101844.

References


