The relationship between symbolic play and language acquisition: A meta-analytic review

Sara Quinn[a,b,*], Seamus Donnelly[a,b], Evan Kidd[a,b,c]

[a] *The Australian National University, Canberra, Australia*
[b] *ARC Centre of Excellence for the Dynamics of Language, Canberra, Australia*
[c] *Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*

A B S T R A C T

A developmental relationship between symbolic play and language has been long proposed, going as far back as the writings of Piaget and Vygotsky. In the current paper we build on recent qualitative reviews of the literature by reporting the first quantitative analysis of the relationship. We conducted a three-level meta-analysis of past studies that have investigated the relationship between symbolic play and language acquisition. Thirty-five studies ($N = 6848$) met the criteria for inclusion. Overall, we observed a significant small-to-medium association between the two domains ($r = .35$). Several moderating variables were included in the analyses, including: (i) study design (longitudinal, concurrent), (ii) the manner in which language was measured (comprehension, production), and (iii) the age at which this relationship is measured. The effect was weakly moderated by these three variables, but overall the association was robust, suggesting that symbolic play and language are closely related in development.

Introduction

Across several disciplines (e.g., psychology, education) there is a general although not uncontroversial assumption that symbolic (or 'pretend/fantasy/imaginary') play has an important role in development. Numerous classic and modern theories incorporate this form of play as a key theoretical concept in development, although each theory differs significantly on specific details (e.g., Piaget, 1951; Rakoczy, 2006, 2008; Vygotsky, 1978). In the current paper we focus on the developmental relationship between symbolic play and language acquisition. Specifically, we take stock of the existing literature that has focused on the topic, and rigorously test the strength of the relationship using quantitative meta-analytic techniques.

Symbolic play

Symbolic play reveals a child's emerging capacity for symbolic representation. Although a standard definition has eluded the field, Lillard (1993) operationalised symbolic play as "the projecting of a supposed situation onto an actual one, in the spirit of fun rather than for survival" (p. 349). Lillard identified five features differentiating play and non-play episodes, which are both necessary and sufficient for play to be considered pretend. These include: (i) a pretender, (ii) a reality (a pervasive and obvious essential component), (iii) a mental representation that is different from reality, (iv) a layering of representation over the reality, such that they exist within the same space and time (also see Bretherton, O'Connell, Shore, & Bates, 1984), and (v) an awareness on the part of

the pretender of the components of play described in (ii)–(iv). Children are typically good at distinguishing pretence and reality, suggesting that pretend scenarios are a meaningful and mostly separate context in development (Leslie, 1987; for additional discussion see Bourchier & Davis, 2002; Lillard, Pinkham, & Smith, 2010). The capacity for symbolic representation as revealed through symbolic play is argued to reflect the child's emerging cognitive competence, a view that has framed the majority of theory and research into the relationships between symbolic play and language (McCune, 1995).

Symbolic play and language

The relationship between symbolic play and language acquisition has been a particularly prominent past research topic. Both domains share many conceptual commonalities; in particular, they are both *symbolic* and *communicative* (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Werner & Kaplan, 1963). Consistent with these commonalities, many studies have identified statistically significant concurrent and longitudinal associations (i.e., symbolic play predicting language) between the two domains (e.g., Bates et al., 1979; Hall, Rumney, Holler, & Kidd, 2013; Kelly & Dale, 1989; Kirkham, Stewart, & Kidd, 2013; Lewis, Boucher, Lupton, & Watson, 2000; Lyytinen, Laakso, Poikkeus, & Rita, 1999; McCune-Nicolich, 1981; Ogura, 1991). This relationship has been observed when the two behaviors emerge around the second year of life, as well as in older children. For instance, Bates et al. (1979) reported longitudinal associations between early symbolic play and language in children aged 9–13-months. Similarly, McCune (1995) reported that symbolic play and early language development were closely coupled, with development in symbolic play predicting development in spoken language longitudinally from 8 to 24 months (see also Ogura, 1991). In older children, Kirkham, et al. (2013) reported that symbolic play and language were significantly associated in 3–4-year-old children over and above the influence of non-verbal IQ and age, suggesting that the two domains are significantly related until elementary school age.

Despite reports of positive associations between symbolic play and language acquisition, not every study has reported the effect. For instance, Bornstein, Vibert, Tal, and O'Donnell (1992) reported a concurrent association at 13-months, but not longitudinal associations from 13- to 20-months. Furthermore, Lyytinen et al. (1999) reported a longitudinal symbolic play–language association from 14- to 24-months, but in the same sample found that symbolic play at 18-months did not predict language at 24-months. In a sample of slightly older children, Shore, O'Connell, and Bates (1984) reported no longitudinal symbolic play-language association from 20 to 28 months, although Shore (1986) reported a significant concurrent association in a sample of 18–24 month-olds.

There are several possible explanations for these inconsistencies. Firstly, it may be that the effect is weak in the population, and is thus particularly sensitive to problems associated with operationalizing key concepts. This may be especially problematic for measuring symbolic play, which is both difficult to define (Sutton-Smith, 2001) and has been operationalized in many different ways. Secondly, the likelihood of observing the effect (if it exists) may have been significantly affected by sample size, which in symbolic play-language studies have been notoriously variable and often low, as has also been the case in studies that have investigated the role of symbolic play in other cognitive domains (Lillard et al., 2013). This mirrors the general tendency for developmental research to be underpowered. In a recent analysis of 12 standardised meta-analyses investigating various aspects of language development, Bergmann et al. (2018) reported that most studies test too few participants, reporting a median power of 44% (i.e., 44% of identifying a true effect if it exists, range: 6–99%). Problems with power weaken the reliability and replicability of data, and can lead to both false positives and false negatives.

Therefore, as with many cases in the field, the true strength of the relationship between symbolic play and language is difficult to determine. Such inconsistencies raise the potential problem of publication bias, which may be particularly problematic in studies of pretend play. Smith (1988, 2010), for instance, has suggested that the role of play in development has been overstated because many researchers follow a "play ethos", whereby they weigh evidence according to an *a priori* belief in the intrinsic value of the behavior. The belief that play behaviors are beneficial for development is a common theme in the psychological and biological sciences. From an evolutionary perspective, it has long been argued that play enables juveniles to hone complex skills within the relatively safe confines of the play context. Indeed, many unrelated species engage in some form of play (e.g., physical play, object play), which begs the question as to why similar behaviors would independently evolve if they had no function. Conclusive evidence in support of this general assertion is very difficult to find, potentially because play likely has multiple functions and outcomes (Bateson & Martin, 2013). However, as Smith has argued, it is dangerous to assume a developmental function for behaviors without rigorously testing them, and this may be have influenced the past literature due to the possible presence of confirmation bias.

Several recent qualitative reviews of the symbolic play literature have considered the play-language relationship. In a wide-ranging review of the role of playful learning in preschool, Hirsh-Pasek, Golinkoff, Berk, and Singer (2009) concluded that play (broadly construed) was an important context for language and literacy development. Similarly, Weisberg, Zosh, Hirsh-Pasek, and Golinkoff (2013) reviewed correlational and intervention research investigating the role of play on language development, and concluded that "play is highly beneficial to children's language skills and provides a supportive context for language learning" (p. 49). Concentrating specifically on symbolic play, Lillard et al. (2013) (including commentaries by Bergen, 2013; Walker & Gopnik, 2013; Weisberg, Hirsh-Pasek, & Golinkoff, 2013) briefly reviewed correlational (concurrent and longitudinal) and experimental intervention research investigating the relationship between symbolic play and early language.[1] They concluded that the concurrent correlational evidence for the relationship was compelling, and that there was some evidence for a longitudinal relationship where play predicted language across the first 2 years of life.

---

[1] Early language was defined as "first words and syntax" (Lillard et al., 2013, p. 18).

The current study

While qualitative reviews provide an important first step in determining the existence of a developmental relationship between symbolic play and language development, they cannot reliably estimate the strength of the effect or determine what may moderate it, and they are notably weak in addressing problems associated with small sample sizes and publication bias. In the current study we employed meta-analytic techniques to provide a cumulative statistical assessment of the symbolic play-language literature. Thirty-five ($N = 35$) correlational studies that examined the symbolic play-language relationship were included in the analysis, for a combined sample of 6848 children. Three potential moderator variables were considered: (i) study design (e.g., concurrent or longitudinal), (ii) the modality of language assessment within the studies (e.g., comprehension or production), and the age at which this relationship was measured (e.g., younger or older than 36 months).

The first moderator, study design, was included to assess concurrent and longitudinal symbolic play-language associations. These designs represent the most common design type in the past literature, and there have been some suggestions that longitudinal relationships may attenuate after early development (specifically, from age 2 years onwards, Lillard et al., 2013). Thus there are suggestions that the symbolic play–language relationship changes with age, and so age-related changes were therefore examined as the second moderator. The inclusion of age was also motivated by differences in how classic developmental theory has treated symbolic play. The Piagetian constructivist approach views symbolic play as evidence for the child's emerging capacity for symbolic representation (Belsky & Most, 1981; McCune, 1995; McCune-Nicolich, 1981; Nicolich, 1977). Piaget (1962) argued that both symbolic play and language require the capacity for symbolic understanding of objects and events, and that growth in both domains is dependent on the development of this symbolic capacity (the 'semiotic function') (Bornstein, Haynes, O'Reilly, & Painter, 1996). The suggestion is that any relationship between symbolic play and language is an epiphenomenon deriving from the development of this general symbolic capacity, specifically from 'local homologies' shared across domains (Bates et al., 1979). As such, while Piaget saw symbolic play as an important demonstration of symbolic thought, he did not regard it as a contributing factor to the development of cognitive processes such as language. Therefore, the Piagetian approach does not predict a stable positive relationship between symbolic play and language acquisition across development because the two domains are outcomes of general symbolic development. Insofar as the two domains are dependent on general symbolic development, any observable relationship between the two is predicted to be limited to the period of transition to symbolic functioning (9–36 months). In contrast, Vygotskian theory predicts a significant association across a wider developmental period. Specifically, Vygotsky (1962) argued that play is a driving force in development, providing a context in which children acquire a range of cognitive skills, including language. Thus the prediction from the Vygotskyan approach is that symbolic play and language are closely coupled throughout childhood.

The third moderating variable, modality of language assessment, aimed to resolve inconsistent results in the literature. Whereas symbolic play is always measured via production, language has been measured using tests that measure either comprehension or production, or both. For instance, Lyytinen et al. (1999) reported that symbolic play was more strongly associated with language comprehension than production in children aged 18-months. In contrast, in a sample of 3–4-year-olds, Kirkham et al. (2013) reported stronger associations between symbolic play and language production. Therefore it is unclear as to whether the symbolic play-language relationship is more strongly observed when language is measured through comprehension or production. Although studies of the symbolic play-language relationship rarely discuss methodological choice of language measurement, it is clear that language comprehension and production are related yet not equivalent constructs in development. For instance, Bates, Bretherton, and Snyder (1988) showed that the two are dissociated early in development (10–13 months), but it is likely that the two systems become more integrated across development, since psycholinguistic models typically assume that the two processes rely on common underlying representations (e.g., Chang, Dell, & Bock, 2006). Developmental differences in the degree to which measurements of language comprehension or production tap into underlying linguistic representations could be one potential reason underlying mixed past results. The use of meta-analysis allows for a more reliable assessment of whether any symbolic-play-language relationship differs according to how language is measured.

Method

Study selection

Digital databases (Proquest Psychology Journals, PsycINFO, PsycARTICLES, PsycBOOKS and Linguistics and Language Behavior Abstracts) were searched for English-language articles (peer-reviewed journals, books, and book chapters) using various combinations of the following keywords: *symbolic play, preten* (i.e., wildcard for pretend, pretence, and pretense), fantasy, sociodramatic, language (spoken, comprehension, production, receptive, expressive), object substitution, typical, development, children.* Additional studies were identified using the ancestral method (i.e., from reference lists). Finally, unpublished sources (e.g., dissertations, presentations, unpublished studies) were searched for using Google Scholar, contacting researchers in the field via personal email, the Info-CHILDES listserv, and digital databases of dissertations (e.g., ProQuest Dissertations and Theses, and Educational Resources Information Centre).

There were several criteria for study inclusion. Firstly, each study required an independent measure of symbolic play. To date, there is little consensus in the way that symbolic play ought best be operationalised (e.g., Lillard et al., 2013; Weisberg et al., 2013), and this analysis was bound to past operationalised definitions of the concept. The following definition of symbolic play was used to select the sample of studies:

the nonliteral use of objects, actions or persons (e.g., object substitution, imagining absent objects, assigning absent attributes, as in sociodramatic play).

This definition encompasses the major ways in which symbolic play has been operationalised in past studies (e.g., Bates et al., 1979; McCune, 1995; Tamis-LeMonda, Bornstein, Cyphers, Toda, & Ogino, 1992), in addition to being consistent with current standardized measures of pretence (e.g., the *Test of Pretend Play*, Lewis & Boucher, 1997). Differences in the measures of symbolic play used across different studies were accounted for in the statistical analyses using random-effects models (Field & Gillett, 2010; Rosenthal & DiMatteo, 2001). The remaining criteria for inclusion were: (ii) an independent measure of language comprehension or production (measuring vocabulary or grammatical knowledge, or both); and (iii) report of a correlational relationship, or data that would allow a correlation to be calculated between symbolic play and language measures.

Only correlational studies were included, for the following reasons. First, this is the most frequent type of study in the literature, enabling the most reliable estimate of the symbolic play-language relationship (the reliability of a meta-analysis is itself influenced by the number of effect sizes that can be included). In contrast, the number of past intervention studies is small. Note that these cannot be included in a meta-analysis of correlational data because the effect sizes are not equivalent. While past intervention studies have produced suggestive results, their results are difficult to interpret for various reasons (e.g., no statistics reported, Smilansky, 1968; problems with controls groups that complicate interpretations of the source of the influence of the play intervention on language, Levy, Schaefer, & Phelps, 1986). More recent studies have corrected many of these problems (e.g., Nicolopoulou et al., 2015; Stagnitti, Bailey, Stevenson, Reynolds, & Kidd, 2016; also see Weisberg et al., 2015); however, the evidence base for these studies is still small. Using correlational studies is also the best way to mitigate one other major problem identified in the literature on symbolic play, namely, the fact that few studies use blind experimenters (Smith, 2010). Since correlational studies collect independent measures of both symbolic play and language, the likelihood that experimenters can subtly influence the outcome of a study is reduced.

The initial exclusion criteria were: (a) participants with mental or developmental disorders, and (b) studies in which symbolic play was constructed as a context rather than a variable. Studies that did not include sufficient data to calculate effect sizes were excluded (e.g., Corrigan, 1982; Lifter & Bloom, 1989; Ogura, 1991). Two studies (Rutherford, Young, Hepburn, & Rogers, 2007; Youngblade & Dunn, 1995) did not directly report a relation between symbolic play and language. The authors of these studies were contacted to obtain this information, but the data were unavailable.

Following these search criteria a list of candidate studies were finalised for inclusion, which were then independently verified by a senior researcher (blind to study authorship). Any disagreements were resolved through discussion. A final list of 35 studies (*N* = 6848) published between 1978 and 2016 were identified as eligible for inclusion in the meta-analysis. A detailed summary of how symbolic play and language were operationalised and participant characteristics in this final candidate list of studies is provided in Tables S1 and S2.

Seventeen of the 35 studies did not directly report the native spoken language by participants, although the majority of these studies were conducted in English-speaking countries (i.e., Australia, United Kingdom, Unites States) and by inference a majority of the children were acquiring English as a main language. One study conducted in Canada reported that ten of the 30 participants in their study spoke English as their second language (Astington & Jenkins, 1995). Two American studies reported a mixed sample; one comprised of 12 Italian and 13 English children (Bates et al., 1979) and the other comprised 38 American and 40 Japanese speaking children (Tamis-LeMonda et al., 1992). Finnish was the native language spoken by all participants in the five studies conducted in Finland (Laakso, Poikkeus, Katajamäki, & Lyytinen, 1999; Lyytinen, 1983; Lyytinen, Poikkeus, & Laakso, 1997; Lyytinen, Poikkeus, & Laakso, 1999; Lyytinen, Poikkeus, Laakso, Eklund, & Lyytinen, 2001). The children in the studies were drawn from the full range of socio-economic backgrounds. The number of participants in the studies ranged from 12 to 5070, and ranged in age from 8 to 73 months.

Coding the studies

The following study characteristics were coded: (i) study design (concurrent or longitudinal), (ii) age at symbolic play and language measurement, (iii) the difference between these ages in the longitudinal studies (time span), (iv) measurement of symbolic play and the way that symbolic play was operationalised (e.g., Symbolic Play Test, object substitution, sociodramatic play), (v) language measurement (i.e., type of test used and whether the test measured comprehension and/or production), (vi) the method of measurement (e.g., naturalistic or elicited), and (vii) year of publication and publication status (e.g., peer review journal article, book chapters, unpublished data). Table 1 reports sample characteristics and effect sizes of the studies included in the meta-analysis.

Overall, three studies determined the level of symbolic play and subsequently measured language in a longitudinal design, whereas 22 studies measured symbolic play and language concurrently. Ten studies reported both concurrent and longitudinal data. In these cases, effect sizes were computed for both design types. It is important to emphasise that, since the focus is the predictive association between symbolic play and language, the effect sizes from the longitudinal studies represent symbolic play predicting language. That is, effect sizes coded as *longitudinal* exclusively refer to symbolic play predicting language development. Predictive analyses of the symbolic play-language relationship (with play predicting language and language in turn predicting play) would provide insight into how language originates and then acts to support social contexts such as symbolic play. However, this hypothesis was could not be tested, since there are not enough studies that measured both symbolic play and language longitudinally.

To examine the hypothesis derived from the Piaget's (1962) suggestion that the concurrent symbolic play-language relationship is limited to early childhood, effect sizes were also coded according to the age at which the symbolic play-language relationship was

Table 1
Effect sizes and moderators for the 35 studies.

Article	r	n	Design	Language modality	SE	Age
Acredolo and Goodwyn (1988) (Study 1)	0.53	38	C	Prod	0.14	1
Acredolo and Goodwyn (1988) (Study 2)	0.48	16	L	Prod	0.23	1
Astington and Jenkins (1995)	0.26	30	C	Both	0.18	2
Astington and Jenkins (1995)	0.32	30	C	Both	0.18	2
Astington and Jenkins (1995)	0.03	30	C	Prod	0.19	2
Bates et al. (1979)	0.41	25	C	Prod	0.19	1
Bates et al. (1979)	0.58	25	C	Prod	0.17	1
Bates et al. (1979)	0.53	25	C	Prod	0.18	1
Bates et al. (1979)	0.40	25	C	Prod	0.19	1
Bates et al. (1979)	0.45	25	C	Prod	0.19	1
Bates et al. (1979)	0.37	25	C	Prod	0.19	1
Bates et al. (1979)	0.46	25	C	Prod	0.19	1
Bates et al. (1979)	0.57	25	C	Prod	0.17	1
Bates et al. (1979)	0.62	25	C	Prod	0.16	1
Bates et al. (1979)	0.41	25	C	Prod	0.19	1
Bates et al. (1979)	0.44	25	C	Prod	0.19	1
Bates et al. (1979)	0.43	25	L	Prod	0.19	1
Bates et al. (1979)	0.57	25	L	Prod	0.17	1
Bates et al. (1979)	0.36	25	L	Prod	0.19	1
Bates et al. (1979)	0.42	25	L	Prod	0.19	1
Bates et al. (1979)	0.44	25	L	Prod	0.19	1
Bates et al. (1979)	0.44	25	L	Prod	0.19	1
Bates et al. (1979)	0.38	25	L	Prod	0.19	1
Bates et al. (1979)	0.45	25	L	Prod	0.19	1
Bates et al. (1979)	0.37	25	L	Prod	0.19	1
Bates et al. (1979)	0.42	25	L	Prod	0.19	1
Bates et al. (1979)	0.35	25	L	Prod	0.20	1
Bates et al. (1979)	0.57	25	L	Prod	0.17	1
Bates et al. (1980)	0.29	32	C	Comp	0.17	1
Bates et al. (1980)	0.02	32	C	Comp	0.18	1
Bates et al. (1980)	0.00	32	C	Comp	0.18	1
Bates et al. (1980)	0.46	32	C	Comp	0.16	1
Bates et al. (1980)	0.25	32	C	Comp	0.18	1
Bates et al. (1980)	0.16	32	C	Comp	0.18	1
Bates et al. (1980)	0.07	32	C	Comp	0.18	1
Bates et al. (1980)	0.47	32	C	Comp	0.16	1
Bates et al. (1980)	0.32	32	C	Comp	0.17	1
Bates et al. (1980)	0.21	32	C	Comp	0.18	1
Bates et al. (1980)	0.03	32	C	Comp	0.18	1
Bates et al. (1980)	0.30	32	C	Comp	0.17	1
Bates et al. (1980)	0.36	32	C	Prod	0.17	1
Bates et al. (1980)	0.51	32	C	Prod	0.16	1
Bates et al. (1980)	−0.22	32	C	Prod	0.18	1
Bates et al. (1980)	0.05	32	C	Prod	0.18	1
Bates et al. (1980)	0.40	32	C	Prod	0.17	1
Bates et al. (1980)	0.54	32	C	Prod	0.15	1
Bates et al. (1980)	−0.1	32	C	Prod	0.18	1
Bates et al. (1980)	0.05	32	C	Prod	0.18	1
Bates et al. (1980)	0.29	32	C	Prod	0.17	1
Bates et al. (1980)	0.49	32	C	Prod	0.16	1
Bates et al. (1980)	−0.28	32	C	Prod	0.18	1
Bates et al. (1980)	0.08	32	C	Prod	0.18	1
Bornstein et al. (1996)	0.10	138	C	Both	0.09	1
Bornstein et al. (1996)	0.00	138	C	Both	0.09	1
Callaghan and Rankin (2002)	0.07	16	C	Prod	0.27	1
Callaghan and Rankin (2002)	−0.23	16	C	Prod	0.26	1
Callaghan and Rankin (2002)	−0.11	16	C	Prod	0.27	1
Callaghan and Rankin (2002)	−0.49	16	C	Prod	0.23	1
Callaghan and Rankin (2002)	0.47	16	C	Comp	0.24	1
Callaghan and Rankin (2002)	0.52	16	C	Comp	0.23	1
Casby and Corte (1987)	0.90	15	C	Prod	0.12	1
Casby and Corte (1987)	0.84	15	C	Prod	0.15	1
Hall et al. (2013)	0.43	50	C	Prod	0.13	1
Hall et al. (2013)	0.26	50	C	Prod	0.14	1
Hall et al. (2013)	0.29	50	C	Prod	0.14	1
Hall et al. (2013)	0.59	50	C	Prod	0.12	1
Hall et al. (2013)	0.25	50	C	Prod	0.14	1
Hall et al. (2013)	0.07	50	C	Prod	0.14	1

Table 1 (*continued*)

Article	r	n	Design	Language modality	SE	Age
Hughes et al. (2006)	0.40	111	C	Prod	0.09	1
Hughes et al. (2006)	0.57	111	C	Prod	0.08	1
Jurkovic (1978)	0.45	26	C	Both	0.18	2
Jurkovic (1978)	0.00	26	C	Comp	0.20	2
Jurkovic (1978)	0.54	26	C	Prod	0.17	2
Kirkham et al. (2013)	0.37	60	C	Comp	0.12	2
Kirkham et al. (2013)	0.21	60	C	Comp	0.13	2
Kirkham et al. (2013)	0.42	60	C	Prod	0.12	2
Kirkham et al. (2013)	0.39	60	C	Prod	0.12	2
Kirkham et al. (2013)	0.42	60	C	Prod	0.12	2
Kirkham et al. (2013)	0.39	60	C	Prod	0.12	2
Kirkham et al. (2013)	0.56	31	L	Comp	0.18	2
Kirkham et al. (2013)	0.2	31	L	Comp	0.18	2
Kirkham et al. (2013)	0.45	31	L	Prod	0.18	2
Kirkham et al. (2013)	−0.06	31	L	Prod	0.18	2
Kirkham et al. (2013)	0.49	31	L	Prod	0.18	2
Kirkham et al. (2013)	0.27	31	L	Prod	0.18	2
Kirkham (2010)	0.41	31	C	Comp	0.17	2
Kirkham (2010)	0.08	31	C	Comp	0.19	2
Kirkham (2010)	0.64	31	C	Prod	0.14	2
Kirkham (2010)	0.26	31	C	Prod	0.18	2
Kirkham (2010)	0.46	31	C	Prod	0.16	2
Kirkham (2010)	0.22	31	C	Prod	0.18	2
Laakso et al. (1999a)	0.72	111	C	Comp	0.05	1
Laakso et al. (1999a)	0.41	111	C	Prod	0.08	1
Laakso et al. (1999a)	0.34	111	L	Comp	0.08	1
Laakso et al. (1999b)	0.14	111	C	Comp	0.09	1
Laakso et al. (1999b)	0.20	111	C	Prod	0.09	1
Laakso et al. (1999b)	0.15	111	C	Prod	0.09	1
Laakso et al. (1999b)	0.41	111	L	Comp	0.08	1
Laakso et al. (1999b)	0.26	111	L	Prod	0.09	1
Laakso et al. (1999b)	0.28	111	L	Comp	0.09	1
Lewis et al. (2000)	0.35	40	C	Comp	0.15	2
Lewis et al. (2000)	0.46	40	C	Comp	0.14	2
Lewis et al. (2000)	0.47	40	C	Prod	0.14	2
Lewis et al. (2000)	0.68	40	C	Prod	0.12	2
Lillard and Kavanaugh (2014)	0.24	57	L	Both	0.13	2
Lillard and Kavanaugh (2014)	0.18	57	L	Both	0.13	2
Lillard and Kavanaugh (2014)	0.11	58	C	Prod	0.13	1
Lyytinen (1983)	0.33	79	C	Comp	0.11	1
Lyytinen (1983)	0.45	79	C	Comp	0.10	1
Lyytinen (1983)	0.40	79	C	Comp	0.10	1
Lyytinen (1983)	0.55	79	C	Prod	0.10	1
Lyytinen (1983)	0.58	79	C	Prod	0.09	1
Lyytinen et al. (1997)	0.16	110	C	Prod	0.09	1
Lyytinen et al. (1997)	0.40	110	C	Comp	0.08	1
Lyytinen et al. (1997)	0.47	110	C	Comp	0.07	1
Lyytinen et al. (1997)	0.22	110	C	Prod	0.09	1
Lyytinen et al. (1997)	0.23	110	C	Prod	0.09	1
Lyytinen et al. (1997)	0.19	110	C	Prod	0.09	1
Lyytinen et al. (1997)	0.20	110	C	Prod	0.09	1
Lyytinen et al. (1997)	0.16	110	C	Prod	0.09	1
Lyytinen et al. (1997)	0.18	110	C	Prod	0.09	1
Lyytinen et al. (1997)	0.48	110	C	Comp	0.07	1
Lyytinen et al. (1999)	0.41	171	C	Both	0.07	1
Lyytinen et al. (1999)	0.19	171	C	Prod	0.08	1
Lyytinen et al. (1999)	0.21	171	L	Prod	0.08	1
Lyytinen et al. (2001)	0.32	94	L	Prod	0.09	1
Lyytinen et al. (2001)	0.29	94	L	Prod	0.09	1
Lyytinen et al. (2001)	0.34	94	L	Prod	0.09	1
Lyytinen et al. (2001)	0.39	94	L	Comp	0.09	1
Lyytinen et al. (2001)	0.27	94	L	Prod	0.10	2
Lyytinen et al. (2001)	0.30	94	L	Comp	0.09	2
Lyytinen et al. (2001)	0.28	94	L	Comp	0.10	2
McCune (1995)	0.45	72	C	Prod	0.11	1
McCune (1995)	0.18	72	C	Prod	0.12	1
McCune (1995)	0.48	72	C	Prod	0.10	1
McCune (1995)	0.33	72	C	Prod	0.11	1
McCune (1995)	0.41	72	C	Prod	0.11	1

Table 1 (*continued*)

Article	r	n	Design	Language modality	SE	Age
McEwen et al. (2007)	0.33	5070	C	Prod	0.01	1
Nielsen (unpublished data)	0.51	75	C	Both	0.10	1
Nielsen (unpublished data)	0.54	75	C	Both	0.10	1
Nielsen (unpublished data)	0.23	75	C	Both	0.11	1
Nielsen (unpublished data)	0.42	51	L	Prod	0.13	1
Nielsen (unpublished data)	0.42	51	L	Prod	0.13	1
Nielsen (unpublished data)	0.36	51	L	Prod	0.13	1
Nielsen and Dissanayake (2000)	0.36	31	C	Prod	0.17	2
Nielsen and Dissanayake (2000)	0.34	31	C	Prod	0.17	2
Nielsen and Dissanayake (2000)	0.41	31	C	Prod	0.17	2
Nielsen and Dissanayake (2000)	0.19	31	C	Prod	0.18	2
Nielsen and Dissanayake (2000)	0.23	31	C	Prod	0.18	2
Nielsen and Dissanayake (2000)	−0.03	31	C	Prod	0.19	2
Nielsen and Dissanayake (2000)	−0.4	31	C	Prod	0.17	2
O'Reilly et al. (1997) (Study 1)	0.60	30	C	Comp	0.15	1
O'Reilly et al. (1997) (Study 1)	0.40	30	C	Comp	0.17	1
O'Reilly et al. (1997) (Study 1)	0.41	28	L	Comp	0.18	1
O'Reilly et al. (1997) (Study 2)	0.57	27	C	Comp	0.16	1
O'Reilly et al. (1997) (Study 2)	0.32	27	C	Comp	0.19	1
O'Reilly et al. (1997) (Study 2)	0.26	27	C	Comp	0.19	1
Pellegrini et al. (1991)	0.45	12	L	Prod	0.28	2
Pellegrini et al. (1991)	0.44	12	L	Prod	0.28	2
Pellegrini et al. (1991)	0.63	12	L	Prod	0.25	2
Pellegrini et al. (1991)	0.62	12	L	Prod	0.25	2
Shore (1986)	0.49	30	C	Prod	0.16	1
Shore (1986)	0.41	30	C	Prod	0.17	1
Shore (1986)	0.31	30	C	Prod	0.18	1
Shore (1986)	0.41	30	C	Prod	0.17	1
Shore (1986)	0.44	30	C	Prod	0.17	1
Smith and Jones (2011)	0.29	63	C	Prod	0.12	1
Stagnitti et al. (2016)	0.37	28	C	Comp	0.18	2
Stagnitti et al. (2016)	0.27	28	C	Comp	0.19	2
Stagnitti et al. (2016)	0.47	28	C	Comp	0.17	2
Stagnitti et al. (2016)	−0.06	28	C	Comp	0.20	2
Stagnitti et al. (2016)	0.03	28	C	Comp	0.20	2
Stagnitti et al. (2016)	0.19	28	C	Comp	0.19	2
Stagnitti et al. (2016)	0.38	28	L	Comp	0.18	2
Stagnitti et al. (2016)	0.43	28	L	Comp	0.18	2
Stagnitti et al. (2016)	0.59	28	L	Comp	0.16	2
Stagnitti et al. (2016)	0.33	28	L	Comp	0.18	2
Stagnitti et al. (2016)	0.29	28	L	Comp	0.19	2
Stagnitti et al. (2016)	0.31	28	L	Comp	0.19	2
Tamis-LeMonda et al. (1992)	0.47	78	C	Comp	0.10	1
Tamis-LeMonda and Bornstein (1994)	0.33	41	C	Comp	0.15	1
Tamis-LeMonda and Bornstein (1994)	0.10	41	C	Prod	0.16	1
Tamis-LeMonda and Bornstein (1994)	0.51	41	C	Prod	0.14	1
Tamis-LeMonda and Bornstein (1994)	0.24	41	C	Prod	0.16	1
Tamis-LeMonda and Bornstein (1994)	0.17	41	C	Prod	0.16	1
Tamis-LeMonda and Bornstein (1994)	0.03	41	L	Comp	0.16	1
Tamis-LeMonda and Bornstein (1994)	0.42	41	L	Prod	0.15	1
Tamis-LeMonda and Bornstein (1994)	0.43	41	L	Prod	0.14	1
Tamis-LeMonda and Bornstein (1994)	0.07	41	L	Prod	0.16	1
Tamis-LeMonda and Bornstein (1994)	0.07	41	L	Prod	0.16	1
Tamis-LeMonda and Bornstein (1990)	0.42	43	C	Comp	0.14	1
Tamis-LeMonda and Bornstein (1990)	0.34	43	C	Comp	0.15	1
Tamis-LeMonda and Bornstein (1990)	−0.02	43	C	Prod	0.16	1
Tamis-LeMonda and Bornstein (1990)	−0.06	43	C	Prod	0.16	1
Youngblade and Dunn (1995)	0.50	50	C	Prod	0.13	1
Youngblade and Dunn (1995)	0.53	50	C	Prod	0.12	1

Note. C = Concurrent, L = Longitudinal, Comp = Comprehension, P = Production, Both = Comprehension and Production, Age 1 = < 36 mths, Age 2 = > 36 mths.

measured. Age was categorised into two groups: effect sizes derived from studies examining the relationship between symbolic play and language in children who were (i) 36 months or younger (k = 137) or (ii) older than 36 months (k = 59). Any overlap in this arbitrary age range was resolved by classifying the study into the age range that represented the majority of the sample. For example, as ages ranged from 12 and 72 months of age in the study conducted by Lewis et al. (2000), it was assigned to the "greater than 36 months" group. However, in the case of the Callaghan and Rankin (2002) study, as the sample of 28–42 month-old children fell

evenly between the age-range groups, it was classified into the lower age-range group.

Calculation of effect sizes

When study designs differ, the calculated effect sizes are likely to describe different population parameters (Morris & DeShon, 2002). For example, in a between-groups design (e.g., language compared across high pretence vs low pretence groups), the effect size is a description of the difference in the outcome variable between the two groups relative to the variability within the groups. In contrast, in a single-group longitudinal design (e.g., correlational analysis of data from the same individual measured at two different times) the effect size describes the change in the outcome variable (e.g., measure of language) relative to the variability of change scores (e.g., measure of symbolic play), and the effect size is derived from standard deviations of the change score. As such, effect sizes are heterogeneous when derived from studies using different designs. Therefore, in this analysis, effect sizes were transformed to the same metric (Pearson's *r*). Effect sizes were calculated performing as few transformations as possible to ensure that metrics and meaning remained as close to original data as possible. For example:

- Acredolo and Goodwyn (1988) interpreted a strong negative effect size to mean "the larger the number of object gestures a child developed, the younger the child was when the 10-word milestone was reached" (p. 461). For the purpose of the meta-analysis, this effect size was converted to a positive value to indicate that children demonstrating greater proficiency in symbolic play also showed greater proficiency in spoken language.
- Shore, O'Connell, and Bates (1984) reported the mean and maximum sequence lengths in language and (symbolic play) gesture and completed multiple regression analyses, but did not report correlations or corresponding effect sizes (p. 876). As multiple regression does not isolate the contribution of a single variable to the outcome of another variable, for the purpose of running the meta-analysis, Pearson's *r* was calculated for the mean length of utterance and the mean length of gestural sequence at both 20 and 28 months. These measures, as described by the authors, were determined to be the most reliable indicators of the child's proficiency in both symbolic play and language.
- In a cross-sectional design, Jurkovic (1978) compared the linguistic maturity of children aged between 52 and 66 months of age who exhibited high or low levels of free play, and found that children in the 'high play group' also had higher verbal production scores (as measured by the verbal expression subscale of the Illinois Test of Psycholinguistic abilities). A significant *F* test was reported with this finding, with 1 degree of freedom (e.g., $F(1, 44) = 17.88$, $p < .001$). Using the equation, $r = \sqrt{\frac{F}{F + df_{error}}}$, *r* was calculated to equal .54.

Positive effect sizes indicated that children demonstrating greater proficiency in symbolic play also showed greater proficiency in spoken language. Negative effect sizes indicated that children with greater proficiency in symbolic play exhibited comparatively lower proficiency in language.

Results

In the present analysis, 31 of the 35 studies contributed multiple effect sizes (median number of effect sizes per study = 4). Because effect sizes from the same study might be more similar than effect sizes from different studies, we used three-level meta-analysis, which extends random effects meta-analysis by including random effects at the level of study.[2] While this approach can accommodate correlated effect sizes within studies, it cannot accommodate the correlated sampling errors within studies that are due to effect sizes being calculated on the same participants (Cheung, 2015). Therefore, all confidence intervals and standard errors for test statistics used cluster robust standard errors, which correct standard errors for dependence between observations whose covariance structure is unknown (Hedges, Tipton & Johnson, 2010). Because of the correction to the standard errors, test statistics follow *t* sampling distributions, rather than normal distributions. All meta-analyses were conducted using the *metafor* package (Viechtbauer, 2010) in R (R Development Core Team, 2016).

The first model estimated the overall effect size for all studies, including those that reported on dependent measures containing both production and comprehension data. The overall effect size was small-to-medium and highly significant ($r = .35$, $t(33) = 13.91$, $p < .001$, $CI = [.30, .40]$). The significant effect size is unlikely to be due to publication bias. First, consider the funnel plot in Fig. 1, which plots each individual effect size on the *x*-axis against the effect size standard error on the *y*-axis. Funnel plots assume that studies with low variance and with large sample sizes are likely to be published irrespective of the statistical significance of the published effect, whereas studies with small sample sizes and high variance are only likely to be published if the reported effects are statistically significant. Therefore, if effect sizes from each study are symmetrically distributed around the estimated average effect size, publication bias is not present. In contrast, publication bias is indicated if low-variance effect sizes are near, or below, the estimated average effect size and high-variance effect sizes appear only above the estimated average effect size. Fig. 1 shows a symmetrical distribution of effects sizes, and therefore suggests an absence of publication bias. Second, Rosenthal's (1991, 1995) failsafe N indicated that 89,156 studies averaging null effects would need to be added to the data set in order to render the average

---

[2] One set of studies reported multiple effect sizes from different subsets of the same sample, the Jyvaskyla Longitudinal Study of Dyslexia (Lyytinen et al., 1997, 1999; Lyytinen, Poikkeus, Laakso, Eklund, & Lyytinen, 2001; Laakso et al., 1999; Laasko et al., 2000). Because these were different studies, we treated these samples as independent, which we consider a limitation of our study.
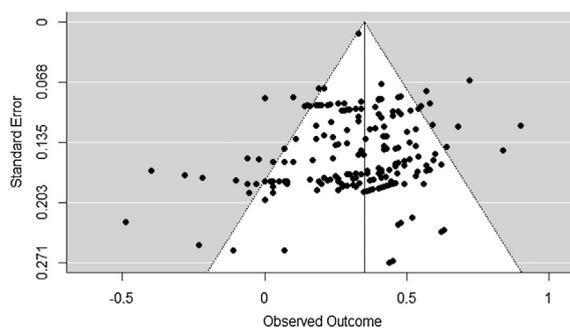
Fig. 1. Funnel plot of all effect sizes ($k = 196$).

effect size non-significant. While this number is probably an overestimate, as it does not account for the dependence between observed effect sizes, together with the funnel plot, it suggests a lack of publication bias.

Fig. 2 represents the effect size magnitudes from each study in a forest plot. As can be seen, there was a great deal of variability in effect sizes, which was confirmed by a significant Q statistic [$Q(195) = 403.99, p < .001$], suggesting significant heterogeneity.[3] To examine the impact of various moderators, several meta-regressions with moderator variables were fit. Since language modality (comprehension versus production) was one of our moderator variables, we excluded from these analyses those effect sizes where the dependent measure included a composite language measure that combined comprehension and production. The total number of effect sizes included was $k = 185$.

Models 1, 2 and 3 considered only the main effects of each of the predictor variables. Model 1 tested the effect of study design (longitudinal vs concurrent), which did not improve fit relative to a null model with no predictors ($F(1, 31) = .00, p = .948$). Model 2 tested the effect of age, and did not improve fit relative to the null model ($F(1, 31) = .54, p = .470$). Model 3 tested the effect of language modality, and also did not improve fit relative to the null model ($F(1, 31) = 1.07, p = .308$). Overall, Models 1–3 suggest that each potential moderating variable does not, on its own, influence the observed relationship between symbolic play and language.

Models 4, 5 and 6 included each of the two-way interactions between each pair of predictor variables. Model 4 included the interaction between design and modality, which did not significantly improve model fit ($F(1, 29) = .02, p = .883$). Model 5 included the interaction between design and age, which did not significantly improve fit ($F(1, 29) = .74, p = .397$). Model 6 included the interaction between age and modality, which did significantly improve fit ($F(1, 29) = 5.97, p = .021$).

To clarify this interaction, the effects of modality were tested for studies with younger and older participants separately. Among studies with younger participants, effect sizes on production tasks were numerically smaller than those on comprehension tasks, though this effect was not significant ($b = -.09, t(22) = 1.77, p = .091$). Among studies with older participants, effect sizes were numerically larger on production tasks, though this was not significant ($b = .06, t(8) = -1.31, p = .226$). We further examined the effect of age for production and comprehension effect sizes separately. For production tasks, effect sizes were numerically larger for older children than younger children, though this difference did not approach significance ($b = .01, t(27) = .19, p = .849$), and for comprehension tasks, effect sizes were numerically smaller for older children, though this difference did not approach significance ($b = -.07, t(15) = -1.51, p = .150$). Thus, while we observed a significant interaction between age and modality, follow-up analyses suggest that the interaction was driven by marginally different results across the two age groups that trended in the opposite direction.

Finally, Model 7 included a three-way interaction between age, design and modality. The three-way interaction significantly improved model fit ($F(1, 25) = 8.40, p = .008$). Just as a two-way interaction can be conceived of as a difference in mean differences, a three-way interaction can be conceived of as a difference in two-way interactions. Therefore, in order to examine the nature of this interaction, we fit several models testing for two-way interactions, first splitting the data by age groups, then by modality and then by design. We first split the data set by age group and tested for the interaction between design and modality in each age group separately. This interaction was non-significant in both younger and older samples ($F(1, 20) = .78, p = .39$, and $F(1, 6) = 3.94$, $p = .094$, respectively). We then split the data by language modality and tested for the interaction between age and design. This interaction was non-significant amongst the production and comprehension effect sizes ($F(1, 13) = 1.81, p = .201$, and ($F(1, 25) = 1.76, p = .197$, respectively). We then split the data by design and tested for the interaction between modality and age. This interaction was significant among effect sizes calculated on concurrent studies ($F(1, 26), = 5.17, p = .006$) and non-significant among effect sizes from longitudinal studies ($F(1, 8) = .00, p = .978$). Therefore, the three-way interaction appears to qualify the two-way interaction detected in Model 6 by restricting it to concurrent studies. The effect of modality was examined for younger and older participants in concurrent studies separately. As was the case with Model 6, the effect of production was numerically positive for the young children and numerically negative for the older children; however, it was non-significant for both ($b = -.11, t(20) = -1.72, p = .110$ and $b = .10, t(6) = 1.97, p = .097$, respectively). This is confirmed by Table 2, which shows the effect sizes and confidence intervals for each cell in the three-way interaction, with numerically larger comprehension than production effect

---

[3] This Q statistic is based on a model without cluster robust variance estimation as the Q statistic is not available in the output of the robust function.
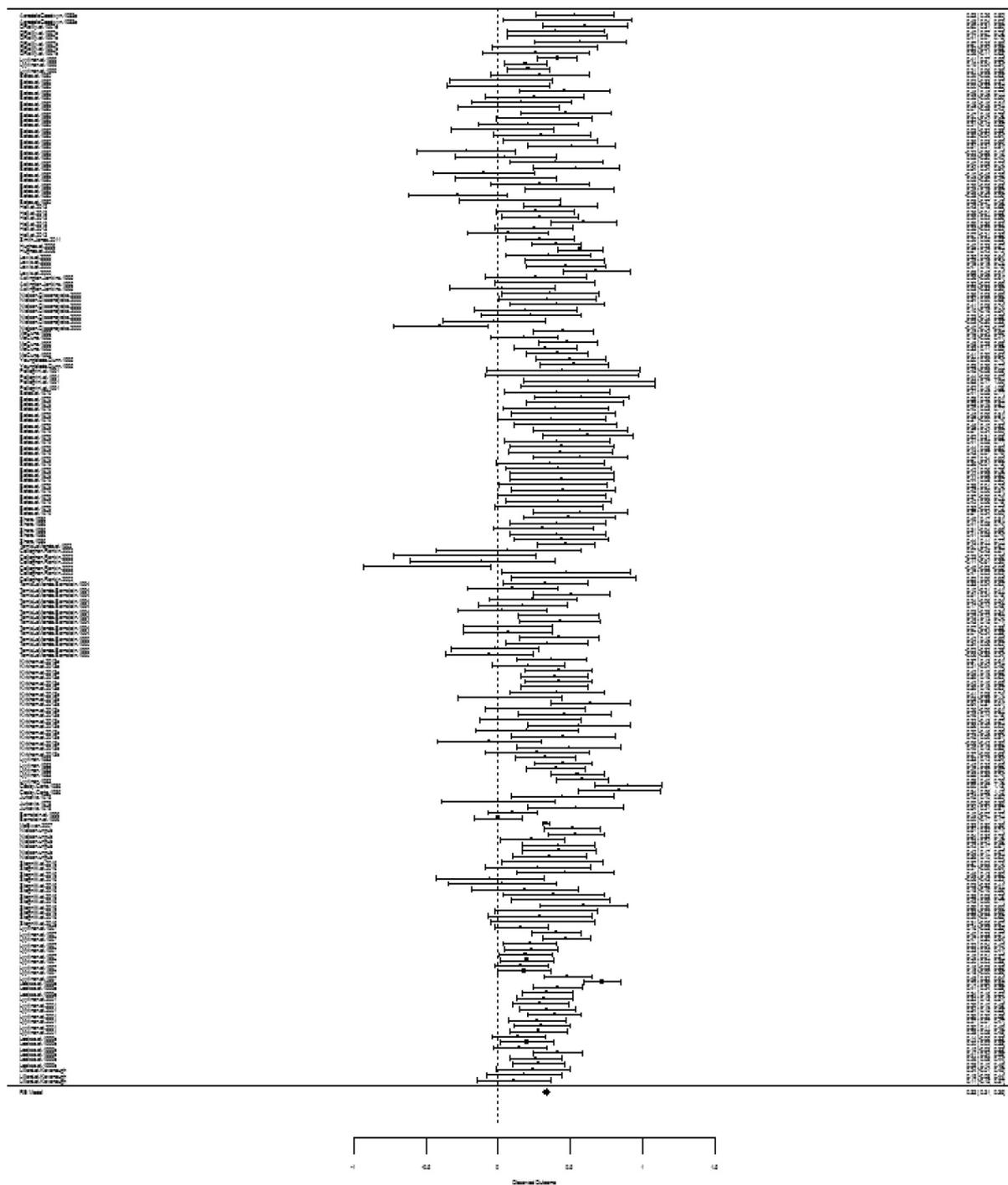
Fig. 2. Forest plot of all effect sizes (*k* = 196).

sizes for younger children in concurrent studies and numerically smaller comprehension than production effect sizes for older children in the concurrent studies.

Discussion

In the current study we used meta-analytic techniques to estimate the strength of the developmental relationship between symbolic play and language from infancy to early childhood. Although past qualitative reviews of the literature exist (Hirsh-Pasek

Table 2
Estimated effect sizes and confidence intervals by study design, language modality, and age.

	≤ 36 mths		> 36 mths	
	N	r	n	r
Concurrent-comprehension	28	.41 (.29, .54)	16	.29 (.19, .39)
Concurrent-production	72	.33 (.23, .43)	19	.36 (.16, .56)
Longitudinal-comprehension	6	.34 (.26, .41)	10	.35 (.17, .53)
Longitudinal-production	25	.33 (.24, .42)	9	.36 (.01, .70)

Notes. 95% Confidence intervals are in parentheses.

et al., 2009; Lillard et al., 2013; Weisberg, Zosh, et al., 2013), our quantitative assessment allows confidence in the data not possible with qualitative reviews because it directly addresses problems commonly associated with research on this topic – small sample sizes and the possibility of publication bias. To this effect, we found a significant small-to-medium association between symbolic play and language acquisition. This relationship was largely consistent regardless of study design (concurrent or longitudinal), the modality in which language was measured (comprehension or production), and the age at which this relationship is measured in early childhood. We can be further confident in the attested relationship because: (i) it was observed despite considerable variability in how individual studies operationalised both symbolic play and language, and (ii) there was little evidence of publication bias.

On the basis of these data we can therefore conclude that there are significant concurrent and longitudinal associations between symbolic play and language across early development (i.e., 1–6 years). The presence of significant heterogeneity justified the exploration of moderators, which might have helped us explain the nature of this relationship. We observed a significant three-way interaction between age, language modality, and study design type, which was driven by small age-related differences in the strength of the symbolic play-language relationship when language was measured through comprehension or production in concurrent designs. Although the interaction was significant, follow-up analyses revealed that the differences across age groups were relatively minor. While there is no obvious theoretical explanation for why symbolic play would be more strongly related to comprehension than production in young children but the opposite in older children, there are possible practical modality-based methodological explanations.

The interactions were driven by a difference in effect sizes between symbolic play and concurrent language comprehension in young (i.e., < 3 years) versus older children (i.e., > 3 years) (the effect sizes for language production remained relatively stable). One potential explanation for this effect derives from challenges in measuring comprehension versus production. Language abilities can be difficult to estimate in young children through comprehension. When measured via parental report (e.g., via the MacArthur-Bates CDI, Fenson et al., 2007), comprehension estimates are likely to be less accurate than production estimates because estimating comprehension requires an inference from behavior (i.e., that a child responds behaviorally to language) whereas production simply requires that the child says the word or structure (i.e., an observed behavior) (Tomasello & Mervis, 1994). When measured through behavioral tasks, comprehension requires children to produce overt responses, for instance, selecting a picture from an array. Such measures have tasks demands that may add error to estimates of language ability. Our suggestion here is that the difficulties in estimating language proficiency through comprehension could have contributed to more variability in the estimation of effect sizes, leading to the interaction. Importantly, we emphasise that our follow up analyses revealed no significant differences across age. Thus we caution against drawing any substantial conclusions from the result, beyond the general observation that the effect size between symbolic play and language production is less variable.

This one result notwithstanding, one notable result was the lack of developmental change in the symbolic play-language relationship. That is, we found evidence for the attested relationship across the full range of ages included in our analyses. This result is inconsistent with the Piagetian prediction that the symbolic play-language relationship should attenuate across development (Piaget, 1962). In contrast, it is consistent with Vygotsky's (1962) suggestion that symbolic play and language are closely coupled throughout childhood. However, it is important to note that these data fall short of demonstrating that symbolic play has a key determining role in language development. Instead, the data simply suggest that symbolic play and language continue to be associated around the age when children enter elementary school (5–6 years). Thus we can conclude that the relationship between symbolic play and language is attested beyond 3 years, and that the weight of evidence suggests that the strength of the relationship does not significantly change at least up until the age of 6 years, the age range represented in the studies on which we report.

A clear limitation of this study was that additional potential mediating variables were not included in the analyses (e.g., non-verbal IQ, maternal sensitivity, socio-economic status), which may have accounted for some variance in the symbolic play-language relationship. Their absence in this analysis was for good reason. Many past studies have not included these control variables, and some studies that did find significant symbolic play-language associations after controlling for some of these factors could not be integrated into the meta-analysis because the statistics provided were not available in the appropriate metric (e.g., SES, e.g., Jurkovic, 1978; Lyytinen et al., 2001; general cognitive development, e.g. Lewis et al., 2000; Lyytinen et al., 1999; Shore, 1986).

While these variables may contribute to language acquisition, the available evidence suggests that they would not reduce the symbolic play-language relationship to zero. One obvious third variable that could mediate the symbolic play–language relationship is non-verbal IQ. However, there appears good evidence to suggest that the relationship holds even after controlling for this variable. For instance, Kirkham et al. (2013) reported significant concurrent associations between symbolic play and two measures of language in children aged 3–4–years after controlling for non-verbal IQ. Lyytinen et al. (1999) found that symbolic play and language at

14 months significantly and independently predicted general cognitive development at 2 years, as measured by the Bayley Scales of Infant Development (Bayley, 1993). Similarly, variables related to socio-economic status (SES), such as parental education, affect both symbolic play and language development (Fein & Stork, 1981; Hoff, 2003). However, Lyytinen et al. (1999, 2001) reported significant longitudinal relations between symbolic play and language after controlling for parental education levels.

Correlational designs are an important part of developmental research, but third variables are always problematic. Therefore it is important for future studies to scrutinize the symbolic play–language relationship while controlling for any variables that may share variance with the two.

Outstanding issues

Overall, these results resolve any ambiguity around the presence and strength of the symbolic play-language relationship across infancy and early childhood, and provide a foundation for research to establish which factors influence the nature of this robust relationship. One primary factor that could influence the symbolic play-language relationship is the way in which both domains are measured. Most studies included in these analyses used measures of solitary symbolic play elicited under laboratory conditions. If symbolic play provides a conducive *context* for language learning, then ecologically valid measures of symbolic play may be more appropriate. In infancy, the most common form of symbolic play is social, with infants typically spending the majority of their symbolic play time interacting with their primary caregiver (Haight & Miller, 1993). Language acquisition is a similarly interactive process. Therefore understanding the social dynamics of the symbolic play–language relationship is one important avenue for future research.

One possibility is that symbolic play provides a particularly fertile context that fosters skills important for language development. For instance, symbolic play appears to elicit greater amounts of joint attention than comparable non-symbolic activities (Lillard & Witherington, 2004; Quinn & Kidd, 2018). Joint attention is arguably an essential component of early language development (Tomasello, 2003), and thus one potential contributing factor in the early symbolic play–language coupling could be the increased time children spend in joint attentional frames during symbolic play. Increased time spent in joint attention no doubt increases the chances of children learning from the input because it increases the quantity of language to which they directly attend, but symbolic play may also result in *qualitatively* different input. Quinn (2016) studied child-directed speech in infant-caregiver dyads during symbolic and functional play, and found that symbolic play contained significantly greater proportion of questions and conversational turns in comparison to functional play, both of which have been shown to predict language development (Romeo, 2018; Rowe, Leech, & Cabrera, 2016; Romeo et al., 2018). The increase in these variables likely reflects the inherent ambiguity present in play, where the identity of objects must be negotiated (e.g., in this context the block is a car, Rakoczy, 2006). This challenging context appears to also elicit more sophisticated language from children (e.g., Fekonja, Umek, & Kranjc, 2005; for discussion see Weisberg, Zosh, et al., 2013), although more controlled studies that isolate these effects to symbolic play as opposed to generic play contexts would be useful. Overall, the social dynamics of symbolic play and how they relate to language warrant further attention.

There are additional general issues associated with meta-analysis that deserve discussion. The technique provides a mostly objective method of evaluating the presence of a cumulative effect in the population. However, combining studies that use different measures for the same concept has been criticised because the results may mean different things (however slight the difference may be) (Rosenthal & DiMatteo, 2001). This is relevant to both independent variables in this study, since a number of different measures were used to measure these concepts (as evident in Table S1). With respect to symbolic play, a very strict definition was adhered to in order to ensure comparability across studies. With respect to language, the final study list included tests of comprehension and production of both vocabulary and grammar, which depended on both the age of the children and the type of test used in each individual study. Few studies separated the influence of symbolic play on vocabulary and grammar separately, which removed any option of isolating an analysis to examine a specific relationship to these sub-components of language. This is a matter for future research. Suffice to say, vocabulary and grammar are tightly coupled during acquisition (e.g., Bates & Goodman, 1997; Dixon & Marchman, 2007; Hoff, Quinn, & Giguere, 2017), making it possible that symbolic play is related to both.

Finally, these analyses were limited to correlational studies. Training studies of symbolic play and language exist, and consistent with our findings, have often reported positive effects of symbolic play on language (e.g., Baumer, Ferholt, & Lecusay, 2005; Levy et al., 1986; Smilansky, 1968). However, some of these studies have suffered methodological problems (e.g., lack of control groups, failure to control the role and the quality of adult intervention, lack of blind experimenters, Lillard et al., 2013). Although the past research has suffered from methodological shortcomings, we suggest that the results of the current meta-analysis show promise for future efforts aimed at modernising this evidential base. The current study has identified a robust relationship between symbolic play and language in development and has done so using rigorous statistical techniques, allaying any doubt that this association exists. Future studies can dissect and scrutinise the relationship further using a mixture of correlational and training studies.

Conclusion

The results of the current meta-analysis investigating the relationship between symbolic play and language suggests a robust association between the two domains in preschool children. The results provide a largely unbiased and theoretically neutral basis for future research. In practical terms, an accurate understanding of the size of the symbolic play-language association will allow *a priori* power analyses, which enables the calculation of adequate sample sizes. This is an important contribution, since many past studies have lacked adequate power. In theoretical terms, the field can move beyond debating the existence of the relationship and begin to design studies that scrutinise this relationship.

ARTICLE IN PRESS

S. Quinn et al.   Developmental Review xxx (xxxx) xxx–xxx

Acknowledgements

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.dr.2018.05.005.

ReferencesReferences marked with an asterisk indicate studies included in the meta-analysis

*Acredolo, L., & Goodwyn, S. (1988). Symbolic gesturing in normal infants. *Child Development, 59*(2), 450–466. http://dx.doi.org/10.1111/j.1467-8624.1988.tb01480.x.

*Astington, J. W., & Jenkins, J. M. (1995). Theory of mind development and social understanding. *Cognition & Emotion, 9*(2–3), 151–165. http://dx.doi.org/10.1080/02699939508409006.

*Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy.* New York, NY: Academic Press.

Bates, E., Bretherton, I., & Snyder, L. S. (1988). *From first words to grammar: Individual differences and dissociable mechanisms.* Cambridge: Cambridge University Press.

*Bates, E., Bretherton, I., Synder, L., Shore, C., & Volterra, V. (1980). Vocal and gestural symbols at 13 months. *Merrill-Palmer Quarterly, 26*(4), 407–423.

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes, 12*(5–6), 507–584. http://dx.doi.org/10.1080/016909697386628.

Bateson, P., & Martin, P. (2013). *Play, playfulness, creativity and innovation.* Cambridge: Cambridge University Press.

Baumer, S., Ferholt, B., & Lecusay, R. (2005). Promoting narrative competence through adult-child joint pretense: Lessons from the Scandinavian educational practice of playworld. *Cognitive Development, 20*, 576–590. http://dx.doi.org/10.1016/j.cogdev.2005.08.003.

Bayley, N. (1993). *Bayley scales of infant development* (2nd ed.). San Antonio, TX: The Psychological Corporation, Harcourt Brace.

Belsky, J., & Most, R. K. (1981). From exploration to play: A cross-sectional study of infant free play behavior. *Developmental Psychology, 17*(5), 630–639. http://dx.doi.org/10.1037/0012-1649.17.5.630.

Bergen, D. (2013). Does pretend play matter? Searching for evidence: Comment on Lillard et al. (2013). *Psychological Bulletin, 139*(1), 45. http://dx.doi.org/10.1037/a0030246.

Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M. B., Frank, M. C., & Cristia, A. (2018). Promoting replicability in developmental research through meta-analyses: Insights from language acquisition research. *Child Development.* http://dx.doi.org/10.1111/cdev.13079.

*Bornstein, M. H., Vibbert, M., Tal, J., & O'Donnell, K. (1992). Toddler language and play in the second year: Stability, covariation and influences of parenting. *First Language, 12*(36), 323–338. http://dx.doi.org/10.1177/014272379201203607.

*Bornstein, M. H., Haynes, M. O., O'Reilly, A. W., & Painter, K. M. (1996). Solitary and collaborative pretense play in early childhood: Sources of individual variation in the development of representational competence. *Child Development, 67*(6), 2910–2929. http://dx.doi.org/10.1111/j.1467-8624.1996.tb01895.x.

Bourchier, A., & Davis, A. (2002). Children's understanding of the pretence–reality distinction: A review of current theory and evidence. *Developmental Science, 5*(4), 397–413. http://dx.doi.org/10.1111/1467-7687.00236_1.

Bretherton, I., O'Connell, B., Shore, C., & Bates, E. (1984). The effect of contextual variation on symbolic play development from 20 to 28 months. *Symbolic Play, 271–298.* http://dx.doi.org/10.1016/b978-0-12-132680-7.50014-4.

*Callaghan, T. C., & Rankin, M. P. (2002). Emergence of graphic symbol functioning and the question of domain specificity: A longitudinal training study. *Child Development, 73*(2), 359–376. http://dx.doi.org/10.1111/1467-8624.00412.

*Casby, M. W., & Corte, M. D. (1987). Symbolic play performance and early language development. *Journal of Psycholinguistic Research, 16*(1), 31–42 Retrieved from < https://search-proquest-com.virtual.anu.edu.au/docview/1300148636 > .

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234–272. http://dx.doi.org/10.1037/0033-295X.113.2.234.

Cheung, M. W. (2015). *Meta-analysis: A structural equation modeling approach* (1st ed.). New York: John Wiley & Sons, Incorporated.

Corrigan, R. (1982). The control of animate and inanimate components in pretend play and language. *Child Development, 53*(5), 1343–1353. http://dx.doi.org/10.1111/j.1467-8624.1982.tb04174.x.

Dixon, J. A., & Marchman, V. A. (2007). Grammar and the lexicon: Developmental ordering in language acquisition. *Child Development, 78*(1), 190–212. http://dx.doi.org/10.1111/j.1467-8624.2007.00992.x.

Fekonja, U., Umek, L., & Kranjc, S. (2005). Free play and other daily preschool activities as a context for child's language development. *Studia Psychologica, 47*(2), 103–117 Retrieved from < http://search.proquest.com/docview/220156292/fulltextPDF?accountid=8330 > .

Fein, G. G., & Stork, L. (1981). Sociodramatic play: Social class effects in integrated preschool classrooms. *Journal of Applied Developmental Psychology, 2*(3), 267–279. http://dx.doi.org/10.1016/0193-3973(81)90006-X.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates communicative development inventories – User's guide and technical manual* (2nd ed.). Baltimore: Paul Brookes.

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology, 63*(3), 665–694. http://dx.doi.org/10.1348/000711010x502733.

Haight, W. L., & Miller, P. J. (1993). *Pretending at home: Early development in a sociocultural context.* Albany, NY, US: State University of New York Press.

*Hall, S., Rumney, L., Holler, J., & Kidd, E. (2013). Associations among play, gesture, and early spoken language acquisition. *First Language, 33*, 294–312. http://dx.doi.org/10.1177/0142723713487618.

Hedges, L., Tipton, E., & Johnson, M. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65. http://dx.doi.org/10.1002/jrsm.5.

Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E., & Singer, D. G. (2009). *A mandate for playful learning in preschool: Applying the scientific evidence.* New York: Oxford University Press, USA.

Hoff, E. (2003). The specificity of environmental influence: Socio-economic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368–1378. http://dx.doi.org/10.1111/1467-8624.00612.

Hoff, E., Quinn, J. M., & Giguere, D. (2017). What explains the correlation between growth in vocabulary and grammar? New evidence from latent chance score analyses of simultaneous bilingual development. *Developmental Science, 21*(2), http://dx.doi.org/10.1111/desc.12536.

*Hughes, C., Fujisawa, K. K., Ensor, R., Lecce, S., & Marfleet, R. (2006). Cooperation and conversations about the mind: A study of individual differences in 2-year-olds and their siblings. *British Journal of Developmental Psychology, 24*(1), 53–72. http://dx.doi.org/10.1348/026151005x82893.

*Jurkovic, G. J. (1978). Relation of psycholinguistic development to imaginative play of disadvantaged preschool children. *Psychology in the Schools, 15*(4), 560–564. http://dx.doi.org/10.1002/1520-6807(197810)15:4<560::aid-pits2310150420>3.0.co;2-3.

Kelly, C. A., & Dale, P. S. (1989). Cognitive skills associated with the onset of multiword utterances. *Journal of Speech, Language, and Hearing Research, 32*(3), 645–656. http://dx.doi.org/10.1044/jshr.3203.645.

*Kirkham, J. A. (2010). The relationship between development in the graphic, language and symbolic play domains during early and middle childhood. Unpublished PhD Thesis The University of Manchester.

*Kirkham, J., Stewart, A., & Kidd, E. (2013). Concurrent and longitudinal relationships between development in graphic, language and symbolic play domains from the fourth to the fifth year. *Infant and Child Development, 22*(3), 297–319. http://dx.doi.org/10.1002/icd.1786.

Laakso, M., Poikkeus, A., Katajamäki, J., & Lyytinen, P. (1999a). Early intentional communication as a predictor of language development in young toddlers. *First Language, 19*(56 part 2), 207–231. http://dx.doi.org/10.1177/014272379901905604.

Laaskso, M., Poikkeus, A., Eklund, K., & Lyytinen, P. (1999b). Social interaction behaviors and symbolic play competence as predictors of language development and their associations with material attention-directing strategies. *Infant Behavior and Development, 22*(4), 541–556. http://dx.doi.org/10.1016/S0163-6383(00)00022-9.

Leslie, A. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review, 94*(4), 412–426. http://dx.doi.org/10.1037//0033-295x.94.4.412.

Levy, A. K., Schaefer, L., & Phelps, P. C. (1986). Increasing preschool effectiveness: Enhancing the language abilities of 3- and 4-year-old children through planned sociodramatic play. *Early Childhood Research Quarterly, 1*(2), 133–140. http://dx.doi.org/10.1016/0885-2006(86)90024-4.

*Lewis, V., & Boucher, J. (1997). *The test of pretend play.* London: Psychological Corporation.

Lewis, V., Boucher, J., Lupton, L., & Watson, S. (2000). Relationships between symbolic play, functional play, verbal and non-verbal ability in young children. *International Journal of Language & Communication Disorders, 35*(1), 117. http://dx.doi.org/10.1080/136828200247287.

Lifter, K., & Bloom, L. (1989). Object knowledge and the emergence of language. *Infant Behavior and Development, 12*(4), 395–423. http://dx.doi.org/10.1016/0163-6383(89)90023-4.

Lillard, A. S. (1993). Pretend play skills and the child's theory of mind. *Child Development, 64*(2), 348–371. http://dx.doi.org/10.2307/1131255.

*Lillard, A. S., & Kavanaugh, R. D. (2014). The contribution of symbolic skills to the development of an explicit theory of mind. *Child Development, 85*(4), 1535–1551. http://dx.doi.org/10.2307/1131255.

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2013). The impact of pretend play on children's development: A review of the evidence. *Psychological Bulletin, 139*(1), 1. http://dx.doi.org/10.1037/a0029321.

Lillard, A., Pinkham, A. M., & Smith, E. (2010). Pretend play and cognitive development. In U. Goswami (Ed.). *The Wiley-Blackwell handbook of childhood cognitive development* (pp. 285–311). (2nd ed.). Oxford, UK: Wiley-Blackwell. http://dx.doi.org/10.1002/9781444325485.ch11.

Lillard, A. S., & Witherington, D. C. (2004). Mothers' behavior modifications during pretense and their possible signal value for toddlers. *Developmental Psychology, 40*(1), 95–113. http://dx.doi.org/10.1037/0012-1649.40.1.95.

*Lyytinen, P. (July-August, 1983). Symbolic play and early language. Paper presented at the biennial meeting of the International Society for the Study of Behavioral Development. Munich, West Germany.

*Lyytinen, P., Laakso, M., Poikkeus, A., & Rita, N. (1999). The development and predictive relations of play and language across the second year. *Scandinavian Journal of Psychology, 40*, 177–186. http://dx.doi.org/10.1111/1467-9450.00115.

*Lyytinen, P., Poikkeus, A., & Laakso, M. (1997). Language and symbolic play in toddlers. *International Journal of Behavioral Development, 21*(2), 289–302. http://dx.doi.org/10.1080/016502597384875.

*Lyytinen, P., Poikkeus, A.-M., Laakso, M.-L., Eklund, K., & Lyytinen, H. (2001). Language development and symbolic play in children with and without familial risk for dyslexia. *Journal of Speech, Language, and Hearing Research, 44*(4), 873–885. http://dx.doi.org/10.1044/1092-4388(2001/070).

McCune-Nicolich, L. (1981). Toward symbolic functioning: Structure of early pretend games and potential parallels with language. *Child Development, 52*(3), 785–797. http://dx.doi.org/10.1111/j.1467-8624.1981.tb03115.x.

*McCune, L. (1995). A normative study of representational play at the transition to language. *Developmental Psychology, 31*(2), 198–206. http://dx.doi.org/10.1037/0012-1649.31.2.198.

*McEwen, F., Happe, F., Bolton, P., Rijsdijk, F., Ronald, A., Dworzynski, K., & Plomin, R. (2007). Origins of individual differences in imitation: Links with language, pretend play, and socially insightful behavior in two-year-old twins. *Child Development, 78*(2), 474–492. http://dx.doi.org/10.1111/j.1467-8624.2007.01010.x.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105–125. http://dx.doi.org/10.1037/1082-989X.7.1.105.

Nicolopoulou, A., Cortina, K. S., Ilgaz, H., Cates, C. B., & de Sá, A. B. (2015). Using a narrative- and play-based activity to promote low-income preschoolers' oral language, emergent literacy, and social competence. *Early Childhood Research Quarterly, 31*, 147–162. http://dx.doi.org/10.1016/j.ecresq.2015.01.006.

Nicolich, L. M. (1977). Beyond sensorimotor intelligence: Assessment of symbolic maturity through analysis of pretend play. Retrieved from *Merrill-Palmer Quarterly of Behavior and Development, 23*(2), 89–99.

*Nielsen, M., & Dissanayake, C. (2000). The investigation of pretend play, mental state terms and false belief understanding: In search of a metarepresentational link. *The British Journal of Developmental Psychology, 18*, 609–624. http://dx.doi.org/10.1348/026151000165887.

*O'Reilly, A. W., Painter, K. M., & Bornstein, M. H. (1997). Relations between language and symbolic gesture development in early childhood. *Cognitive Development, 12*(2), 185–197. http://dx.doi.org/10.1016/S0885-2014%2897%2990012-5.

Ogura, T. (1991). A longitudinal study of the relationship between early language development and play development. *Journal of Child Language, 18*(2), 273–294. http://dx.doi.org/10.1017/S0305000900011065.

*Pellegrini, A. D., Galda, L., Dresden, J., & Cox, S. (1991). A longitudinal study of the predictive relations among symbolic play, linguistic verbs, and early literacy. *Research in the Teaching of English, 25*(2), 219–235.

Piaget, J. (1951). The child's conception of the world (J. Tomlinson & A. Tomlinson, Trans.). London: Routledge & K. Paul.

Piaget, J. (1962). *Play, dreams, and imitation in childhood.* New York: Norton.

Quinn, S. (2016). Learning to play and playing to learn: the role of symbolic play in language acquisition. Unpublished PhD thesis: The Australian National University.

Quinn, S., & Kidd, E. (in press). Symbolic play promotes non-verbal communicative exchange in infant-caregiver dyads. *British Journal of Developmental Psychology.*

R Development Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from <https://www.r-project.org/>.

Rakoczy, H. (2006). Pretend play and the development of collective intentionality. *Cognitive Systems Research, 7*(2–3), 113–127. http://dx.doi.org/10.1016/j.cogsys.2005.11.008.

Rakoczy, H. (2008). Pretence as individual and collective intentionality. *Mind & Language, 23*(5), 499–517. http://dx.doi.org/10.1111/j.1468-0017.2008.00357.x.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science.* http://dx.doi.org/10.1177/0956797617742725.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (rev. ed.). Thousand Oaks, CA, US: Sage Publications Inc.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*(2), 183–192. http://dx.doi.org/10.1037/0033-2909.118.2.183.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*(1), 59–82. http://dx.doi.org/10.1146/annurev.psych.52.1.59.

Rowe, M., Leech, K., & Cabrera, N. (2016). Going beyond input quantity: *Wh*-questions matter for toddlers' language and cognitive development. *Cognitive Science, 41*, 162–179. http://dx.doi.org/10.1111/cogs.12349.

Rutherford, M. D., Young, G. S., Hepburn, S., & Rogers, S. J. (2007). A longitudinal study of pretend play in autism. *Journal of Autism and Developmental Disorders, 37*, 1024–1039. http://dx.doi.org/10.1007/s10803-006-0240-9.

Shore, C., O'Connell, B., & Bates, E. (1984). First sentences in language and symbolic play. *Developmental Psychology, 20*(5), 872–880. http://dx.doi.org/10.1037/0012-1649.20.5.872.

*Shore, C. (1986). Combinatorial play, conceptual development, and early multiword speech. *Developmental Psychology, 22*(2), 184–190. http://dx.doi.org/10.1037/0012-1649.22.2.184.

Smilansky, S. (1968). *The effects of sociodramatic play on disadvantaged preschool children.* New York: Wiley and Sons Inc.

Smith, P. K. (1988). Children's play and its role in early development: A re-evaluation of the "play ethos". In A. D. Pellegrini (Ed.). *Psychological bases for early education* (pp. 207–226). New York, NY: Wiley.

Smith, P. K. (2010). *Children and play: Understanding children's worlds.* Chicester: Wiley-Blackwell.

ARTICLE IN PRESS

S. Quinn et al.   Developmental Review xxx (xxxx) xxx–xxx

*Smith, L. B., & Jones, S. S. (2011). Symbolic play connects to language through visual object recognition. *Developmental Science, 14*(5), 1142–1149. http://dx.doi.org/10.1111/j.1467-7687.2011.01065.x.

*Stagnitti, K., Bailey, A., Hudspeth Stevenson, E., Reynolds, E., & Kidd, E. (2016). An investigation into the effect of play-based instruction on the development of play skills and oral language. *Journal of Early Childhood Research, 14*(4), 389–406. http://dx.doi.org/10.1177/1476718X15579741.

Sutton-Smith, B. (2001). *The ambiguity of play.* Cambridge, MA, US: Harvard University Press.

*Tamis-LeMonda, C. S., Bornstein, M. H., Cyphers, L., Toda, S., & Ogino, M. (1992). Language and play at one year: A comparison of toddlers and mothers in the United States and Japan. *International Journal of Behavioral Development, 15*(1), 19–42. http://dx.doi.org/10.1177/016502549201500102.

*Tamis-LeMonda, C. S., & Bornstein, M. H. (1990). Language, play, and attention at one year. *Infant Behavior and Development, 13*(1), 85–98. http://dx.doi.org/10.1016/0163-6383(90)90007-u.

*Tamis-LeMonda, C. S., & Bornstein, M. H. (1994). Specificity in mother-toddler language-play relations across the second year. *Developmental Psychology, 30*(2), 283–292. http://dx.doi.org/10.1037/0012-1649.30.2.283.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, Mass: Harvard University Press.

Tomasello, M., & Mervis, C. (1994). The instrument is great, but measuring comprehension is still a problem. *Monographs of the Society for Research in Child Development, 59*(5), 174–179. http://dx.doi.org/10.1111/j.1540-5834.1994.tb00186.x.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48 Retrieved from < https://doaj.org/article/4a1fc553f6324f54951f17ec3a02c080 > .

Vygotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Walker, C. M., & Gopnik, A. (2013). Pretense and possibility-A theoretical proposal about the effects of pretend play on development: Comment on Lillard et al. (2013). *Psychological Bulletin, 139*(1), 40. http://dx.doi.org/10.1037/a0030151.

Weisberg, D. S., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Embracing complexity: Rethinking the relation between play and learning: Comment on Lillard et al. (2013). *Psychological Bulletin, 139*(1), 35–39. http://dx.doi.org/10.1037/a0030077.

Weisberg, D. S., Ilgaz, H., Hirsh-Pasek, K., Golinkoff, R., Nicolopoulou, A., & Dickinson, D. K. (2015). Shovels and swords: How realistic and fantastical themes affect children's word learning. *Cognitive Development, 35*, 1–14. http://dx.doi.org/10.1016/j.cogdev.2014.11.001.

Weisberg, D. S., Zosh, J. M., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Talking it up: Play, language development, and the role of adult support. *American Journal of Play, 6*, 39–54.

Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic developmental approach to language and the expression of thought.* New York, NY: John Wiley & Sons.

*Youngblade, L. M., & Dunn, J. (1995). Individual differences in young children's pretend play with mother and sibling: Links to relationships and understanding of other people's feelings and beliefs. *Child Development, 66*(5), 1472–1492. http://dx.doi.org/10.1111/j.1467-8624.1995.tb00946.x.