# Play it High, Play it Low
## Examining the Reliability and Validity of a New Observation Tool to Measure Children's Make-Believe Play

•

Carrie Germeroth, Elena Bodrova,
Crystal Day-Hess, Jane Barker, Julie Sarama,
Douglas H. Clements, and Carolyn Layzer

The authors consider mature make-believe play a critical component of childhood that helps children develop new skills and learn to communicate. They argue that, although theoretical accounts of play have emphasized the importance of make-believe play for children to achieve social and academic competence, the absence of a reliable and valid measure of children's mature make-believe play has hampered the evaluation of such claims. They seek to address this shortcoming with a review of the psychometric characteristics of existing assessments and with their findings from a new assessment using the Mature Play Observation Tool (MPOT), which they administered during a multiyear longitudinal study of twenty-six early-childhood classrooms. They found that children in classrooms scoring well on the MPOT better perform such skills as self-regulation, literacy, and numeracy. **Key words:** make-believe play; Mature Play Observation Tool (MPOT); scaffolding; skills assessment

## Introduction

Educators vigorously debate how to balance play with formal classroom instruction in early-childhood education. For many years, play received minimal attention in child development books (Pellegrini 2010) and diminished attention in educational settings (Hirsh-Pasek et al. 2009). A renewed interest in play has led educators to develop a number of play-based preschool interventions intended to support self-regulation and learning, and subsequent research has begun to investigate the relationship between play and the classroom.

Following the seminal work of Vygotsky (1967), developmental theorists and classroom-based practitioners frequently posited that make-believe play supports children's acquisition of social and self-regulatory skills in early childhood (e.g., Bateson 2000; Bodrova and Leong 2007; Brown 2009; Bruner 1976; Diamond 2007; Erikson,1972; Freud 2003; Leslie 1987; Piaget 1962; Saifer 2009). According to Vygotskians, make-believe play, when it reaches its mature state, fosters self-regulation and provides the basis for other activities or interactions that in turn foster the learning of symbolic and emotional thinking, spoken language, and the beginnings of literacy. From this theoretical perspective, mature make-believe play becomes a critical driver of learning in childhood that provides an opportunity for children to push their individual "developmental edge" (Elkonin 1978).

During mature make-believe play, children create imaginary situations, take on explicit roles (using the language and rules of the roles), and use objects symbolically (Bodrova and Leong 2011; Elkonin 1978). Thus, play affords children opportunities to project an internal, mental representation on the external environment (Leslie 1987; Lillard 1993) and to adopt roles and responsibilities that extend beyond their typical daily activities and represent everyday objects or environments in nonconventional ways. For example, when pretending to be a shopkeeper, a child might engage in new forms of cognitive and social interaction (counting money, for example, or interacting with a pretend patron); develop and maintain internal goals in the presence of conflicting environmental signals (inhibiting the use of a playmates' given name and instead using a pretend name, for example, or using playing cards as money and ignoring their typical use); and practice perspective taking (considering, for example, what another child might do in an imagined scenario).

Several interventions have attempted to support self-regulation and early literacy through play. Although some of these interventions have yielded benefits to child self-regulation and socioemotional skills (Han et al. 2010; Toub et al. 2018; Thibodeau et al. 2016), others have not. One factor that may complicate studies of preschool play is the absence of instruments to measure the quality of mature play in preschool classrooms. Although past work has defined, measured, and espoused the effect of play on education (Pellegrini 2010) and developed taxonomies for categorizing types of play (e.g., Burghardt 2010), no instruments have been developed to evaluate the quality of children's make-believe play in typical classroom settings. In the absence of valid, reliable measures of play, it is difficult to determine whether curricular interventions supported predicted

changes in child play. For this reason, most extant studies of classroom play interventions have not tested for proximal effects on child behavior, which are critical precursors to change in child self-regulation.

## Defining Mature Make-Believe Play

With such an important role attributed by Vygotskians to make-believe play in child development, we think it important to determine if there is a specific level of play that needs to be reached for it to become beneficial. Elaborating on Vygotsky's insights about the nature of play, Elkonin (1978, 2005a, 2005b) introduced the idea of mature play, claiming that only this kind of play can be a source of development in early childhood. Elkonin defined mature play (he used such terms as "advanced play" or "fully developed play") as a "unique form of children's activity, the subject of which is the adult—his work and the system of his relationships with others" (Elkonin 2005a, 19) thus distinguishing this form of play from other playful activities in which children engage. Although Vygotsky himself never used such terms as mature or advanced, the play vignettes in his writings seem to describe play we would consider fairly advanced. Based on the works of Vygotsky and Elkonin, as well as the work of their students, it is possible to identify several components of mature play (Bodrova and Leong 2007).

According to Vygotskian theory, make-believe play, when it reaches its mature state, fosters self-regulation and provides the basis for other activities or interactions that, in turn, foster the learning of symbolic and emotional thinking, spoken language, and the beginnings of literacy. Mature make-believe play may afford children opportunities to function within their zones of proximal development (ZPD), where newly developing skills can be completed but only with external support (Barnett, et al. 2008; Berk, Mann, and Ogan 2006; Bodrova and Leong 2007; Diamond and Lee 2011; Karpov 2005; Vygotsky 1978, 1987). These observations have been supported by experimental research. For example, Manuilenko (1975) found that children asked to act as a "lookout" in a pretend play scenario remained at their posts and did not move for a longer period of time than did children who were simply asked to stand still. Notably, the gap between play and nonplay performance was strongest in five-year-old children (assumed to be at the peak of mature play), and weaker for three-year-old children (who were still developing advanced

forms of play) and for seven-year-old children (who were less dependent on external support for behavioral regulation).

## Mature Make-Believe Play and Child Outcomes

Theoretical claims about the importance of play have been mixed. They are most often supported by correlational findings that children who actively participate in make-believe play during preschool and early elementary years display better language and social skills. For example, improvements in children's vocabulary size and syntax development have been linked to the frequency and quality of their play (e.g., Bergen 2002; Christie and Roskos 2006; Lyytinen et al. 1999; McCune 1995; Neuman and Roskos 1992; Roskos and Neuman 1998). Theory of mind (TOM), or the understanding of others' mental states, is also predicted by play behaviors—children who more often engage in fantasy-oriented play become better at discerning whether another individual entertains false beliefs about an event or circumstance—even after researchers control for potential confounding factors such as age and intelligence (Taylor and Carlson 1997). Similar patterns have been observed between play and broader social competencies (Lindsey and Colwell 2003, 2013; Colwell and Lindsey 2005).

Despite strong theoretical claims and a large pool of suggestive, correlational evidence, surprisingly few studies have causally linked make-believe play to child outcomes, and those that do are frequently subject to methodological limitations. In a recent meta-analysis of the play literature, Lillard and colleagues (2013) reviewed support for the claim that play drives improvements in a number of theoretically relevant domains, including language, emotion regulation, theory of mind, social skills, and executive functions (EF) (or the cognitive control processes that regulate thought and action in support of goal-directed behavior). For several domains, including language and emotion regulation, the authors concluded that research was insufficient to differentiate whether pretend play causally influences skill acquisition. They also found insufficient evidence to determine whether make-believe play was one of many activities promoting skill development in young children (i.e., an equifinal role) or was instead an epiphenomenal activity representing a by-product of other drivers of developmental change. For other emerging skills, such as executive functions and social skills, evidence supporting a uniquely causal role proved less convincing than evidence supporting either equifinal or ephiphenomal explanations.

In each domain, interpretation of the effects of play was complicated by the methodological limitations of existing studies, including differing definitions of play and a paucity of well-controlled empirical investigations.

One potential limitation of extant play studies, which may complicate attempts to identify causal relationships between play and developmental outcomes, is that these studies have not focused on children's mature make-believe play, which has been theorized to provide unique benefits to children's developing executive functions and self-regulation. Thus, development of a reliable, classroom-based measure of make-believe play will benefit both research and practice. First, robust observational instruments will allow us to measure better the effects of intervention on mature play and the extent to which high-quality play may produce social, self-regulatory, and academic benefits in children. Second, such an assessment can be used to determine how the quality of play might be improved in early-childhood classrooms, yielding benefits to classroom practice, where play is increasingly used to meet subject-specific instructional objectives.

### Measuring Mature Make-Believe Play

If we find defining play complicated, we find measuring it just as complicated. The observational methods of studying children's play have been influenced by ethological methods developed in the 1930s. The direct observation of play had been put aside during the midnineteenth century in favor of methods relying on questionnaires, interviews, and tests in controlled laboratory settings. More recently, studies of children's play again widely use observational methods. Most observational studies of children's play tend to be nonparticipant, meaning that the observer stands apart from those they observe with minimal to no interaction (Smith 2010). Some participant studies, however, have been used in the study of play, primarily with the observer combining observation of and interview with the subject to inquire about play behaviors. Direct observation provides more valid data than other methods such as interviews or questionnaires. Furthermore, settings such as classrooms, day care centers, and individual homes where children's play typically occurs has been seen as having an ecological advantage that might make the results more representative of real life compared to a controlled or laboratory setting.

Smith (2010) suggests conducting observational studies in a systematic way that can be objectively replicated requires several elements: theoretical presuppositions, pilot work, and familiarization; the development of categori-

Figure 1. Comparison of MPOT to Existing Play Assessments

| Measure | Age (mo.) | Setting/ Play Partner | Child Play Constructs | | | | | Teacher/Adult Facilitation Constructs | Reliability evidence | Validity evidence |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Social | Cognitive | | | Other | | | |
| | | | Interaction | Objects | Actions/ Agents | Roles/ Schemes | | | | |
| Make-Believe Play Observational Tool (MPOT) | 36–60 | Naturalistic (classroom) | | | | | | Center management; Make-believe play time; Teacher Intervention | Inter-rater (α > .9) | Predictive |
| Smilansky Scale for the Evaluation of Dramatic and Sociodramatic Play (SSEDSP; 1968/1990) | -- | Peers | With others (4 levels) | Make believe (4 levels) | Make believe (4 levels) | Imitative (4 levels) | Persistence in role (4 levels) Communication (4 levels) | -- | No .73-.92[a] .84[b] | Construct |
| Symbolic Play Test (SPT; 1976) | 12–36 | One-on-one with assessor | | Symbolic reasoning using realistic objects | | | | -- | Yes .52-.92[c] N = 137 | Concur (RDLS) |
| Howes Peer Play Scale (HPPS; 1980) | 18–43 | Class/Peers | Parallel; Parallel aware; Simple social | Complementary and reciprocal Contingent social | | | | -- | No .87-.93[a] N = 52 | Construct |
| Play Observation Scale (POS; 1989) | 24–72 | Class/Peers | Solitary; Parallel; Group | Functional Dramatic | | Constructive Games w/rules | | -- | No .80-.95[a] N = ? | None |
| Transdisciplinary Play-Based Assessment (TPBA; | 0–72 | Class/Peers or Assessor | Interaction | Exploratory through games with rules (no distinct levels) | | | Attention span Language use | -- | No None | None |

| Instrument | Age | Administration | Play categories | | | | Reliability | Validity |
|---|---|---|---|---|---|---|---|---|
| 1992) | | | | | | | | |
| Penn Interactive Peer Play Scale (PIPPS; 1995) | 60-85 | Teacher recall survey | Interaction; Disruption; Disconnection | | | -- | No .89-.90[c] N=493 | Construct Concur (SSRS) Predict |
| Test of Pretend Play (ToPP; 1997) | 12-72 | One-on-one w/ assessor | Self with everyday objects / Toy & non-representational materials / Representational toy alone / Self alone | | | -- | Yes .87[b] N=513 | Construct Concur (PPVT, Leiter) |
| PLAY (1999) | 36-60 | Class/Peers | Solitary; Parallel; Associative; Cooperative | Functional / Dramatic | Constructive Games w/rules | -- | No .92[a] N=42 | Concur (SSRS, BDI) |
| Child-Initiated Pretend Play Assessment (ChIPPA; 2000) | 36-84 | One-on-one w/ assessor | Object substitutions | Imitated actions | Elaborate pretend play actions | -- | No .97-1.0[a] .56-.84[b] N=82[a] 38[b] | Concur (PIPPS) Discrim |
| Play In Early Childhood Evaluation System (PIECES; 2005) | 19-52 | One-on-one w/ assessor | Exploratory through sequenced pretend play (13 levels) | | | -- | No .90-1.0[a] .48-.58[b] N=32 | None |

[a] Inter-rater reliability; [b] Test-retest correlation; [c] Internal consistency; BDI = Battelle Developmental Inventory; PIPPS = Penn Interactive Peer Play Scale; PPVT = Peabody Picture Vocabulary Test; RDLS = Reynell Developmental Language Scales; SSRS = Social Skills Rating Scale

cal dimensions and measures of behavior; the establishment of observational recording techniques and sampling methods; the analysis of behavior; and the assessment of reliability and validity. We discuss the extent to which we have been able to address each of these issues with a new observational measure of children's make-believe play. But before turning to how we developed this new measure, we offer a brief overview of existing observational assessments of play to further establish the need for such an instrument.

## Overview of Existing Play Assessments

In the absence of a tool to measure make-believe play effectively—and in particular the type of mature make-believe play that may have the greatest developmental impacts—we cannot empirically determine if play has educational value, or, if so, its nature and extent. We therefore conducted a review of the reliability and validity of peer reviewed observational measures for evaluating children's mature play, summarized in figure 1. We evaluated each assessment with respect to the Make-believe Play Observational Tool (MPOT) by asking several questions: Does the measure assess each of the key elements of mature play as defined by Vygotskian theory? Was the measure developed for administration in authentic play contexts—for example, does the protocol specify that each child should be observed during typical play with other children? Does the measure capture information about the play context, the physical and social environment in which play takes place? Does the measure evaluate both child and teacher involvement and, thus, capture teacher behaviors that scaffold or facilitate mature play? Does the measure demonstrate adequate reliability? Does the measure demonstrate adequate content validity?

### Intended Administration Setting

To assess play development adequately, we must observe how and when children engage in distinct play behaviors (e.g., motor, constructive, and dramatic play) in social settings with other children. Thus, evaluations of mature play should ideally be conducted in authentic social contexts. Evaluating play within a classroom setting can facilitate improvements in both its quality and a teacher's scaffolding of mature play. It embeds assessment within a supportive system without pulling children from classroom activities, singling them out, asking them direct questions, or giving them decontextualized instructions. The MPOT

was developed to assess children's make-believe play abilities in classroom settings to allow observation of child play in authentic social contexts and support improvements in play instruction. By contrast, most of the existing play measures we reviewed—Test of Pretend Play (ToPP), Transdisciplinary Play-Based Assessment (TPBA), and Child-Initiated Pretend Play Assessment (ChIPPA)—all evaluate laboratory-based child play with an adult. One instrument, Penn Interactive Peer Play Scale (PIPPS), queries teachers about child play behaviors and captures information limited to the play's context.

### Incorporation of Information about the Context of Play

The assessments we reviewed varied in the methods they used to capture information about the context of the play—that is, the physical and social environments in which the play takes place and is observed. Ideally, assessments developed for regular play capture information about play conditions (e.g., staged or authentic play scenarios, live observations or recall of general play behavior, the extent to which cognitive components of play are included) (Barnett, 1991; Linder, 1993; Nicolich, 1977; Power and Radcliffe, 1991). Current play assessments lack this information, but it is embedded in the MPOT.

### Evaluation of Teacher Behavior

Another important aspect of an effective play assessment concerns teacher involvement in play. Research suggests that children's mature play improves with classroom-based play interventions that include targeted suggestions from adult teachers (Berk and Winsler 1995; Bredekamp 2005; Dickinson and Tabors 2001; Enz and Christie 1993, 1994; Roskos and Neuman 1993; Stanton-Chapman 2015). More specifically, teacher interactions improve the complexity and duration of children's play (Berk and Winsler 1995; Bredekamp 2005; Dickinson and Tabors 2001; Enz and Christie 1993, 1994; Roskos and Neuman 1993). For example, teachers can intervene in children's play activities to encourage positive social behavior (for example, by modeling dispute resolution or engaging sidelined children), suggest modifications to existing activities that support academic skills, and increase play complexity when children's play is immature (e.g., when children persist in a single imitation, such as pretending to be a super hero and do not generate new ideas or scenarios). As teachers monitor children's progress during mature play time, their behaviors can support children within their zone of proximal development. Unlike other current play assessments, the MPOT records information about teachers that takes into account their intervention during play, affording us the

opportunity to examine the relationship between adult intervention and the level of play and its associated outcomes for the children.

Reliability. Reliability is a necessary but not sufficient condition for attaining psychometric quality. Of the five measures involving direct observation in situ, four have adequate reliability data. The fifth, the Transdisciplinary Play Based Assessment (TPBA) is well-liked and used within the field of special education even though it has no reliability data (Linder 1993).

Content validity. Although each measure we reviewed was designed to evaluate children's level of social or cognitive development through play (i.e., the overall construct), variety exists in the content each of these assessments actually measures. In 1968 the Smilanksy Scale for Evaluation of Dramatic and Sociodramatic Play (SSEDSP) began with five elements of play, including four levels within each of these subdomains. The interim observational measures have limited the scope to two: interaction level (i.e., social development) and symbolic level. Although reducing the content included in the measure dramatically simplifies the challenge of attaining psychometric quality, it also reduces the overall value of the measure. For example, "persistence" is not included as a subdomain of interest on most of the measures; yet, this may be the one element that captures a child's level of self-regulation—an important component thought to be supported by mature make-believe play. This is particularly important in early childhood, because self-regulation has consistently predicted academic success above and beyond intelligence (Barnett et al. 2008; Blair and Razza, 2007; Diamond et al. 2007; McClelland et al. 2007; Raver and Knitzer 2002; Rothbart, Posner, and Kieras 2006; Saarni, Mumme, and Campos 1998).

Normative data and concurrent validity. Of the measures reviewed, only one has been adjusted to evaluate play among children between the ages of two and seven years—the Test of Pretend Play (ToPP, formerly the Symbolic Play Test). Only half of the measures report some form of concurrent validity (i.e., PLAY and ChIPPA). Reviews of the two measures found in the *Mental Measurements Yearbook* (Lindsey and Soares 1998; Paolitto and Switzky 1995) suggest that this validity is based on a small subsample of the original sample and hence is of questionable validity. A similar pattern is true for two of the other three as well—the concurrent validity is based on a small follow-up study with a very small sample (Farmer-Dougan and Kaszuba 1999; McAlonely

and Stagnitti 2009). The only exception is the PIPPS, but again this measures only children's social play development and is a teacher-rating scale.

Because there is no widely used, reliable observational measure of classroom play that captures information about children's make-believe play, we developed a sensitive and reliable measure of it that can be used to identify implementation features and in causal evaluations of classroom-based play interventions. Such an instrument will fill a critical gap and help teachers to support and scaffold this play. The MPOT serves this purpose.

## Development of a Mature Make-Believe Play Observational Instrument

For this study, we sought to develop a sensitive and reliable measure of mature make-believe play that could be used to identify implementation features and program fidelity for causal evaluations of classroom-based play interventions. To ensure systematic and reliable observation, assessments must provide observers with a list of critical elements or categories to document and an organized means to record the information gathered (Christie and Roskos 2006). We discuss our theoretical rationale for selecting observation categories included in the MPOT. Following recommendations from Snow and Van Hemel (2008) and the National Research Council (2008), we considered MPOT's implications for children with disabilities and for members of minority language and cultural groups; we incorporated assessment of nontraditional areas of development, including affect, interpersonal interaction, and opportunities for self-expression; and we sought to include information about how to use the results productively and appropriately. We also worked to ensure our assessment did not impose additional structural burdens on children and teachers. To this end, we attempted to develop an assessment that did not require children be pulled from classroom activities, singled out, asked direct questions, or given decontextualized instructions.

To identify observable behaviors characteristic of mature make-believe play, we drew on the work of Elkonin (1978, 2005a, 2005b) and Vygotsky (1978, 1987; Bodrova and Leong 2007), who provide rich descriptions of developmental changes in play. Figure 2 delineates several differences between mature and immature make-believe play that provide the theoretical foundation for our observation categories. First, in mature make-believe play, children use substitute objects that often little resemble the objects they symbolize (e.g., they use

a stick as a horse or a box as a train car). During this symbolic play, children use gestures to represent actions with real or imaginary objects (props). For two- and three-year-olds, play appears largely object centered—the children determine their choice of roles through available objects. For those between the ages of three and five, this immature form of play gradually transitions to more elaborate, relationship-centered play. Such mature play is characterized by well-defined imaginary roles. Additionally, children engaged in mature play can communicate their reasons for adopting specific rules and can adhere to these rules instead of submitting to their immediate desires.

In mature play, children are able to step out of the pretend scenario and discuss various aspects of their play (metaplay). Children engaged in mature play participate more in social play, interacting with one another in the same play context (play interactions). Mature play is also defined by the ability of those children engaged in actions and interactions to take on and to sustain specific roles (role playing) and the speech associated with these roles (role speech). The more mature the play, the richer children's role representations. Mature role playing also involves the ability to follow the rules associated with pretend scenarios in general (e.g., playing hospital versus playing school) and with chosen characters in particular (e.g., playing a doctor versus playing a teacher).

Thus, according to Elkonin, play starts with the object-centered role play of two- and three-year olds (level 1) during which object-oriented actions determine their choice of roles, and such play evolves gradually to become the elaborate relationship-centered play of kindergarten-aged children (level 4), a play characterized by well-defined roles as well as by children's awareness of the reasons for the rules they adopt. The ability to follow rules in play rather than to submit to immediate desires seems first to appear at level 2 but does not fully develop until level 4. Combined with the changes in the use of play props and in the relationships between play roles and play actions, this evolution of play rules allows us to consider level 4 the level of fully developed or mature play. Although a relationship exists between a child's age and the level of his or her play with older children and being able to engage in more mature play, not every child reaches the highest level of play by the end of kindergarten.

## MPOT Dimensions

We developed child- and adult-focused dimensions of the MPOT to reflect

Figure 2. Characteristics of Mature and Immature Make-Believe Play

| Characteristics | Mature | Immature |
|---|---|---|
| **Props**<br><br>*Bodrova & Leong, 2007; Elkonin, 1978* | • Create props to fulfill the need of a play role<br>• Use speech or symbolic gestures in their place of props, to fit their roles<br>• Solve disputes and disagreements by inventing props rather than fighting over them<br>• Teacher-created props are accepted and used, as long as there are child-created ones as well<br>• Re-created props are also used | • Use props only realistically and cannot invent new ones<br>• Play actions using props imitate isolated everyday actions (e.g., feeding the doll)<br>• Children fight and argue over props and roles<br>• Teacher-made props are predominant |
| **Metaplay**<br><br>*Bodrova & Leong, 2007; Elkonin, 1978* | • Children use language to choose, to define, and to negotiate roles without being prompted<br>• Children discuss and create a pretend scenario and act out a scene developed in that scenario<br>• Children have extended conversations about their play | • Children don't or can't describe what they will do before beginning the action *or* anytime throughout play |
| **Play Interactions**<br><br>*Bodrova & Leong, 2007; Elkonin, 1978* | • Children engage in cooperative play<br>• Children engage in associative play | • Children engage in parallel play<br>• Children play alone |
| **Role Playing**<br><br>*Bodrova & Leong, 2007; Elkonin, 1978* | • Children use sequenced actions in a longer play scenario<br>• Most children follow play scenario rules, applying them to themselves and others<br>• Other-regulation precedes self-regulation<br>• Children may play several roles at once | • Play is primitive<br>• Play is based on one role or a single prop<br>• Pretend actions are repetitive and stereotypical<br>• Play episodes are short<br>• Children maintain a role for 5-10 minutes before moving on to another non-play activity |
| **Role Speech**<br><br>*Bodrova & Leong, 2007; Elkonin, 1978* | • Children use theme-related words associated with their roles<br>• Some children use gestures to communicate due to language barriers<br>• Children extend or improvise new scenarios through role speech<br>• Some children may adjust language and actions to indicate a new role | • Children use little or no role speech<br>• Children use one-word utterances or repetitive phrases<br>• Children's role speech uses only labels for the person or their role (e.g., "I'm the mom.") |

research and theory regarding critical components of mature play. We now summarize subcomponents of each dimension and discuss how they capture aspects of mature play emphasized in Vygotskian theory. We offer the basis we used for the ratings in level 4 in order to provide greater detail in the MPOT scoring shown on figure 3.

## Child Dimensions

Child-created props (CCP). Child-created props, gestures, and behaviors support and extend children's mature role play. Vygotsky's theory of imaginative play delineated how the use of props become more elaborate as play matures. At the highest level of mature play, children may enlist one

prop for several purposes and invent new props to resolve disputes or sup-
port new storylines (Vygotsky 1978, 1987). For example, a child might use a
wooden block as a telephone or a credit-card scanner at the grocery. We base
the MPOT CCP rating on the quality and quantity of props (real or symbolic)
that children create or repurpose in the context of a specific play scenario
rather than props created by the entire class as a group project. In mature play,
CCPs might also include gestures that imitate an action or communicate an
aspect of the role play in place of physical props (e.g., raising a hand to an ear
to simulate a phone call).

Child metaplay (CMP). Others have broadly defined child metaplay as
"the act of stepping out of a make-believe role to think or communicate about
play" (Trawick-Smith 1998, 433; also see Fein and Schwartz 1986). In the context
of make-believe play as we define it here, CMP refers to a child's ability to discuss
play scenarios and his or her role within them. CMP involves a conversation,
not merely a child stating what he or she does; children should demonstrate an
understanding of the various aspects of the roles they are playing. For example,
a child might say "I'm the waiter, I'm going to take the customer's order and
bring the food."

Play Interaction (PI). Play interaction refers to the level of action
between children in the play scenario. This includes descriptions of four levels
of interaction in play: alone (a child in the same location as others but clearly
engaging in a unique activity); parallel (a child playing independently on
task that others are also playing); associative (children—with or without the
teacher—interacting in the context of a scenario without a predetermined out-
come or rules); and cooperative (children acting out a planned scenario together
in sequence with clear goals) (Parten 1932). Ideally, children will engage in
cooperative play and follow a prescribed plan for play with common goals and
understanding of rules and roles.

Children's role playing (CRP). Children's role playing refers to the
extent children maintain their decided roles during a scenario. Role maintenance
draws on children's ability to ignore external, conflicting signals in the environ-
ment (e.g., viewing a block as a block) and instead maintain past rules (the block
is a phone) as they interact with their peers in ways that support pretense. The
ability to maintain imaginative play constitutes a critical component of mature

pretend play. Children may take on various roles in a play scenario because peers to play such role are absent, but they should not switch roles without purpose.

Child role speech and communication (CS). Child role speech and communication refers to the language and gestures children use during the play scenario. This involves children's ability to maintain speech consistent with their role. For example, a child playing a doctor might ask, "What hurts?" CS also captures gestures children might use in their roles and their abilities to change their language based on the roles they play (for example using baby talk with an imagined infant). Gestures are an important part of CS because children with language problems will likely use gestures as a substitute for role speech.

*Adult Dimensions*

In addition to creating an MPOT that captures key aspects of children's make-believe play, we also created a tool that captures information about child-teacher interactions and teacher support during children's play—unique information not typically captured in play assessments.

Center management (CM). Center management refers to a teacher's system for managing center rotations and includes such supports as a planning wheel or chart and color-coded cards. CM also includes activities such as teacher behavior management within the center.

Planned play time (PPT). Planned play time indicates the cumulative time given to uninterrupted planned, structured play (ranging from less than thirty minutes to more than one hour each day). Short or interrupted play intervals typically impede children's ability to develop scenarios reflective of mature play.

Teacher intervention (TI). Teacher intervention refers to how teachers interact with children in the play scenario. Although teacher intervention can benefit children, theoretical and pedagogical accounts often emphasize the importance of ensuring that teacher interventions during child play be brief and targeted rather than intrusive (Stanton-Chapman 2015). For example, some interventions, such as *Tools of the Mind* (TofM) (Bodrova and Leong 2007) offer three guidelines for teacher intervention during play. First, teachers should take a minor role in children's play scenarios (e.g., a passenger, rather than a pilot,

during an imaginary plane ride). Second, teachers should watch for opportunities to model sociodramatic play skills (e.g., role playing and using symbolic props) and authentic, play-related academic activities (e.g., in literacy, writing an imaginary shopping list or ordering imaginary food from a menu or, in math, counting the number of imaginary apples a customer buys at an imaginary grocery store). Third, teachers should intervene only to model play roles, extend role speech and scenarios, or exit play scenarios when their goals have been completed. In keeping with this theoretical perspective, teachers receive high scores for this MPOT item when they intervene only to guide or scaffold children in their play, then leave.

## Evaluation of Content Validity and Instrument Refinement

The MPOT construction followed procedures customary to develop high-quality assessments (Kline 2005; Snow and Van Hemel 2008). The construction of the MPOT involved multiple cycles of development and trials, including the construction of an assessment blueprint, the development of more criteria (items) and refinement of redundant components and indicators, the refinement of user materials, and the completion of beta- and pilot testing prior to field testing. We invested significant time in developing and field testing to achieve a high-quality assessment tool.

   First we refined the design and items, added items where appropriate, and developed criteria, user instructions, and a training workshop. We drafted an initial scoring design including scales. We asked for input and review of our procedures and content from an advisory group (AG) made up of early childhood practitioners, teachers, and experts in the field of play. We derived targets for cumulative mature play time from Christie and Wardle (1992), which the AG members reviewed to establish initial construct and content validity. We used the feedback from AG members to develop a test of the assessment that we used in a small-scale beta-test study with ten classrooms.

   The beta testing provided a way to study assessment processes; to obtain user input on usability, clarity of criteria, and accuracy; and to gather a second round of construct- and content-validity findings. We videotaped all beta-test assessments for in-depth analysis. Assessors also completed a brief survey of their experiences observing constructs and implementation and a questionnaire evaluating them.

The AG reviewed beta-test data to refine the assessment, protocols, scoring, user manual, and training workshop in preparation for their use in the study. The AG scrutinized the viability of the test items in the revised protocol (individually and collectively), the assessment design, and the scoring system and metrics. In consultation with the AG, we also developed a video example (from the beta-test) of each level of play and each element of play. We used this video to train assessors and to contribute to inter-rater reliability studies (i.e., allowing two raters to score the video-recorded performance of one classroom).

## Evaluation of the Mature Play Observation Tool

To evaluate the reliability and validity of the MPOT, we examined correlations between child play (as indexed by the MPOT) and relevant child outcomes, including self-regulation, language, and math skills. Specifically, we tested whether differences in the maturity of child pretend play related to differences in outcomes, as predicted by Vygotskian theory. We also examined the relationship of teacher-supporting behaviors to child outcomes.

### Method

We conducted our evaluation of the MPOT as part of a larger study examining the effects of a math- and play-based intervention. The study was implemented as a three-armed cluster randomized control trial with an intervention condition, an active control condition (a math-only intervention), and a standard classroom context (the business-as-usual condition). Because the MPOT study focuses on the implementation and evaluation of a play-based assessment, we have based our analyses on the combined math- and play-based intervention group (an integrated group of twenty-eight teachers). Briefly, teachers in the integrated group comprising the study population administered Building Blocks (BB) (Clements and Sarama 2013), a research-based mathematics curriculum that addresses geometric, spatial, and quantitative competencies, in combination with a second scaffolding component designed to improve children's self-regulation. Make-believe play was a primary mechanism for the scaffolded improvement of children's self-regulatory skills. As part of the scaffolding component, teachers implemented specific play themes over four to six weeks. Teachers received extensive training in the math and self-regulation curriculum. During the two-year implementation period, teachers participated in eight sessions of

professional development, each lasting between two and four days. Substitute teachers were provided for teachers participating in the study. Instructional coaches and mentors also supported teacher learning and intervention implementation. Implementation came in two parts, the first year as a "gentle" (or practice) implementation and the second year as a full implementation.

Teachers in the both groups received training during each of the two years of their involvement. Both groups received the same training in math education with the Building Blocks curriculum—two days in the first month of school of year one and two school days in the fall and two in the spring (with the project paying for substitutes) of year two. Although comprehensive, this research-based plan was designed to be commensurate with support the program received under normal but adequate conditions of adoption. The course included learning trajectories for each math topic, learning trajectories for observation and other authentic assessment strategies, supporting mathematical development in the classroom, recognizing and supporting math throughout the day, setting up math learning centers, teaching with computers (including the use of the management system and research-based teaching strategies), small-group activities, and supporting mathematical development in the home. The main technological tool was the Building Blocks Learning Trajectory (BBLT) web application. BBLT provides scalable access to the learning trajectories in descriptions, videos, and commentaries. All aspects of the learning trajectories—developmental progressions of children's thinking and connected instruction—are linked (Clements and Sarama 2013; Sarama and Clements 2004).

The Building Blocks Scaffolding Executive Function (BBSEF) group also included additional training on scaffolding executive function (SEF) (i.e., those in the group received twice as much professional development). The SEF training, delivered by authors of Tools of the Mind (Bodrova 2007) and their colleagues, included an additional six days of training each of the two years the teachers were involved. The SEF training followed the general organizational structure of the BB training and included the topics of development in executive function (EF) in early childhood, how dramatic play supports EF, and how teachers can scaffold mature and intentional dramatic play. Like BB professional development, SEF professional development combined enhancing teachers' knowledge about childhood learning and development with helping them master effective instructional strategies designed to support this learning and development. The idea of mature dramatic play as a critical component in promoting EF is new to most teachers, and so is the idea that mature play may not emerge spontaneously but requires teacher scaffolding. Accordingly, SEF staff paid special attention to

teachers' understanding of these new concepts. Videotapes illustrating various stages in play development as well as best practices for scaffolding play led to discussions during training. In addition to SEF-specific training, the BBSEF teachers received training on modified BB instructional strategies redesigned especially to promote EF.  Further, the EF strategies added to the math curriculum combined activities focused on implementation outside the math block with instructional strategies embedded in math activities.

We coached teachers throughout the two years within each classroom. Coaches observed the implementation of BB and the self-regulation (SR) curriculum and offered feedback, and they helped acquire material and provided technical troubleshooting. They lent on-site coaching support for hour-long, biweekly sessions and offered teachers feedback using the Classroom Needs Assessment (CNA). Coaches completed the CNAs during the observation period and sent them to the lead coach at the conclusion of each visit. They also provided coaching support off site and made themselves available to teachers and research coordinators via email, phone, and FAX. All three districts employed instructional coaches for mathematics, and the same coaches within each district worked with teachers concerning as many as three of the study conditions.

Coaches reported to a lead coach as well as to mentors (in the second year) after each visit, who in turn reported to site coordinators. This real-time, classroom-based observation, demonstration, coaching, and mentoring, which included shared decision making, one-to-one consultation, monitoring, and reinforcement, was considered indispensable.

### Procedure

Independent observers conducted MPOT monitoring in all classrooms at the end of the second year of implementation. Additionally, independent consultants, primarily former educators, conducted assessments of children's executive functioning, mathematical achievement, vocabulary, and spoken language comprehension and production. All children completed assessments in the following order: head-toes-knees-shoulders (HTKS), pencil tap, backward digit span, Test of Early Achievement in Mathematics (TEAM), and the Renfrew bus story.

## Characteristics of Observed Classrooms

MPOT observations were carried out within the classrooms of teachers ran-

domly assigned to the integrated group (N= 28). Scheduling constraints prevented observation in two assigned classrooms, so a total of twenty-six teachers were observed. At the time of observation, teachers were approximately two weeks into a new play theme (M = 1.87 weeks). In a baseline survey, the integrated-group teachers reported their highest level of education and average years of instructional experience (M = 11.3 years). Six teachers held a High School Diploma or GED (22 percent), fifteen teachers held a Bachelor's degree (54 percent) and two held a Master's degree (7 percent). Two-thirds of the teachers indicated they held specific certifications in early childhood instruction (N = 18).

In total, 286 study-enrolled children were present during the observations of integrated-group classrooms (females = 142. Gender was not reported for 8 children; age was not reported for 3 children.) Although child participants spanned a wide age range (3.46 – 8.89 years), 94 percent of child participants were four to five years old ($M_{age}$ = 4.67 years; 2 percent of participants < four years; 4 percent of participants > five years). Average class size for integrated group classrooms totaled 23.31 students, and approximately half of students in each classroom spoke English as a second language (average N = 12.62).

*Mature Play Observation Measure*
As previously mentioned, the MPOT categories (Bodrova et al. 2012) include both teacher and child and components. Detailed and leveled items guide observers through systematic evaluation of child and teacher actions during play-based activities. This measure of children's behaviors, and teacher's behaviors (as presented in figure 3) was designed to be low-inference and anchored with specific instructional behaviors and characteristics.

Moreover, the MPOT was developed to evaluate the quality of mature make-believe play occurring in preschool classrooms. It identifies specific aspects of children's mature make-believe play and supporting adult behaviors, as described in extant theoretical accounts of play behaviors and contexts (Bodrova and Leong 2007; Bodrova, Germeroth, and Leong 2013; Elkonin, 1978, 2005a, 2005b; Vygotsky, 1978, 1987). Observers rate the authentic, classroom play using a variety of factors that include child behavior and teacher involvement during play. We instructed observers to watch for one hour or the extent of the classroom dramatic play period. The adult and child behaviors they observed are defined by tiered components to provide a systematic means to assess each, with scores ranging from 1 to 4 for each set of behaviors and a sum score of these for each. Rated child behavioral components include child-created props,

Figure 3. Description of Adult and Child MPOT Dimensions

| Child Dimensions | Level 4 Anchors |
|---|---|
| **1.1 Child-Created Props (CCP)**<br>The extent to which props are created to support children's play roles. Excludes props created as part of group activities in which each child constructs the same prop (e.g., pirate hat). | • More than 5 or more child-created props to fit children's play roles.<br>• Re-created props, newly created props, and play without props (indicated by word or gesture) are used by all children.<br>• Children create props as needed or use speech or gestures in their place.<br>• Props are not teacher organized. |
| **1.2 Child Meta-Play (CMP)**<br>Children talk about how and what they will be playing. Talk about roles, rules, and props that are needed. Meta-play can occur during play planning, as well as during play itself. | • Children engage in dialogues greater than 15 minutes about what the play scenario will be, and how the scenario will unfold.<br>• Children engage in extended discussions about their roles, actions, and the use of props prior to starting their play as well as when the play scenario is about to be changed. |
| **1.3 Play Interactions (PI)**<br>The extent to which children interact with each other during play. | • Children must have specific roles and a "script" (predetermined sequences).<br>• Children act out a planned scenario together in sequence with a clear goal.<br>• Children generally follow predetermined rules of the roles and those rules govern the steps or sequence of their behaviors. |
| **1.4 Children's Role Playing (CRP)**<br>The extent to which children maintain their roles and associated rules during mature play. | • Children are able to sustain play for longer than 15 minutes.<br>• More than two actions in play scenario are sequenced.<br>• Most children follow play scenario rules. Rules are followed-including children applying rules to selves, apply rules to others, and general awareness of rules. At this level there are fewer occurrences of other regulation because rules are not broken.<br>• Optional: Children play several roles at once |
| **1.5 Child Role Speech and Communication (CS) During Play**<br>Includes oral language and gestures to communicate role meaning. | • Children extend or improvise new scenario through role speech, they take on suggestions of other children and incorporate the unexpected into play.<br>• Optional: Children adjust their language and actions to indicate a new role. (Ex. Changing register to sound like a baby v mom) |
| **Adult Dimensions** | **Level 4 Anchors** |
| **2.1 Center Management (CM)**<br>The extent to which the teacher uses a management system to support children's self-regulated play. | • A system for making center choices is used daily (i.e., necklaces, clips, etc.).<br>• More than two visual mediators such as concept maps, role cards, or pictures to support play are displayed and accessible. |
| **2.2 Make-Believe Play Time (MBPT)**<br>The amount of time for uninterrupted play. | • One hour of uninterrupted play time occurs every day. |
| **2.3 Teacher Intervention (TI)**<br>The extent to which the teacher or assistant teacher intervenes during play. | • Teacher intervenes in play and models play roles.<br>• Teacher intervention includes language extensions.<br>• Teacher stays in play scenario briefly and purposefully.<br>• Teacher uses multiple strategies for intervening, such as using visual mediators or children to intervene in play scenarios.<br>• No teacher intervention because the play is already functioning at high levels and teacher intervention is not needed. |

child metaplay, play interactions, children's role playing, and child role speech and communication. Those for teachers include center management, planned play time, and teacher intervention.

We trained our MPOT observers—using still photos, video examples, and live observations—to recognize the essential elements of mature play. Observers had to achieve 100 percent reliability to become certified and, in this study,

achieved 97 percent reliability for data collection. The average observation length was 51.18 minutes.

*Self-Regulation Measures*

HEAD-TOES-KNEES-SHOULDERS TASK. The HTKS directly measures behavioral regulation that requires children to demonstrate three skills—inhibitory control, attention, and working memory. The task is administered frequently in classroom intervention studies. It predicts performance on other self-control measures and activities and shows high reliability (McClelland et al. 2007; Ponitz et al. 2008). Children are taught a game in which they are to do the opposite of what the examiner tells them to do. For example, if the examiner instructs a child to touch his or her head, the child should touch his or her toes. Children completed four practice tests and heard up to three repetitions of the instructions before the testing portion began. Testing included ten items in random order with possible scores of 0, 1, or 2. Higher scores indicate higher levels of regulation. A score of 0 indicates an incorrect answer, a score of 1 indicates that the child self-corrected (defined as any movement toward an incorrect answer that the child then corrected to provide a correct answer), and a score of 2 indicates a correct answer without any hesitation or movement toward an incorrect answer.

BACKWARD DIGIT SPAN TASK. Backward digit span (Wechsler 2003) measures a child's short-term working memory. Children are asked to repeat digits in the reverse order from which they heard them. Children had to pass a pretest of at least one of two practice trials to move on to scored items. Testing continued until a child missed two items in a row. Scores reflect working memory and executive functioning skills, with backwards span capturing information about children's ability to maintain and manipulate information in working memory.

PENCIL TAP TASK. The pencil tap task (Luria 1966; Diamond and Taylor 1996) is a normed, widely used measure of inhibitory control. Children are asked to tap a pencil on a desk either once or twice after watching the assessor tap. The child must tap once if the assessor taps twice and tap twice if the assessor taps once. Children must attend to the instructions and his or her response, while inhibiting the desire to tap the same number of times as the assessor. The test is individually administered and takes approximately five to eight minutes, depending on the ability of the child.

### Math Measures

Tools for early assessment in math (TEAM). TEAM (Sarama, Wolfe, and Clements 2011) assesses children's foundational math skills based on the developmental trajectories of the Building Blocks framework. TEAM involves interviews that follow a specific protocol. Fluidity of presentation for TEAM administrators requires extensive training. TEAM includes two parts. Part A assesses number concepts including subitizing (the ability to identify quickly without counting the number of items in a small set), verbal counting, number comparison, number sequencing, number composition and decomposition, adding and subtracting, place value, and multiplication and division. Part B assesses shape recognition, shape composition and decomposition, congruence, construction of shapes, and mental imagery. Part B also looks at concepts that are secondary to early math development such as geometric measurement and patterning using geometric shapes.

### Literacy Measures

Phonological awareness literacy screening (PALS). PALS (University of Virginia 2004) is an individually administered assessment of emergent literacy skills that can be used by trained classroom teachers. PALS includes six subtests assessing progressively advanced skills including name writing, alphabet knowledge, beginning sound awareness, print and word awareness, rhyme awareness, and nursery rhyme awareness. For this study, we used scores from the name-writing and alphabet knowledge subtests. The alphabet knowledge subtest includes three sections that assess uppercase letter recognition, lowercase letter recognition, and knowledge of letter sounds. Subtests were scored individually.

Renfrew bus story-North America (RBS-NA). The RBS-NA (Cowley and Glasgow 1994) measures expressive and receptive language abilities for children aged three years to six years and eleven months. Designed as a narrative retell, the test requires children first to listen to a story then to repeat it using their own language. This is a fast and enjoyable assessment for children because it is based on the familiar framework of storytelling. Originally developed and standardized in the United Kingdom, the RBS-NA incorporates vocabulary and concepts familiar to North American children.

Retelling stories during the RBS-NA requires students to coordinate visual and auditory input. As children listen to the narrative, they must attend to and

comprehend the original story. During the second phase, they must recall the story, formulate sentences, and replicate the narrative. Scoring provides indicators of these abilities and a broad overview of a child's language abilities. The RBS-NA can be used to monitor speech and language disorders via in-depth language evaluation and to identify norm-referenced language abilities. The RBS-NA takes about ten minutes to administer and can be used by special-education teachers, researchers, or speech and language therapists to determine language abilities. For the study described in this article, we evaluated child performance using standardized quantitative (standard scores and percentile ranks) and qualitative scores.

# Results

Our preliminary data demonstrates the reliability and validity of the MPOT. We also present exploratory analyses examining how child and teacher responses on the MPOT relate to children's performance in assessments of self-regulation, literacy, and numeracy.

## MPOT Validity and Reliability

Criteria for each MPOT set of responses were designed to provide adequate content coverage for mature play behaviors typically demonstrated by the three- to five-year-olds in standard early childhood classrooms. Following field testing and refinement, trained MPOT observers performed reviews in twenty-six of twenty-eight classrooms; observations could not be conducted in two classrooms because of scheduling and logistical constraints. The MPOT demonstrated high inter-rater reliability among fourteen trained raters (Cronbach's α = .909).

We tested the predictive validity of the MPOT by evaluating how classroom play ratings related to children's self-regulatory, math, and literacy performance. Because data in this study are hierarchically structured to reflect MPOT administration at the classroom level and self-regulation, math, and literacy assessment administration at the student level, we tested relationships across variables using multilevel models with teachers as random intercepts. All models were generated in the R lme4 package (Bates et al. 2015), with effect sizes reported as standardized betas ($\beta$). P-values were generated using Satterthwaite's method for degree of freedom approximation via the R package lmerTest (Kuznetsova, Brockhoff, and Christensen 2017). Finally, given that the majority of children

in the classrooms we observed fell between the ages of three and five years, we limited our analyses to this subset of students, excluding children older than six (N=15) or for whom age was missing (N=3). Descriptive statistics for each measure and correlations with age are presented in figure 4.

## MPOT Measures and Self-Regulation Outcomes

To evaluate relationships between children's make-believe play and their ability to regulate their behavior, we related classroom scores on the MPOT to three measures of self-regulation: head-toes-knees-shoulders (HTKS), backward digit span task, and the pencil-tap task. Intraclass correlation coefficients (ICCs) were computed for each measure using unconditional intercept-only models. ICCs ranged from .11 (pencil-tap) to .18 (HTKS), demonstrating that 11 percent to 18 percent of variance in child scores could be accounted for by classroom differences before other predictors were added to the model.

On average, children in classrooms that rated higher on child portions of MPOT showed better backward digit span performance ($\beta$ = .192, t = 2.13; 95

Figure 4. Characteristics of MPOT, Self-regulation, Math and Literacy Measures

| Measure | Mean | Range | Standard Deviation | Correlation with Age |
|---|---|---|---|---|
| *MPOT Child Dimensions* | | | | |
| Child-Created Props | 1.91 | 1.00 - 4.00 | 0.95 | ns |
| Child Metaplay | 1.80 | 1.00 - 4.00 | 0.87 | ns |
| Play Interactions | 2.48 | 1.50 - 4.00 | 0.72 | ns |
| Children's Role Playing | 2.34 | 1.00 - 4.00 | 1.14 | ns |
| Child Role Speech and Communication | 2.24 | 1.00 - 4.00 | 1.01 | ns |
| Sum | 10.77 | 5.50 - 20.00 | 4.17 | ns |
| *MPOT Adult Dimensions* | | | | |
| Center Management | 2.63 | 1.00 - 4.00 | 1.12 | ns |
| Planned Play Time | 2.34 | 1.00 - 4.00 | 1.09 | ns |
| Teacher Intervention | 2.91 | 1.00 - 4.00 | 1.00 | .174* |
| Sum | 7.88 | 4.00 - 12.00 | 2.85 | ns |
| *Head-Toes-Knees-Shoulders* | 15.87 | 0.00 - 40.00 | 14.36 | .232** |
| *Backward Digit Span* | .50 | 0.00 - 4.00 | 0.99 | .163* |
| *Pencil Tap* | 10.36 | 0.00 - 16.00 | 6.05 | .294*** |
| *TEAM Scaled Score* | 365.7 | 35.0 - 586.0 | 95.30 | .375*** |
| *PALS Literacy Measures* | | | | |
| Uppercase | 16.26 | 0.00 - 26.00 | 9.77 | .249*** |
| Lowercase | 21.57 | 0.00 - 26.00 | 4.42 | ns |
| Letter Sounds | 15.33 | 0.00 - 26.00 | 8.34 | ns |
| Name Writing | 5.98 | 0.00 - 7.00 | 1.49 | .276*** |
| *RBS-NA* | 46.19 | 11.00 - 55.00 | 9.97 | .132* |

Figure 5. Relationships between MPOT Dimensions and Self-regulation Performance

|  | Head Toes Knees Shoulders | Backward Digit Span | Pencil Tap |
|---|---|---|---|
| *Child Dimensions* |  |  |  |
|    Child-Created Props | - | 0.192* | - |
|    Child Metaplay | - | - | - |
|    Play Interactions | - | 0.230* | - |
|    Children's Role Playing | - | - | - |
|    Child Role Speech and Communication | 0.215* | 0.216* | - |
|    Sum | - | 0.192* | - |
| *Adult Dimensions* |  |  |  |
|    Center Management | - | - | - |
|    Planned Play Time | - | - | 0.189* |
|    Teacher Intervention | - | 0.191* | - |
|    Sum | - | 0.191* | - |

*Note.* Values represent standardized betas. * $p < .05$ ** $p < .01$

percent CI = [.016 – .369]; $p = .044$). Individually, child-created props ($\beta = .192$, $t = 2.14$; 95 percent CI = [.016 – .368]; $p = .043$), play interactions ($\beta = .230$, $t = 2.70$; 95 percent CI = [.063 – .397]; $p = .013$), and role speech and communication ($\beta = .216$, $t = 2.51$; 95 percent CI = [.047 – .384]; $p = .019$) showed significant, positive relationships to child backward digit span performance (Table 5). Other self-regulation measures showed weaker associations with MPOT. Although HTKS and pencil-tap performance showed trend relationships with cumulative MPOT child scores, neither relationship met significance criteria ($p$'s < .10). Child role speech and communication did positively predict HTKS performance, however ($\beta = .215$, $t = 2.26$; 95 percent CI = [.029 – .402]; $p = .034$)).

Similarly, the sum of the adult portions of MPOT positively predicted backward digit span performance ($\beta = .191$, $t = 2.10$; 95 percent CI = [.013 – .369]; $p = .048$) driven by a positive relationship between teacher-intervention scores and backward digit span performance ($\beta = .191$, $t = 2.13$; 95 percent CI = [.015 – .368]; $p = .045$). Cumulative adult MPOT scores showed nonsignificant trends with HTKS and pencil-taps ($p$'s < .10), although planned play time did predict pencil-tap performance ($\beta = .189$, $t = 2.22$; 95 percent CI = [.022 – .356]; $p = .039$).

## MPOT and Math Outcomes

To understand the relationship between mature play and mathematics achievement, we examined relationships between the TEAM and MPOT results. Approximately 32 percent of variance in TEAM scores was explained by class-

Figure 6. Relationships between MPOT Dimensions and Math Performance

| MPOT Dimension | TEAM Scaled Score |
|---|---|
| *Child Dimensions* | |
| Child-Created Props | 0.299* |
| Child Metaplay | - |
| Play Interactions | 0.266* |
| Children's Role Playing | 0.258* |
| Child Role Speech and Communication | 0.309* |
| Sum | 0.300* |
| *Adult Dimensions* | |
| Center Management | 0.310* |
| Planned Play Time | 0.262* |
| Teacher Intervention | 0.327** |
| Sum | 0.337** |
| *Note.* Values represent standardized betas. * p < .05 ** p < .01 | |

room level variance in a no-intercept model (ICC = .32). In models adding MPOT elements as predictors, many MPOT components positively predicted children's cumulative TEAM scores (figure 6).

The MPOT child sum score positively predicted child TEAM performance ($\beta$ = .300, t = 2.56; 95 percent CI = [.070 – .529]; p = .018), as did four of the five child measures, including child-created props ($\beta$ = .299, t = 2.56; 95 percent CI = [.070 – .529]; p = .018), play interactions ($\beta$ = .266, t = 2.22; 95 percent CI = [.308 – .501]; p = .038), children's role playing ($\beta$ = .258, t = 2.11; 95 percent CI = [.018 – .497]; p = .047), and child role speech and communication ($\beta$ = .309, t = 2.77; 95 percent CI = [.090 – .528]; p = .011). The MPOT adult dimension cumulative score also showed a strong relationship with children's TEAM scores ($\beta$ = .337, t = 3.06; 95 percent CI = [.121 – .553]; p = .006), reflected across the three adult dimension subcomponents: center management ($\beta$ = .310, t = 2.65; 95 percent CI = [.081 – .538]; p = .015), planned play time ($\beta$ = .262, t = 2.24; 95 percent CI = [.032 – .491]; p = .036), and teacher intervention ($\beta$ = .327, t = 2.95; 95 percent CI = [.109 – .544]; p = .008).

## MPOT and Literacy Outcomes

To evaluate relationships between play and literacy outcomes, we examined relationships between MPOT, PALS, and RBS-NA results. The specific variables of interest include children's ability to identify uppercase letters, lowercase letters, and letter sounds, and name writing from the PALS. ICCs for each variable ranged from .05 (RBS-NA) to .45 (PALS letter sounds), indicating substantial

Figure 7. Relationships between MPOT Dimensions and Literacy Performance

| | PALS Literacy Measures | | | | RBS-NA |
|---|---|---|---|---|---|
| | Uppercase Letters | Lowercase Letters | Letter Sounds | Name Writing | |
| *Child Dimensions* | | | | | |
| Child-Created Props | - | - | - | 0.252** | - |
| Child Metaplay | - | - | - | 0.197* | - |
| Play Interactions | - | - | - | 0.236** | - |
| Children's Role Playing | - | - | - | .225** | - |
| Child Role Speech and Communication | 0.274* | 0.375** | 0.351** | 0.249** | - |
| Sum | - | 0.249* | - | 0.260** | - |
| *Adult Dimensions* | | | | | |
| Center Management | - | - | 0.318* | 0.280** | - |
| Planned Play Time | - | - | 0.290* | 0.213** | - |
| Teacher Intervention | - | - | - | 0.319** | - |
| Sum | - | - | 0.325* | 0.302** | - |
| *Note.* Values represent standardized betas. * p < .05 ** p < .01 | | | | | |

variability in the proportion of variance explained by classroom differences across literacy variables. Correlations between items on the MPOT and PALS assessments are summarized in figure 7.

Patterns of correlations across child MPOT and PALs suggest that mature play as assessed by the MPOT predicts some emerging literacy skills. The strongest relationships were observed between MPOT results and name-writing ability. MPOT child sum scores ($\beta$ = .260, t = 3.75; 95 percent CI = [.124 – .396]; p < .001) and each of the five dimension scores positively predicted name-writing performance, including child-created props ($\beta$ = .252, t = 3.64; 95 percent CI = [.116 – .387]; p = .001), child metaplay ($\beta$ = .197, t = 2.54; 95 percent CI = [.045 – .349]; p = .018), play interactions ($\beta$ = .236, t = 3.31; 95 percent CI = [.096 – .375]; p = .003), children's role playing ($\beta$ = .225, t = 2.96; 95 percent CI = [.076 – .374]; p = .007), and child role speech and communication ($\beta$ = .249, t = 3.52; 95 percent CI = [.110 – .388]; p = .002). Additionally, MPOT child sum scores positively predicted children's lowercase letter identification ($\beta$ = .249, t = 2.08; 95 percent CI = [.015 – .483]; p = .049). For individual categories, only child role speech and communication predicted child performance on uppercase letter identification ($\beta$ = .274, t = 2.48; 95 percent CI = [.057 – .491]; p = .021), lowercase letter identification ($\beta$ = .375, t = 3.49; 95 percent CI = [.164 – .586]; p = .002), and letter sounds ($\beta$ = .351, t = 2.67; 95 percent CI = [.093 – .609];

p = .014). We did not find any significant correlations between RBS-NA and MPOT scores.

Adult MPOT measures showed similar relationships to literacy outcome measures. Adult cumulative scores predicted child ability to produce letter sounds ($\beta$ = .325, t = 2.41; 95 percent CI = [.060 – .590]; p = .025) and name writing ($\beta$ = .302, t = 4.94; 95 percent CI = [.182 – .422]; p < .001), but not uppercase or lowercase letter identification. The individual MPOT adult measures showed positive associations with letter sounds and name writing performance apart from teacher intervention, which predicted name writing ($\beta$ = .319, t = 5.25; 95 percent CI = [.200 – .438]; p < .001) but not letter sound performance (figure 7).

In summary MPOT results positively correlated with aspects of children's emergent self-regulatory, math, and literacy skills. MPOT measures correlated most often with math performance, though correlations were also found with some measures of self-regulatory and literacy performance.

## Discussion

In our study, we evaluated the reliability and validity of a new classroom-based observational instrument of children's mature make-believe play, the Make-Believe Play Observational Tool. The MPOT was developed to evaluate the quality of mature make-believe play in preschool classrooms using an integrated, theory-based framework (Bodrova and Leong 2007; Bodrova, Germeroth, and Leong 2013; Elkonin, 1978, 2005a, 2005b; Vygotsky 1978, 1987) and to provide a mechanism for identifying implementation features and program adherence in causal evaluations of classroom-based play. During mature make-believe play, children create imaginary situations, take on explicit roles (using the language and rules of the roles), and use objects symbolically (Bodrova and Leong 2011; Elkonin 1978). Vygotsky saw all three components of play—imaginary situation, roles, and rules—as important in the formation of a child's mind, both the development of abstract thinking and of conscious and voluntary behaviors so critical in the early years. According to Vygotskians, make-believe play, when it reaches its mature state, fosters self-regulation and provides the basis for other activities or interactions that in turn foster the learning of symbolic and emotional thinking, oral language, and the beginnings of literacy.

We provide preliminary support for the MPOT. We found that the MPOT demonstrates high inter-rater reliability in classroom-based field testing (Cronbach's α = .909) and predicts children's performance in domains theorized to benefit from mature play, including self-regulatory, math, and literacy skills. Both child and teacher (adult) dimensions of the MPOT and cumulative scores were found to have moderate correlations with self-regulation outcome measures. Math skills were positively related to all child and adult measures, as well as the cumulative child and adult scores. Child and adult measures and cumulative scores also correlated with emergent literacy as assessed by PALS subscales for letter-sound knowledge and name-writing skills.

To replicate and extend these findings, the psychometric properties of the MPOT should be tested in a broader sample of classrooms, teachers, and students. Our study investigates the reliability and validity of the MPOT within a sample that was largely composed of four- and five-year-old children from three diverse, urban districts. Future investigations would benefit from more robust sampling of three-year-olds, allowing for estimations of sensitivity in a population just beginning to demonstrate mature forms of play and for specific investigations of differential item functioning across subpopulations. We have also investigated the MPOT within a relatively constrained intervention setting that may not generalize to other classroom environments. For this reason, it would be useful to investigate the efficacy of MPOT in more diverse curricular settings, although obtaining a large number of classrooms in which teachers are supporting make-believe play can be challenging. In particular, we have evaluated the MPOT in classrooms using a math curriculum that incorporates a playful educational component. This component may have inflated average classroom ratings in this study's sample relative to alternative curricular settings.

Additionally, though we have presented evidence that children from classrooms that scored highly on MPOT measures showed better self-regulatory, math, and literacy skills than children from classrooms that scored poorly on MPOT, the correlational design of the present study does not allow us to rule out alternative, noncausal interpretations for observed links between children's play and their performance. For example, children who engage in more make-believe play may show better performance on a variety of assessments because of other, unmeasured variables such as motivation or confidence. To build on the present correlational findings, future investigations could explore the directionality of relationships between make-believe play and outcomes by testing

whether changes in the maturity of children's make-believe play drive effects of play-based classroom interventions.

Because the MPOT evaluates play quality at the classroom level, the instrument cannot be used to make direct inferences about how individual differences in mature play ability relate to child outcomes. These inferences could be supported by testing the concurrent and predictive validity of the MPOT in relation to measures that evaluate play quality at the level of the individual. For example, MPOT scores may predict changes in play ability in individual children during a school year (e.g., as indexed by the Scale for Evaluation of Dramatic and Sociodramatic Play), since classroom experiences may scaffold and reinforce emerging mature play in other contexts.

We anticipate that future investigations will explore the usability and efficacy of MPOT as a professional development tool for educators. One goal in developing the MPOT was to create a reliable and valid instrument that could be used by classroom teachers to assess the level of make-believe play of their students. Presumably, having a better understanding of the level of student play should help teachers to improve their instruction. Although the rater-observers in our study were trained, we predict that the MPOT might serve well as a professional development tool to support teachers in conducting their own assessment of levels of play in their classrooms. If the benefits of make-believe play are most pronounced when children engage in mature play, classroom interventions may provide unique opportunities to improve the quality of children's play. In typical preschool classrooms, play rarely reaches the mature level and instead displays characteristics one would expect to see in the play of toddlers, including short play episodes that are repeated day after day, the use of toys that are exact replicas of real objects, and simple scripts (Singer, Golinkoff, and Hirsh-Pasek 2006). Moreover, incorporating mature make-believe play in preschool classrooms may yield benefits in other domains. For example, playful learning environments have been shown to reduce children's stress and increase academic skills (Miller and Almon 2009).

Although we concede that many kinds of play are important in the early-childhood classroom, from a Vygotskian point of view, make-believe play is most critical for supporting cognitive and social-emotional development (Vygotsky 1967). Make-believe play offers a unique opportunity for all children to operate in ways that push their individual developmental edge when practiced at a mature level (Elkonin 1977, 1978). To experience this zone of proximal development, children cannot persist in engaging in play frozen at the same level

throughout early-childhood years. Though children can experience challenge and growth during other play activities (such as building with blocks, or playing movement, board, or computer games), Vygotzky highlighted the uniquely social aspects of make-believe play when he declared it the leading activity of early childhood.

Despite the theoretical importance of mature play, it can be surprisingly difficult to find within contemporary early-childhood classrooms (Gudareva 2005; Levin 2008). Even five- and six-year-old children who, according to Vygotsky and Elkonin, should be at the peak of their play performance, often display signs of immature play more typical for toddlers and younger preschoolers: playing only with realistic props, enacting stereotypical and primitive play scenarios, and displaying a limited repertoire of themes and roles (Miller and Almon 2009; Smirnova and Gudareva 2004). Children who come to preschool with some play skills often do not acquire new skills by the end of the year, and they may even regress to less mature play (Farran and Son-Yarbrough 2001).

Although the decline of make-believe play may have many causes—including an increase in adult-directed forms of children's learning and recreation, the proliferation of toys and games that limit children's imagination, and the safety limits set by parents and teachers on where and how children are allowed to play (Chudacoff 2007)—the most important factor is likely the decrease in adult mediation of make-believe play (Karpov 2005). The idea that we need to teach young children how to play is not a new one. However, until recently, adult-scaffolded play has been discussed primarily in the context of special education. Although children with language delays or emotional disorders were thought to benefit from play interventions, typically developing children were expected to develop play skills on their own. Viewing play as a cultural phenomenon, however, we conclude that adult mediation will likely be a critical precursor to restoring children's historical levels of mature play.

We provide preliminary evidence that MPOT offers a reliable, classroom-based instrument positioned to help researchers and educators identify specific elements of mature play within early-childhood settings. Make-believe play provides an important and unique context that gives children opportunities to learn skills not afforded by other classroom activities. Too often, time for mature play is cut to accommodate more time for academic skills or incorporated only to add entertainment value to inherently boring and decontextualized drills. We suggest that play should instead be preserved and nurtured as

one of the "uniquely preschool" activities—in the words of Vygotsky's colleague and student Alexander Zaporozhets—that provide the most beneficial context for children's development.

## REFERENCES

Barnett, Lynn A. 1991. "The Playful Child: Measurement of a Disposition to Play." *Play and Culture* 4:51–74.

Barnett, W. Steven, Kwanghee Jung, Donald J. Yarosz, Jessica Thomas, Amy Hornbeck, Robert Stechuk, and Susan Burns. 2008. "Educational Effects of the *Tools of the Mind* Curriculum: A Randomized Trial." *Early Childhood Research Quarterly* 23:299–313. http://doi.org/10.1016/j.ecresq.2008.03.001.

Bates, Douglas, Martin Mächler, Benjamin M. Bolker, and Steven C. Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67:1–51. doi:10.18637/jss.v067.i01.

Bateson, Gregory. 1972. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology.*

Bergen, Doris. 2002. "The Role of Pretend Play in Children's Cognitive Development." *Early Childhood Research and Practice* 4. ERIC Number: ED464763.

Berk, Laura E., Trisha D. Mann, and Amy T. Ogan. 2006. "Make-Believe Play: Wellspring for Development of Self-Regulation." In *Play=Learning: How Play Motivates and Enhances Children's Cognitive and Social-Emotional Growth*, edited by Dorothy Singer, Roberta M. Golinkoff, and Kathy Hirsh-Pasek, 74–100.

Berk, Laura. E., and Adam Winsler. 1995. *Scaffolding Children's Learning: Vygotsky and Early Childhood Education.* NAEYC Research into Practice Series, Vol. 7.

Blair, Clancy, and Rachel Peters Razza. 2007. "Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten." *Child Development* 78:647–63.

Bodrova, Elena, Cynthia Björk, Crystal Day-Hess, Carrie Germeroth, Deborah Mazzeo, and S. Isaacs. 2012. "Scaffolding Early Learning: Strategies for Success." Unpublished manual, Mid-continent Research for Education and Learning (McREL), Denver, CO.

Bodrova, Elena, Carrie Germeroth, and Deborah J. Leong. 2013. "Play and Self-Regulation: Lessons from Vygotsky." *American Journal of Play* 6:111–23.

Bodrova, Elena, and Deborah J. Leong. 2007. *Tools of the Mind: The Vygotskian Approach to Early Childhood Education*. 2nd ed.

———. 2011. "Revisiting Vygotskian Perspectives on Play and Pedagogy." In *Rethinking Play and Pedagogy in Early Childhood Education: Concepts, Contexts, and Cultures*, edited by Sue Rogers, 60–72.

Bredekamp, Sue. 2005. "Play and School Readiness." *Educational Perspectives* 38:18–26.

Brown, Stuart, and Christopher Vaughan. 2009. *Play: How It Shapes the Brain, Opens the Imagination, and Invigorates the Soul.*

Bruner, Jerome S., Alison Jolly, and Kathy Sylva, eds. 1976. *Play: Its Role in Development and Evolution.*

Burghardt, Gordon. 2010. "The Comparative Reach of Play and Brain: Perspective, Evidence, and Implications." *American Journal of Play* 2:338–56.

Christie, James F., and Kathleen A. Roskos. 2006. "Standards, Science, and the Role of Play in Early Literacy Education." In *Play=Learning: How Play Motivates and Enhances Children's Cognitive and Social-Emotional Growth*, edited by Dorothy Singer, Roberta M. Golinkoff, and Kathy Hirsh-Pasek, 57–73.

Christie, James F., and Francis Wardle. 1992. "How Much Time Is Needed for Play?" *Young Children* 47:28–33.

Chudacoff, Howard P. 2007. *Children at Play: An American History.*

Clements, Douglas H., and Julie Sarama. 2013. *Building Blocks.* 2 vols.

———. 2014. "Play, Mathematics, and False Dichotomies." Preschool Matters. http://preschoolmatters.org/2014/03/03/play-mathematics-and-false-dichotomies/.

Clements, Douglas H., Julie Sarama, and Christopher B. Wolfe. 2011. *TEAM–Tools for Early Assessment in Math.*

Colwell, Malinda J., and Eric W. Lindsey. 2003. "Teacher-Child Interactions and Preschool Children's Perceptions of Self and Peers." *Early Child Development and Care* 173:249–58. http://doi.org/10.1080/03004430303096.

———. 2005. "Preschool Children's Pretend and Physical Play and Sex of Play Partner: Connections to Peer Competence." *Sex Roles* 52:497–509. http://doi.org/10.1007/s11199-005-3716-8.

Cowley, Judy, and Cheryl Glasgow. 1994. *The Renfrew Bus Story: Language Screening by Narrative Recall. Examiner's Manual.*

Diamond, Adele, W. Steven Barnett, Jessica Thomas, and Sarah Munro. 2007. "Preschool Program Improves Cognitive Control." *Science* 318:1387–88. http://doi.org/10.1126/science.1151148.

Diamond, Adele, and Kathleen Lee. 2011. "Interventions Shown to Aid Executive Function Development in Children 4 to 12 Years Old." *Science* 333:959–64. http://doi.org/10.1126/science.1204529.

Diamond, Adele, and Colleen Taylor. 1996. "Development of an Aspect of Executive Control: Development of the Abilities to Remember What I Said and to 'Do as I Say, Not as I Do.'" *Developmental Psychobiology* 29:315–34.

Dickinson, David K., and Patton O. Tabors. 2001. *Beginning Literacy with Language: Young Children Learning at Home and School.*

Elkonin, Daniil B. 1978. *Psychologija igry* [The psychology of play].

———. 2005a. "Chapter 1: The Subject of Our Research: The Developed Form of Play."

*Journal of Russian and East European Psychology* 43:22–48. http://doi.org/10.108 0/10610405.2005.11059242.

———.2005b. "The Psychology of Play." *Journal of Russian and East European Psychology* 43:11–21. http://doi.org/10.1080/10610405.2005.11059245.

Enz, Billie, and James F. Christie. 1993. "Teacher Play Interaction Styles and Their Impact on Children's Oral Language and Literacy Play." ERIC Number: ED366015.

———.1994. "Teacher Play Interaction Styles: Effects on Play Behavior and Relationships with Teacher Training and Experience." *International Journal of Early Childhood Education* 2:55–75.

Erikson, Erik H. 1975. "Play and Actuality." In *Explorations in Psychohistory: The Wellfleet Papers*, edited by Robert Jay Lifton and Eric Olson, 109–35.

Fantuzzo, John, Brian Sutton-Smith, Kathleen Coyle Coolahan, Patricia Holliday Manz, Sally Canning, and Darlena Debnam. 1995. "Assessment of Preschool Play Interaction Behaviors in Young Low-Income Children: Penn Interactive Peer Play Scale." *Early Childhood Research Quarterly* 10:105–20. http://doi.org/10.1016/0885-2006(95)90028-4.

Farmer-Dougan, Valeri, and Tami Kaszuba. 1999. "Reliability and Validity of Play-Based Observations: Relationship between the PLAY Behaviour Observation System and Standardised Measures of Cognitive and Social Skills." *Educational Psychology* 19:429–40. http://doi.org/10.1080/0144341990190404.

Farran, Dale C., and Whasoup Son-Yarbrough. 2001. "Title I Funded Preschools as a Developmental Context for Children's Play and Verbal Behaviors." *Early Childhood Research Quarterly* 16:245–62.

Fein, Greta G., and Shirley S. Schwartz. 1986. "The Social Coordination of Pretense in Preschool Children." In *The Young Child at Play,* Reviews of Research, Vol. 4, edited by Greta Fein and Mary Rivkin, 95–111.

Freud, Sigmund. 2003. *Beyond the Pleasure Principle*. Translated by John Reddick.

Gudareva, Olga V. 2005. "Psikhologicheskie Osobennosti Suzhetno-rolevoy Igry Sovremennykh Doshkol'nikov." [Psychological Features of Make-Believe Play in Today's Preschoolers]. PhD diss., Moscow City University for Psychology and Education.

Han, Myae, Noreen Moore, Carol Vukelich, and Martha Buell. 2010. "Does Play Make a Difference?: How Play Intervention Affects the Vocabulary Learning of At-Risk Preschoolers." *American Journal of Play* 3:82–105.

Hirsh-Pasek, Kathy, Roberta Michnick Golinkoff, Laura E. Berk, and Dorothy G. Singer. 2009. *A Mandate for Playful Learning in Preschool: Presenting the Evidence.*

Howes, Carollee. 1980. "Peer Play Scale as an Index of Complexity of Peer Interaction." *Developmental Psychology* 16:371–72. http://doi.org/10.1037/0012-1649.16.4.371.

Invernizzi, Marcia, Amie Sullivan, Joanne Meier, and Linda Swank. 2001. *PALS-PreK: Phonological Awareness Literacy Screening.*

Karpov, Yuriy V. 2005. "Three- to Six-Year-Olds: Sociodramatic Play as the Leading Activity during the Period of Early Childhood." In *The Neo-Vygotskian Approach to Child Development*, 139–70.

Kelly-Vance, Lisa, and Bridget O. Ryalls. 2005. "A Systematic, Reliable Approach to Play

Assessment in Preschoolers." *School Psychology International* 26:398–412. http://doi.org/10.1177/0143034305059017.

Kline, Theresa J.B. 2005. *Psychological Testing: A Practical Approach to Design and Evaluation.*

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H.B. Christensen. 2017. "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software* 82:1–26. doi: 10.18637/jss.v082.i13.

Leslie, Alan M. 1987. "Pretense and Representation: The Origins of 'Theory of Mind.'" *Psychological Review* 94:412–26. http://doi.org/10.1037/0033-295X.94.4.412.

Levin, Diane E. 2007. "Problem Solving Deficit Disorder: Programmed Play in Korea and the United States." In *Where Do The Children Play?: A Study Guide to the Film,* edited by Elizabeth Goodenough, 137–41.

Lewis, Vicki, and Jill Boucher. 1997. *The Test of Pretend Play*. Manual.

Lillard, Angeline S. 1993. "Pretend Play Skills and the Child's Theory of Mind." *Child Development* 64:348–71. http://doi.org/10.1111/j.1467-8624.1993.tb02914.x.

Lillard, Angeline S., Matthew D. Lerner, Emily J. Hopkins, Rebecca A. Dore, Eric D. Smith, and Carolyn M. Palmquist. 2013. "The Impact of Pretend Play on Children's Development: A Review of the Evidence." *Psychological Bulletin* 139:1–34.

Linder, Toni W. 1993. *Transdisciplinary Play-Based Assessment: A Functional Approach to Working with Young Children*. 2nd ed.

Lindsey, Eric W., and Malinda J. Colwell. 2013. "Pretend and Physical Play: Links to Preschoolers' Affective Social Competence." *Merrill-Palmer Quarterly* 59:330–60. http://doi.org/10.1353/mpq.2013.0015.

Lindsey, P., and L. Soares. 1998. Review of *Test of Pretend Play* by Vicky Lewis and Jill Boucher. In *The Fifteenth Mental Measurements Yearbook, edited by* Barbara S. Plake, James C. Impara, and Robert A. Spies.

Luria, Aleksandr R. 1966. *Higher Cortical Functions in Man*. Translated by Basil Haigh.

Lyytinen, Paula, Marja-Leena Laakso, Anna-Maija Poikkeus, and Niina Rita. 1999. "The Development and Predictive Relations of Play and Language across the Second Year." *Scandinavian Journal of Psychology* 40:177–86. http://doi.org/10.1111/1467-9450.00115.

Manuilenko, Z. V. 1975. "The Development of Voluntary Behavior in Preschool-Age Children." *Soviet Psychology* 13:65–116.

McAloney, Katherine, and Karen Stagnitti. 2009. "Pretend Play and Social Play: The Concurrent Validity of the Child-Initiated Pretend Play Assessment." *International Journal of Play Therapy* 18:99–113.

McClelland, Megan M., Claire E. Cameron, Carol McDonald Connor, Carrie L. Farris, Abigail M. Jewkes, and Frederick J. Morrison. 2007. "Links between Behavioral Regulation and Preschoolers' Literacy, Vocabulary, and Math Skills." *Developmental Psychology* 43:947–59. http://doi.org/10.1037/0012-1649.43.4.947.

McCune, Lorraine. 1995. "A Normative Study of Representational Play in the Transition to Language." *Developmental Psychology* 31:198–206. http://doi.org/10.1037/0012-1649.31.2.198.

Miller, Edward, and Joan Almon. 2009. "Crisis in the Kindergarten: Why Children Need to Play in School." *Alliance for Childhood.* ERIC Number: ED504839.

National Research Council. 2008. *Early Childhood Assessment: Why, What, and How*, edited by Catherine E. Snow and Susan B. Van Hemel.

Neuman, Susan B., and Kathy Roskos. 1992. "Literacy Objects as Cultural Tools: Effects on Children's Literacy Behaviors in Play." *Reading Research Quarterly* 27:203–25. http://doi.org/10.2307/747792.

Nicolich, Lorraine McCune. 1977. "Beyond Sensorimotor Intelligence: Assessment of Symbolic Maturity through Analysis of Pretend Play." *Merrill-Palmer Quarterly of Behavior and Development* 23:89–99.

Paolitto, A. W., and H. N. Switzky. 1995. Review of *Symbolic Play Test*, 2nd ed., by Marianne Lowe and Anthony J. Costello. In *The Twelfth Mental Measurements Yearbook,* edited by Jane Close Conoley and James C. Impara.

Parten, Mildred B. 1932. "Social Participation among Preschool Children." *The Journal of Abnormal and Social Psychology* 27:243–69.

Pellegrini, Anthony D., ed. 2011. *The Oxford Handbook of the Development of Play*.

Piaget, Jean. 1962. *Play, Dreams, and Imitation in Childhood*.

Ponitz, Claire E. Cameron, Megan M. McClelland, Abigail M. Jewkes, Carol McDonald Connor, Carrie L. Farris, and Frederick J. Morrison. 2008. "Touch Your Toes! Developing a Direct Measure of Behavioral Regulation in Early Childhood." *Early Childhood Research Quarterly* 23:141–58. http://doi.org/10.1016/j.ecresq.2007.01.004.

Power, Thomas S., and Jerilynn Radcliffe. 1991. "Cognitive Assessment of Preschool Play Using the Symbolic Play Test." In *Play Diagnosis and Assessment*, edited by Charles E. Schaefer, Karen Gitlin, and Alice Sandgrund, 87–111.

Raver, C. Cybele, and Jane Knitzer. 2002. "Ready to Enter: What Research Tells Policymakers about Strategies to Promote Social and Emotional School Readiness among Three- and Four-Year-Olds." National Center for Children in Poverty. https://academiccommons.columbia.edu/doi/10.7916/D82V2QVX.

Roskos, Kathy, and Susan B. Neuman. 1993. "Descriptive Observations of Adults' Facilitation of Literacy in Young Children's Play." *Early Childhood Research Quarterly* 8:77–97. http://doi.org/10.1016/S0885-2006(05)80099-7.

———. 1998. "Play as an Opportunity for Literacy." In *Multiple Perspectives on Play in Early Childhood Education*, edited by Olivia N. Saracho and Bernard Spodek, 100–15.

Rothbart, Mary K., Michael I. Posner, and Jessica Kieras. 2006. "Temperament, Attention, and the Development of Self-Regulation." In *The Blackwell Handbook of Early Childhood Development,* edited by Kathleen McCartney and Deborah Phillips, 338–57.

Rubin, Kenneth. 2008. *The Play Observation Scale - Revised*. College Park: University of Maryland.

Rubin, K. H. (2008). *The Play Observation Scale –Revised*. College Park: University of Maryland.

Saarni, Carolyn, Donna L. Mumme, and Joseph J. Campos. 1998. "Emotional Development: Action, Communication, and Understanding." In *Handbook of Child Psychology: Social, Emotional, and Personality Development*, edited by William Damon and Nancy Eisenberg, 237–309. 5th ed.

Saifer, Steffen. 2009. "Higher-Order Play: A Reconceptualization of Play in Development and Early Education from a Neo-Vygotskian Perspective." Paper presented at the 18th annual National Association for the Education of Young Children's National Institute for Professional Development in Charlotte, North Carolina.

Sarama, Julie, and Douglas H. Clements. 2004. "Building Blocks for Early Childhood Mathematics." *Early Childhood Research Quarterly* 19:181–89. http://doi.org/10.1016/j.ecresq.2004.01.014.

Singer, Dorothy G., Roberta Michnick Golinkoff, and Kathy Hirsh-Pasek, eds. 2006. *Play=Learning: How Play Motivates and Enhances Children's Cognitive and Social-Emotional Growth*.

Smilansky, Sara, and Leah Shefatya. 1990. *Facilitating Play: A Medium for Promoting Cognitive, Socio-Emotional, and Academic Development in Young Children*.

Smith, Peter K. 2010. "Observational Methods in Studying Play." In *The Oxford Handbook of the Development of Play*, edited by Anthony D. Pellegrini, 138–49.

Smirnova, Elena O., and Olga V. Gudareva. 2004. "Igra i proizvol'nost u sovremennykh doshkol'nikov" [Play and intentionality in modern preschoolers]. *Voprosy Psychologii* 1:91–103.

Snow, Catherine E., and Susan B. Van Hemel, eds. 2008. *Early Childhood Assessment: Why, What, and How*.

Stagnitti, Karen. 2007. *The Child Initiated Pretend Play Assessment (ChIPPA)*.

Stanton-Chapman, Tina L. 2015. "Promoting Positive Peer Interactions in the Preschool Classroom: The Role and the Responsibility of the Teacher in Supporting Children's Sociodramatic Play." *Early Childhood Education Journal* 43:99–107. http://doi.org/10.1007/s10643-014-0635-8.

Taylor, Marjorie, and Stephanie M. Carlson. 1997. "The Relation between Individual Differences in Fantasy and Theory of Mind." *Child Development* 68:436–55. http://doi.org/10.2307/1131670.

Thibodeau, Rachel B., Ansley T. Gilpin, Melissa M. Brown, and Brooke A. Meyer. 2016. "The Effects of Fantastical Pretend-Play on the Development of Executive Functions: An Intervention Study." *Journal of Experimental Child Psychology* 145:120–38.

Toub, Tamara Spiewak, Brenna Hassinger-Das, Kimberly Turner Nesbit, Hande Ilgaz, Deena Skolick Weisberg, Kathy Hirsh-Pasek, Roberta Michnick Golinkoff, Ageliki Nicolopoulou, and David K. Dickinson. 2018. "The Language of Play: Developing Preschool Vocabulary through Play Following Shared Book-Reading." *Early Childhood Research Quarterly* 45:1–17. https://doi.org/10.1016/j.ecresq.2018.01.010.

Trawick-Smith, Jeffrey. 1998. "A Qualitative Analysis of Metaplay in the Preschool Years." *Early Childhood Research Quarterly* 13:433–52.

Vygotsky, Lev S. 1967. "Play and Its Role in the Mental Development of the Child."

*Soviet Psychology* 5:6–18.

———. 1978. *Mind in Society: The Development of Higher Mental Processes.*

———. 1987. "Thinking and Speech." In *The Collected Works of L.S. Vygotsky*, Vol. 1, edited by Robert W. Rieber and Aaron S. Carton, 375–84.

Wechsler, David. 2003. *WISC-IV: Administration and Scoring Manual.*