University of Reading

A scoping review of methods and measures used to capture children's play during school breaktimes

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

FitzGibbon, L., Oliver, B., Nesbit, R. and Dodd, H. (2025) A scoping review of methods and measures used to capture children's play during school breaktimes. Educational Review, 77 (6). pp. 1991-2016. ISSN 1465-3397 doi: 10.1080/00131911.2024.2306944 Available at https://centaur.reading.ac.uk/115545/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1080/00131911.2024.2306944

Publisher: Routledge

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Educational Review

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/cedr20

A scoping review of methods and measures used to capture children's play during school breaktimes

Lily FitzGibbon, Brooke Oliver, Rachel Nesbit & Helen Dodd

To cite this article: Lily FitzGibbon, Brooke Oliver, Rachel Nesbit & Helen Dodd (15 Feb 2024): A scoping review of methods and measures used to capture children's play during school breaktimes, Educational Review, DOI: 10.1080/00131911.2024.2306944

To link to this article: https://doi.org/10.1080/00131911.2024.2306944

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 15 Feb 2024.

Submit your article to this journal

Article views: 223

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=cedr20

EDUCATIONAL REVIEW
https://doi.org/10.1080/00131911.2024.2306944

Routledge
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

A scoping review of methods and measures used to capture children's play during school breaktimes

Lily FitzGibbon[a], Brooke Oliver[b], Rachel Nesbit[c] and Helen Dodd[c]

[a]Division of Psychology, University of Stirling, Stirling, UK; [b]School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK; [c]Medical School, University of Exeter, Exeter, UK

ABSTRACT

Play is linked to healthy child development and is recognised in the UN Convention on the Rights of the Child. School breaktimes provide regular opportunities for children to play, and as such, they have been the context of a large and interdisciplinary body of research on play. Play research has diverse aims and cuts across many academic disciplines, resulting in a wide range of methods and measurement tools being used in research to capture children's play. In this scoping review, 105 studies of play during school breaktimes were identified and we describe, synthesise and compare methods used to assess play during school breaktimes, bringing together methodologies from different fields for the first time. Specifically, we captured: the aspects of play that have been measured and described; established tools and coding schemes that have been used; what the measures of play have been used for; and what the quality of reporting of play measures has been. In this way, we anticipate that the review will facilitate future play research and support, where appropriate, more consistent use and transparent reporting of methods and measures.

ARTICLE HISTORY
Received 11 November 2022
Accepted 14 January 2024

KEYWORDS
Play; school; breaktime; recess; methodology; measures

Introduction

The importance of play for many aspects of children's lives is increasingly recognised (Lester & Russell, 2010; Yogman et al., 2018) and is enshrined in the UN Convention on the Rights of the Child (UN General Assembly, 1989). For example, play provides myriad opportunities for cognitive and social development (Andersen et al., 2023; Singer et al., 2006) and is important for children's physical (Herrington & Brussoni, 2015; Nijhof et al., 2018) and mental health (Dodd et al., 2022; Whitebread, 2017). Schools are an important context for children's play, with most schools offering some opportunity for play during scheduled breaks in the school day. Formal schooling provides a unique opportunity to address inequity in children's access to quality time and

CONTACT Lily FitzGibbon lily.fitzgibbon@stir.ac.uk Division of Psychology, University of Stirling, Stirling FK9 4LA, UK @FitzyLily

Supplemental data for this article can be accessed online at https://doi.org/10.1080/00131911.2024.2306944

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

space for play (London, 2019). It has been notoriously difficult to define play because of its complexity and ambiguity. Deciding where play begins and ends, what constitutes play and what does not, and how play should be categorised or taxonomised is both challenging and subject to disagreement between researchers (Eberle, 2014; Smith, 2009). This means that the nature and quality of play in schools (and outside of schools) can be difficult to evidence. In this article, our aim is to conduct a scoping review of measures used to assess play during school breaktimes. The review will highlight the range of measures used in the research literature, consider any gaps and make initial suggestions about how the use and reporting of these measures might be improved.

For the purpose of this review, we consider play as an inclusive, umbrella term for the activities and occupations that children choose to engage in for the purpose of enjoyment and recreation rather than for any practical purpose. This broad definition is intentional, with a view to providing an inclusive review of measurement in this area of research. Our focus in this review is on play during school breaktimes. We use the term "breaktime" to include all breaks during the school day including lunchtimes and morning and afternoon breaks between classes. Breaktimes are typically periods of unstructured activity between formal classes in which children may have access to outdoor space and opportunities to interact freely with their peers. These breaks are variously called "playtime", "recess" and "breaktime" across different schools, age-groups and countries, and our search strategy was designed to be inclusive of these different naming conventions (see Method).

Within the school environment, breaktimes have been identified as an important contributing factor to children's classroom behaviour (Jarrett et al., 1998; Massey et al., 2021) and academic achievement (Pellegrini & Bjorklund, 1997; Pellegrini & Bohn, 2005). Furthermore, both children and teachers value breaktimes as vital opportunities for children to socialise and have free time for undirected recreation (Baines & Blatchford, 2019; Evans, 1996). Despite this, educators feel that there is increasing pressure to focus on academic achievement, resulting in the reduction of breaktime by an average of between 45 and 65 minutes per week between 1995 and 2018 in Britain (Baines & Blatchford, 2019). Similarly, in the United States during the 2000s, many school districts reduced the time that elementary school children spent in recess in order to focus on core academic subjects (Henley et al., 2007). Policies protecting adequate and equitable access to recess vary substantially both across and within countries (for example, across the United States; Clevenger et al., 2022).

Public policy, such as the Office for Standards in Education, Children's Services and Skills (OFSTED) framework for assessing UK schools, tends to acknowledge the importance of play for preschool-aged children, but provides minimal guidance regarding the provision of play for school-aged children (OFSTED School Inspection Handbook, 2019). Even where provision for play is mandated in policy (such as the Play Sufficiency Duty in Wales, 2012), schemes for evaluating play are rarely provided, making it challenging for stakeholders to assess their own practice, or judge the effectiveness of novel interventions for improving the quality of school play provision.

In recent years, many programmes have been introduced to improve play provision for children both within and outside of the school context (e.g. Brussoni et al., 2017; Bundy et al., 2017). These programmes often aim to improve the physical and mental health of children through improving play opportunities. In order to properly understand the

impact of these programmes on children's play and the mechanisms through which improvements in health may be achieved, it is essential to have reliable and valid measures of the quality, quantity and content of children's play in schools. Having appropriate measures of play in schools also benefits those designing play space and equipment for play as well as research examining how play changes over time, or in relation to specific events (e.g. Covid-19; natural disasters; policy changes).

Existing measures of play come from diverse fields including town planning; architecture and urban design; education; psychology; public health; and risk management and injury prevention. Often play research lies at the intersection of these fields, involving interdisciplinary collaborations. Perhaps due to this diversity of perspectives as well as the complexity of operationalising play, a wide range of methods and measurement tools have been used in research to capture children's play. For researchers, particularly those new to the field or those working across disciplines, this vast array of approaches can be confusing and potentially prohibitive of good quality research.

The aim of this paper is to conduct a scoping review of measures used to assess play during school breaktimes. Previous reviews of children's breaktime activities have focused solely on physical activity (e.g. Parrish et al., 2020; Ridgers, et al., 2012) or observation measures (Leff & Lakin, 2005). This review brings together methodologies from different fields for the first time with a view to improving understanding of the range of behaviours that can be measured, and the diversity of tools available for capturing different perspectives (i.e. child, teacher, observer). Specifically, we aimed to capture: the aspects of play that have been measured and described; whether there are established tools or coding schemes that have been used; what the measures of play have been used for; and what the quality of reporting of play measures has been. Assessing the quality of the measures themselves is beyond the scope of this article, in part because of the diversity of methodological approaches employed. This kind of critical appraisal of the methods themselves is an important next step for the field, but we believe this would be better suited to reviews focused on studies tackling a single research question or using a single methodological approach in which the same criteria could be used consistently.

Methods

The scoping review was conducted in accordance with the JBI methodology for scoping reviews (Peters et al., 2015) following the Preferred Reporting Items for Systematic Reviews and Meta-analyses extension for Scoping Review (PRISMA-ScR) (Tricco et al., 2018). The protocol including search terms and strategy for this scoping review was pre-registered on the Open Science Framework (OSF; https://osf.io/ftwuk) and updated after a pilot of the full text screening (https://osf.io/83bx4).

Search strategy

The search strategy aimed to identify both published studies and those reported in doctoral dissertations. An initial limited search of PubMed, and ERIC was undertaken to identify articles on the topic. The words contained in the titles, abstracts, author

keywords or full texts of relevant articles, and the index terms used to describe the articles were used to develop a full search strategy for PubMed, PsycINFO, Web of Science, ProQuest's Dissertations database, ERIC and BEI. Grey literature (with the exception of PhD theses) was not included due to feasibility given the number of papers identified for inclusion.

Databases were searched for articles published in English that included "play" in the keywords, database indexing, or MeSH terms, "school" (with truncation) in the title or abstract, and for terms relating to breaktime in all fields (recess, break, lunch, playground and playtime, with truncation to capture plurals and longer forms such as lunchtime). The choice to restrict search results to articles with "play" in the keywords, database indexing, or MeSH terms was a pragmatic one because the word "play" is frequently used when describing the relationship between variables, often in the title or abstract. This resulted in many irrelevant articles being returned. Searchable keywords vary between databases so that for some the author keywords are used and for others the database's descriptors (e.g. MeSH terms) are used. Searches were specified slightly differently for each database dependent on available search fields and syntax. The database searches were conducted on the 28th January 2021. The search was not limited to a specific time period – no start date was specified.

Eligibility criteria

Participants

Studies were required to include at least one measure of play in schools. Because school age differs by location, no specific age eligibility criteria were included. Additionally, although our search terms did not explicitly include the early-years or preschool contexts, because this was not our focus, we adopted an inclusive approach and, if our search identified these studies they were included.

Concept

Studies were required to be empirical and to report the methodology used to measure, describe and understand play. Play was required to be a primary outcome of at least one of the study measures. Note that this does not mean that play needed to be the primary outcome of the *study*. We did not take a strict definition of play in this review, rather we summarised methods and measures that were described as measuring play by the original authors. Our aim was to be inclusive of different definitions of play and of specific types of play being studied. This means that within the articles captured in the review, there may be differences as to which behaviours could be considered as play. For example, some authors may consider engagement in rule-based games as play, whereas others may not. An additional criterion was added after a pilot of the full-text review phase. We noted that while many studies discussed play in the introduction and discussion sections of the report, the methods did not explicitly describe a measure of play, but instead described an independent construct (e.g. social interactions) during breaktime. Thus, we included the criterion that reported measures must be described as measures of play in the method section of the manuscript (or in the description of the methodology if the manuscript did not have an explicit method section). Studies could include general measures of play or measures of

specific types of play, such as imaginary play, creative play, play with peers and active outdoor play.

Context

This review is focused on research conducted in the school context. Specifically, research that measured play during breaks in the school day (i.e. recess), between classes or lunchtimes. This context was chosen because it is when free play is most likely to happen during school and it is typically the target of interventions on children's play (Bundy et al., 2017), although we note that free play does not happen during all school breaktimes, such as when there are structured, adult-led activities during breaktimes.

Types of sources

This scoping review considered quantitative and qualitative research with experimental and observational study designs. Systematic reviews, meta-analyses, text and opinion papers were not included.

Study/source of evidence selection

Following the search, all identified citations were collated and uploaded into Covidence's online review platform (Covidence Systematic Review Software, n.d.). Duplicates were removed automatically by Covidence and, following a pilot test, titles and abstracts were screened by one lead reviewer for assessment against the inclusion criteria for the review. One additional reviewer independently screened 20% of the abstracts. Agreement between reviewers was weak (82% agreement; $\kappa = .46$). For the majority of disagreements (81%), the lead reviewer voted to move the article to the full text review while the additional reviewer voted to exclude the article, suggesting the lead reviewer took a more liberal approach at the screening phase when it was unclear whether the inclusion criteria had been met. Those articles that passed the initial screening phase were retrieved in full. The full text of selected articles was assessed in detail against the inclusion criteria by one reviewer with an additional reviewer screening 20%. Agreement between the reviewers was moderate (82% agreement; $\kappa = .61$). Disagreements that arose between the reviewers at each stage of the selection process were resolved through discussion and consultation with the senior author.

The weak to moderate agreement between reviewers reflects the difficulty assessing whether studies report methodology to measure play, as opposed to a related construct. One source of conflicting decisions was that the same measures were sometimes reported as measures of play and sometimes as measures of physical activity (e.g. SOPLAY; McKenzie et al., 2000). Another source was that play could come out strongly through thematic or content analysis of qualitative data without play having been mentioned in the methods. Finally, in a small proportion of studies, the quality of the description of the methods was poor, making it difficult to determine whether play was being measured.

Reasons for exclusion of sources of evidence at the full text stage that did not meet the inclusion criteria are presented in a PRISMA-ScR flow diagram in Figure 1 (Tricco et al., 2018).

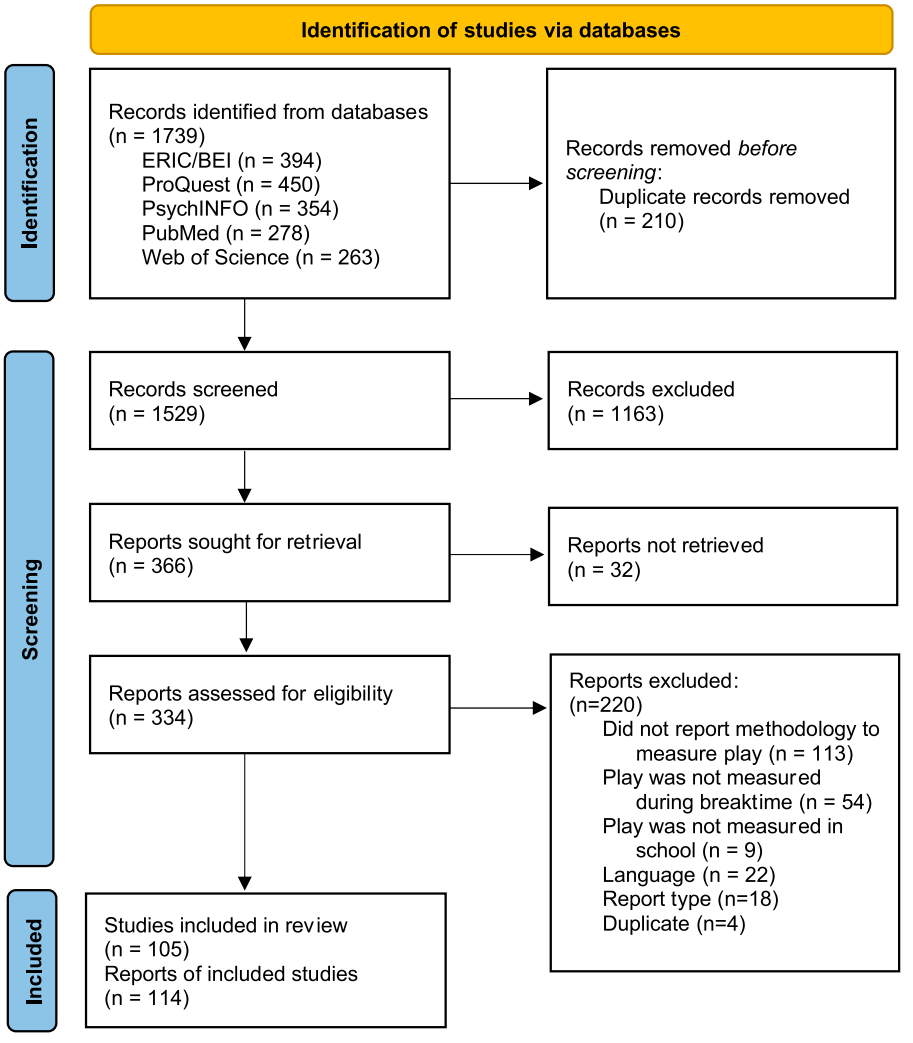


Figure 1. PRISMA-ScR flow diagram for identification of studies.

Data extraction

Data were extracted by hand from reports included in the scoping review by one reviewer, with data recorded in a data extraction form developed by the authors within the Covidence software. A series of questions, some with open-text answers and others with pre-specified choices were used to extract the data from each report. The full form can be found in the supplementary materials. Data were extracted from 114 reports describing 105 unique studies. Where more than one report described the same study, the reports were "merged" in Covidence and data was extracted from all reports at once. The data extracted included information about the participants, the context, the aspects of play being captured, and study methods and outcome measures including whose perspective was taken and how play was measured. The purpose of the measure (e.g. assessment of intervention) was also recorded.

The draft extraction form provided with the protocol document was revised to ensure that it accurately captured the range of studies, and to facilitate the data extraction process. Any ambiguities in the data extraction process were discussed with the other authors. Authors of papers were not contacted for specifics where these were not detailed in the manuscript because of the volume of included reports. While we did not conduct a critical appraisal of the methods employed in each study, we extracted information about reporting of the validity and reliability of measures, and the transparency of reporting more generally.

The supplementary materials (available here: https://osf.io/8rga6/?view_only=859bc5b0037746c4a8e6aa8885a072d0) include the data extraction form, the complete list of included reports, the extracted data reported in this article, and a metadata file describing each of the columns in the extracted data file. Open responses to the data extraction form questions have been removed from the data file because many of these were taken verbatim from the original manuscripts or consist of the first author's personal notes. Where possible, the data from the open responses have been consolidated into categories where pertinent themes emerged.

Results

105 studies reported in 114 reports were included in the review (see the supplementary materials for a full list). The majority were reported in journal articles (80%) and PhD theses (17%) or both (1%). The remaining studies were reported in conference proceedings (1%) and in a report (1%) for the US National Institute of Education. Reports included in the review were published between 1976 and 2020. The field of study was determined by examining the journal scope and/or the authors' departmental affiliations. Most studies were from the fields of education (40%) or psychology (29%) or a combination of the two (5%). Another substantial portion of the studies were from the field of public health (15%). The remaining studies (11%) were from diverse fields including architecture and planning, sociology, anthropology, social work, occupational therapy, and linguistics and language therapy, or interdisciplinary combinations of the above fields.

Study context and participants

The vast majority (95%) of studies were conducted in the Global North, with only 4% conducted in the Global South including Kuwait, Malaysia, the Philippines and Brazil.[1] A further study, conducted in Saint Helena, could not easily be characterised as Global North or South. Most of the studies were based in mainstream primary (elementary) schools, either exclusively (64%), or in combination with early years settings (2%), secondary (middle or high) schools (8%), or special educational needs (SEN) schools (4%). A further 15% of studies were conducted exclusively in early years settings, and 6% in SEN schools. Two further studies (2%) did not report the type of school, but these are assumed to have been conducted in mainstream primary/elementary schools based on the age and description of the children included. Many studies took place across more than one school (55%), with 22% conducted in two or three schools; and 33% in four or more (up to 26) schools. However, a large proportion of studies were conducted in just one school (42%), and 3% did not report the number of schools.

Most studies reported the sample size (93%), although this was sometimes approximate, for example, where scan observations of the whole playground took place. The sample size ranged from a small case study of two children to a large multi-site study of several thousand children. The age of the children included in each study was extracted from the original articles where possible (96%), but for some studies the age range was inferred from the school years included. The ages of children in the studies ranged from 1 year old to 15 years old. Most studies (63%) included only children aged between 5 and 12, with an additional 23% including children younger than 5, and 10% including children older than 12. The majority of studies (66%) did not report the makeup of the study population in terms of neurodiversity or physical disability. A further 4% of studies reported that the study population consisted of typically developing children. The remaining studies reported samples of children exclusively with special educational needs (13%), mixed cohorts (14%) including children with special educational needs and/or physical disabilities as well as typically developing children, and studies that sampled other specific populations (3%) such as "aggressive children".

Methods used to measure play during school breaktimes

For each study we coded whether the following methodological approaches were used: observation (structured and unstructured), questionnaires, qualitative and participatory methods (aside from unstructured observations), and physiological measures such as accelerometers, heart rate monitors, radio-frequency identification (RFID) tags that record instances of close proximity between children, and GPS trackers. The majority of

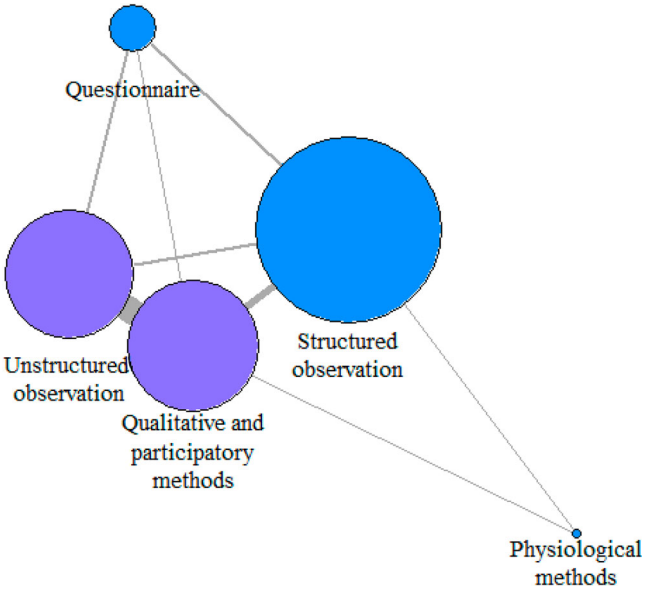


Figure 2. Graph analysis demonstrating frequency and co-occurrence of methodological approaches. Node size represents the frequency of each methodological approach. Edges (lines) between nodes represent that at least one study has used these approaches in combination, and the weight of the edges represents the frequency of each combination. Qualitative approaches are presented in purple (dark grey in print) and quantitative approaches in blue (light grey in print).

studies used a single method (66%) to study play. The remaining studies used multiple methods, either within a single methodological approach (7%), such as using multiple coding schemes, questionnaires, or participatory methods, or a combination of approaches (28%). A graph analysis represented in Figure 2 visualises the frequency and co-occurrence of these methodological approaches. Studies that employed more than one method ranged from using two methods to using 10 different methods including observational, qualitative, and participatory approaches. In the following paragraphs, we describe these methodological approaches and the perspectives that they can bring to the study of play during school breaktimes.

The most common methods for measuring play were observational (87%), of which, 30% used a combination of methodological approaches. Observations were structured (68%) using pre-specified coding schemes or unstructured (30%) using qualitative methods to describe and explain the observed behaviours after the fact. Further information about established coding systems is reported later in the results and in Table 1. Unstructured and structured observations were rarely combined (2%). Across structured and unstructured observations, data were collected in a range of ways, including field notes, video recordings, audio recordings and live coding, or a combination of these.

In 96% of observational studies, the observations were carried out by one or more researcher, who in one study was also a teacher. In the remaining studies, observations were carried out by clinicians (1%), a teacher (1%), both researchers and children themselves (1%), or was not reported (1%). Within the structured observations, these varied as to whether they were "person-based" (63%), following a specific individual and recording their play, "group-based" (5%), following a pair or group of children and recording their play, or "place-based" (23%), scanning across an area of the playground and recording the play of the children in that area. A further 5% used a combination of these approaches, and for 5% it was not clear from the manuscript whether individual children or areas of the playground were observed systematically. Within the unstructured observations, 41% reported some "person-based" observation, in which specific children were observed over a period of time.

Other qualitative and participatory methods such as interviews, focus groups, and walking tours were used in 26% of the included studies, of which 89% used a combination of methodological approaches. Children were the most frequent subjects of these methods either alone (67%), or in combination with teachers and breaktime supervisors or parents (19%). A further 15% of these studies only used teachers as the subject of the qualitative or participatory methods. Within this broad methodological approach, many creative methods were used to enable rich narratives about children's play during breaktimes, often including the child's own perspective. For example, in several studies, children were asked to draw pictures or maps, or annotate plans of the playground to supplement verbal responses. Walking tours of the playground and video diaries made by the children allowed children to explain in their own words and actions what happens where on the playground as well as the affordances of the space for certain preferred or disliked activities. These video diaries and other recordings from the playground were also used to elicit responses and discussion in interviews and focus groups.

Questionnaire or survey methods (13%) were used less often. The respondents in all but one of the studies using questionnaires were children, with just one study surveying teachers. As well as conventional questionnaires with Likert-scale responses, questionnaire

Table 1. Named observation systems used to capture play. Note that these structured observation systems all give rise to quantitative data.

Name	Original reference	Approach	Description	Aspect of play captured	Number of included studies using scheme
Imaginativeness of Play Scale	Singer and Singer (1981)	Person-based	Observational system for determining the imaginativeness of children's play.	Quality (depth)	1
Observation of Playground Play (OPP)	Massey et al. (2018) Adapted from the Activities of Daily Living – Playground Participation (ADL-PP; Stellino & Sinclair, 2014)	Place-based	Observation system to capture the games, activities and sedentary behaviours that children engage in on the playground.	Quantity (play types; activities)	2
Peer Play Scale[a]	Bernardo (1995)	Person-based	Observation system for categorising play into types.	Quantity (play types)	1
Play Observation Scale	Rubin (1989; 2001)	Person-based	Observation system for categorising play into cognitive and social play types and assessing the valance of social interactions in play.	Quantity (play types) and quality (valence)	3
Playground Behaviour Observation Schedule	Charlton et al. (1998)	Person-based	Observation system for recording the occurrence of playground behaviours and group size.	Quantity (play types; activities)	1
Playground Code	Pepler et al. (1998) Adapted from Rusby and Dishion (1990)	Person-based	Observational system for assessing interactions and affect during play.	Quantity (play types) and quality (enjoyment)	1
Playground Observation Checklist	Ingram et al. (2007)	Person-based	Observational system to determine whether neurodiverse and typically developing children differ in their playground behaviour.	Quantity (activities)	1
Playground Observation of Peer Engagement (POPE)	Kasari et al. (2011)	Person-based	Observation system to capture the way that children interact with each other on the playground and engagement in activities.	Quantity (play types)	4
System for Observing Children's Activity and Relationships during Play (SOCARP)	Ridgers et al. (2010)	Person-based	Observation system to capture levels of physical activity, social behaviour and types of play on the playground.	Quantity (play types)	5
System for Observing Outdoor Play (SOOP)	Engelen et al. (2018) Adapted from SOPLAY (McKenzie et al., 2000)	Place-based	Observation system to capture activity level, activity type and equipment use of groups of children on the playground.	Quantity (play types; activities)	1

(*Continued*)

Table 1. Continued.

Name	Original reference	Approach	Description	Aspect of play captured	Number of included studies using scheme
System for Observing Play Engagement Episodes at Recess (SOPEER)	Khorana (2017) Adapted from SOCARP (Ridgers et al., 2010)	Place-based	Observation system to capture social behaviour on the playground.	Quality (valence)	1
System for Observing Play and Leisure in Youth (SOPLAY)	McKenzie et al. (2000)	Place-based	Observation system designed to capture levels of physical activity and predominant activities on the playground. Frequently extended to include further play types.	Quantity (play types; activities)	6
Test of Playfulness	Bundy et al. (2001)	Person-based	Observational system for assessing playfulness in children's playground behaviour.	Quality (depth)	1
Toddler Peer Play Code	Jamison (2010)	Person-based	Observational system for categorising social play types, proximity and joint attention as well as teacher interventions.	Quantity (play types)	1

[a]The description of this scale is limited in the original manuscript. The original author notes that although intended as an observational system, it was not used for this purpose due to practical difficulties using the scheme.

Table 2. Named questionnaires used to capture play.

Name	Reference	Respondent	Description	Outcomes	Number of included studies using scheme
Basic Psychological Needs Satisfaction Scale (BPNS)	Stellino and Sinclair (2013)	Child	21 items examining children's need satisfaction during recess.	Play quality (enjoyment; affordance)	1
Children's Attitudes Towards Play (CATP)	Greer and Stewart (1989) adapted from Webb (1969)	Child	3 items assessing children's motivation for play across three contexts.	Other: child's motivation	1
Lunchtime Enjoyment of Activity and Play (LEAP)	Hyndman (2015)	Child	39 items examining how much children enjoy school play and lunchtime activity.	Play quality (enjoyment)	2
Penn Interactive Peer Play Scale (PIPPS)	Fantuzzo et al. (1995)	Teacher (parent version also available)	32 items assessing interaction, disruption and disconnection from play.	Play quality (valence)	1

methods were adapted for children by using pictorial scales, having items read to the children, and allowing children to respond freely to open-ended questions about their play. Further information about established questionnaire measures is reported later in the results and in Table 2.

The use of physiological methods such as accelerometers, heart rate monitors, RFID tags, and GPS trackers to measure play was rare, with just two studies (2%) using these methods, and both using accelerometers. We acknowledge that these measures are regularly used to capture other aspects of playground behaviour such as social interaction or physical activity *during* play but these are outside the scope of the current review because they are not used to measure play per se (e.g. Fjørtoft et al., 2009; Heravi et al., 2018).

Aspects of play captured

To address what aspects of play were being captured by methods used across studies, we explored the outcomes of the measures of play for each study. Some studies focused on one aspect of play while others included multiple outcomes, so the percentages reported in the following sections are not mutually exclusive.

The most common outcome was a measure of the quantity of play (55%). Within these studies, count, frequency, or duration measures captured the quantity of either (a) play in general (e.g. vs. non-play; 10%), (b) specific types of play such as pretend play or solitary play (62%), (c) specific games or activities during play (17%), or combinations of the above (9%). A further 2% of studies included quantity measures that did not fit into the above categories. The types of play and specific activities and games that were captured are described further below. Most of the quantity measures were derived from structured observation, but in a minority of cases, questionnaire measures assessed the frequency of play. Both person-based and place-based observations can lead to quantitative measures of play. Person-based observations would typically provide frequency or duration measures that represent how much that child engaged in play, in types of play, or specific play activities. In contrast, place-based observations would typically provide counts of the number of children engaging in play, types of play, or specific play activities for each scan of the playground or area thereof.

A second outcome was play quality (22%). Measures of play quality could be broadly categorised into four non-mutually exclusive themes: *enjoyment* (41%) – how much children reported or appeared to be enjoying or appeared to enjoy playing during school breaktimes; *valance* (41%) – the extent to which play was positive or prosocial rather than negative or antisocial[2]; *depth* (19%) – the level of engagement, creativity, or imaginativeness of the play itself; and *affordance* (15%) – the opportunities for, or conversely, restrictions on children's play or types of play. Play quality was captured by a broad range of methods including structured and unstructured observations, questionnaires, and qualitative and participatory methods such as interviews and playground tours. Children's perspectives were sought in the majority of studies assessing play quality using questionnaire (89%) and qualitative and participatory methods (57%). Teachers' perspectives were also sought when assessing play quality by questionnaire (11%) and qualitative and participatory methods (57%). It is noteworthy that many of these assessments (i.e. observation and adult-report) require significant inference regarding children's internal experiences. Different programmes of research handle this subjectivity in different ways; this is discussed at some length in the discussion section.

A third outcome was the location of play or of specific play types (25%). This was most often captured in studies using observational methods by mapping children's locations, recording where events took place in field notes, scanning of areas of the school

playground or monitoring activity on or around specific pieces of play equipment. The location of play events also emerged in narrative accounts captured through interviews and focus groups, and through participatory activities such as map-making and playground tours.

Types of play, discrete activities and games

Of all the studies, 55% assessed types of play, discrete activities, or games. Within these studies, several taxonomies were used to categorise play. Two broad taxonomies were frequently used to categorise types of play: social (Parten, 1932) and cognitive (Piaget, 1962; Smilansky, 1968) play types. These taxonomies were not mutually exclusive of each other and were often used together – a combination that has been formalised in Rubin's (1989; 2001) Play Observation Scale. Detailed descriptions of these two taxonomies and the proportion of included studies using them are reported below. Because of the many variations and modifications of these taxonomies, it was sometimes difficult to determine which studies had used each taxonomy. As a result, we have included two percentages for each taxonomy, the first percentage only includes studies that explicitly referenced each taxonomy or the Play Observation Scale (1989, p. 2001) that includes both. The second percentage (italicised) additionally includes studies that categorised some, or all, of the play types in the taxonomy without direct reference to the framework.

Social play types (14%; *28%*). Play was categorised in terms of the *level of social interaction* it involved. Using various modifications of Parten's (1932) taxonomy, play was categorised as solitary play – play away from and with little or no attention paid to other children; parallel play – independent play near to, with considerable attention to, and/or involving similar toys or activities to other children; or group play (sometimes called cooperative or interactive play) – play with other children in which there is a common goal or purpose to the activity. Additionally, two types of non-play behaviour were often coded in these schemes: unengaged/unoccupied behaviour where children stare blankly into space or wander the play space with no seeming purpose or engagement; and onlooking behaviour where children observe the behaviour of other children without becoming actively involved.

Cognitive play types (9%; *28%*). Based on work from Piaget (1962) and Smilansky (1968), play was categorised according to "*cognitive*" categories. These include *functional play* – play that is done for the enjoyment of the physical sensation it creates, typically simple motor movements with or without objects, such as climbing on playground equipment, making faces, or banging objects together; *constructive play* – play in which objects are manipulated for the purpose of constructing or creating something, such as building a tower from blocks; *pretend/socio-dramatic play* – play that involves an element of pretence or role playing, such as pretending to speak on a telephone, or moving a doll as though it is walking; and *games with rules* – play in which the child accepts and adjusts to pre-arranged rules, such as tag or snap.

Aside from these two taxonomies, play has also been categorised in the following ways. Several studies (15%) categorised play with an emphasis on physical activity. For example, the SOCARP tool (Ridgers et al., 2010), described in more detail in Table 1, categorises children's play into sports, active games, sedentary games and locomotion. Other bespoke taxonomies were used in 12% of studies and typically included more granular categories such as creative play, fantasy play, risky play and nature play. Finally, a small

number of studies (4%) measured only one specific type of play, such as rough and tumble play, or solitary play. Rough and tumble play was also commonly categorised alongside other types of play described above (16%).

Some studies (25%) took an even more granular approach, focusing on discrete activities rather than or as well as types of play. Examples of discrete activities included playing on equipment; skipping games; and running and chasing games; as well as specific games, such as "four square", or play with particular toys.

Established coding schemes and questionnaires

Across all the studies, 18 named tools were used for collecting the data. Of these, 14 were observational coding schemes used for structured observations and a further four were questionnaires. Table 1 summarises each coding scheme, reporting the approach (person- or place-based), a brief description of the scheme, and the aspects of play captured as well as the number of studies included in this review that reported using the scheme. Table 2 summarises each questionnaire, reporting the respondent, a brief description, and the aspects of play captured, and the number of studies included in this review that reported using the questionnaire.

Twelve of the tools were only used in a single study included in this review, suggesting that it is common for researchers to develop bespoke coding schemes for each study. The SOPLAY (McKenzie et al., 2000) and SOCARP (Ridgers et al., 2010) tools were used most frequently. Even with these more commonly used tools, many studies reported modifying the play categories for both tools to capture more play categories than were included in the original scheme, which is focused predominantly on sports. For example, by adding play categories such as imaginative play; play with loose parts equipment; sandpit play; and construction. The additional categories were idiosyncratic to each study.

In this section we focused on named tools because these have, at least to some extent, been designed to be used in future research. These measures are quantitative because qualitative approaches, by their nature, are not designed to be reproducible. Nevertheless, previous qualitative research can and should inspire future research and there are some excellent examples that can serve this purpose. Two such examples are the use of a mosaic of creative multimodal ethnographic methods including child-to-child interviews, drawings and maps, GoPro recordings to represent the "messy" and "kaleidoscopic" nature of children's play (Potter & Cowan, 2020, p. 251); and the use of "go-along" interviews with more than 100 children across 17 school playgrounds to supplement unstructured observations capturing gendered activity patterns during breaktime play (Pawlowski et al., 2015); see supplementary material for full list of papers included in the review.

What were the measures used for?

When assessing the purpose of the measures, we considered how the measures of play were used. This was not always the same as the purpose or aim of the whole study, although there was often overlap. These uses are not mutually exclusive – the same study may have used the measures of play for multiple purposes. More than half of the studies (59%) used the measures of play to describe the play itself either in a certain group of children, for example, a clinical or neurodivergent population, or context, for

example, outdoor play, or to describe a certain type of play, such as rough and tumble play, as it occurs during school recess.

Another large proportion (50%) used the measures to determine group differences, for example between age groups, by gender, or between special populations of interest to the study such as between aggressive children and their non-aggressive peers, or between a clinical or neurodivergent sample and typically developing controls. Finally, a substantial proportion (21%) of the studies assessed the effects of playground interventions, including organised activities, loose parts provision and "peer play" training for children with autism spectrum disorder and their peers.

Quality of reporting

To assess the quality of the reporting of current measures, we first determined whether validity and reliability of measures were explicitly reported for each study, and how these were established. Finally, we comment on the transparency of reporting more generally, and the replicability of the research method. It is important to keep in mind that for some qualitative and ethnographic techniques, validity and reliability are not appropriate measures of quality (Rolfe, 2006), however, we found that reporting of both validity and reliability was not exclusive to the quantitative research captured by the review. Clear reporting of the research process benefits both those interpreting study findings and those conducting further research in the field. In particular, transparent reporting of study design, materials and procedures as well as analyses are recommended across a range of reporting guidelines for both quantitative (Bennett et al., 2011; Klein et al., 2018; Munafò et al., 2017) and qualitative (O'Brien et al., 2014; Rolfe, 2006; Tracy, 2010) research. We note that the aim of this section is not to comment on the quality of the methods themselves, nor the research using these methods; this would depend greatly on the methodological approach and research question, and would be more appropriate for a narrower review aimed at a specific research question or approach.

Validity

Validity of measures was explicitly reported in 25% of included studies, across a range of methodological approaches, including structured observations (54%), questionnaires (12%), unstructured observations (4%) and mixed-methods approaches (27%). Of these studies, validity was established in several, non-mutually exclusive, ways as follows: reference to previous research either using or validating the method (54%); the theoretical or conceptual grounding of the measures (15%); data driven approaches including triangulation of data from multiple methods or multiple informants (multivocality), Rasch analysis, and efficacy for discriminating groups (12%); consultation with experts in child behaviour (4%) or the children themselves (4%); and the use of methodological factors that overcome specific obstacles to validity (4%), such as having a familiarisation period so that children's responses are not affected by the presence of the observer(s). In 11% of cases, the measures were reported to be valid although it was unclear how this validity was established.

Reliability

The reliability of (at least one of) the measures employed was reported in 63% of included studies, 88% of which were supported by statistical analysis (e.g. percent agreement or

Cohen's Kappa, see supplementary data for details). We recognise that reliability is more relevant for quantitative than for qualitative approaches (although see O'Connor & Joffe, 2020 for discussion of the debate around this topic), so we explored reporting of reliability within different methodological approaches. Of all studies using structured observation, 87% reported inter-rater reliability. Of these, inter-rater reliability was approached in one or more of the following ways: through live double coding of observations (54%), where two or more coders conducted synchronised live coding for all or a proportion of the observations; video double coding (30%), where two or more coders coded all or a proportion of the video recordings taken during breaktimes; and through training (45%), where reliability between coders was established prior to data collection in a training or pilot phase. In 2% of cases where reliability was reported, it was not clear how this was established.

For studies using methods other than structured observations, 24% reported at least one form of reliability. Within studies using questionnaire measures, reliability was established with reference to internal consistency of questionnaire items (21%); test-retest correspondence between questionnaire responses at different time points (14%); by reference to prior research establishing the reliability of the measure (7%). Within studies using unstructured observations and other qualitative and participatory methods, inter-coder reliability was established using double coding (19%), where two or more researchers coded or analysed all or a proportion of the observations or materials generated through the study (i.e. videos, transcripts, maps, field notes, etc.), and their correspondence was assessed, with one study (3%) additionally establishing reliability through training of researchers.

Transparency

The methods employed in the included studies were frequently not reported in a way that was sufficiently transparent to allow replication of the study – we noted particular issues with the reporting of study materials and the sampling procedures. While these issues may seem trivial or picky, they have important implications for those wishing to fully understand how the data reported were derived, and for understanding any potential biases or misconceptions that may be inherent in the measures. These are key issues for both quantitative and qualitative research, as reflected by their inclusion in many reporting standards and guidelines (Bennett et al., 2011; Klein et al., 2018; O'Brien et al., 2014; Tracy, 2010). Narrative descriptions of the issues noted are given below, along with proportions where appropriate.

Where structured observation was used, it was often the case that the coding scheme was not included or only partially reported (46%), for example, the categories were given but with no descriptions of how behaviours were categorised. Some studies described the observation procedure precisely while others provided only very sparse descriptions of the observation procedure, for example, not describing how long each child was observed for, or how often observations were recorded. Importantly, even when a complete coding scheme was included, we reflected that it was often not clear how it was operationalised, for example, whether categories were mutually exclusive, or how observation intervals that included multiple play categories, occurring either concurrently (i.e. pretend play in nature) or consecutively were coded.

The lack of transparency around study materials was not limited to structured observation. In many cases, studies using questionnaire methods did not include the questions and response options that participants answered, nor did they provide links to the measures reported elsewhere. Similarly, studies using qualitative methods often did not report whether a topic or interview guide was used to direct the interviews, focus groups and walking tours, and no study made a topic or interview guide that was used available to the reader.

There was also a lack of transparency around reporting of the sampling of children for inclusion in the studies. While there were some good examples of descriptions of purposive, random and other sampling strategies, many studies did not describe the process for sampling or selecting children for observations, interviews, or focus groups. This leaves room for concern about bias when selecting children to take part in the studies that can have implications for the interpretation of the results.

Discussion

The aim of this review was to describe, synthesise and compare methods used to assess play during school breaktimes, bringing together methodologies from different fields for the first time. Taking a systematic approach to our search strategy we identified 105 studies that had evaluated, measured or captured children's play during school breaktimes. Our review highlights some significant strengths in this area of research as well as some challenges and areas for improvement.

A field characterised by its diversity and interdisciplinarity

The range of methods and studies included in the review highlights the truly interdisciplinary nature of play research. Although the majority of studies were from the broad fields of psychology and education, papers came from a wide range of disciplines. Further, a wide range of approaches have been taken to the measurement of play in schools from creative qualitative and ethnographic techniques that provide rich, nuanced insight into individual children's experiences, to structured observations and questionnaire measures that provide quantitative data about children's play in schools and support inferential statistics, reproducibility and larger scale research. It is a strength of this field of research that such diverse methods have been used that strongly complement one another; both approaches and their combination are necessary if we are to have a complete understanding of children's experiences and play environments in schools.

The wide range of methods used also represents a significant challenge for the field, but, once again, reflects the complexity of play itself. Of the 105 studies included, the maximum number of papers using the same measure was six, with the vast majority of measures used in only a single study. Whilst this is to be expected in qualitative research, for quantitative research, using a wide array of methods limits study comparisons and impedes a consolidated understanding of children's play. We recognise that many studies will address different questions, or focus on different aspects of play, and so by necessity will use different measures to capture the aspects of play that are of interest. However, in quantitative research, using standardised measures consistently across

studies would allow the relative impact of interventions to be captured, support comparisons across school systems and ages and would facilitate meta-analyses to be conducted.

Equally, as discussed later, it is unlikely that any single methodological approach could capture the rich and subjective nature of play. Thus, although standardised quantitative tools may have utility for reliably capturing children's activities during play and for comparison across time and contexts, these measures will likely miss important aspects of play that can only be captured using richer, qualitative data collection techniques. Ideally, research takes a mixed-methods approach in order to benefit from the relative strengths of both quantitative and qualitative methods – of the studies reviewed, fewer than one-third of studies took a mixed-methods approach to studying play.

Capturing play

Different approaches to capturing play each have distinct strengths and weaknesses, and the methodological approaches chosen for future studies will necessarily depend on the focus of the research question. Due to the broad and inclusive scope of the current review, assessing the quality of the individual measures was not feasible. Instead, we consider the strengths and weaknesses that some of the commonly used approaches bring with the aim of facilitating selection of appropriate methods for future research.

Observational coding schemes tended to focus on capturing play activities or types of play. The taxonomies used varied in their focus (social or cognitive) and concreteness (abstract categories or discrete, concrete activities or behaviours). The focus on discrete activities (e.g. chasing games, skipping games, play on playground equipment) has the advantage that it can be defined relatively objectively and does not require knowledge of the child's goals or internal states, making it easier to observe reliably as part of a structured observation. It is perhaps not surprising therefore that the most frequently used observation measures focus on these discrete activities. However, play is complex, rich and multi-layered, with children's engagement in surface-level activities often being embedded in a richer socio-dramatic context. As such, the information captured through observation of discrete activities is relatively superficial and misses much of the richness of play experience.

In contrast, observations based on taxonomies of play types often acknowledge the diversity and depth of play, but they are impeded by the fact that it can be difficult to distinguish between certain play categories during observation. For example, considering the "cognitive" play types, a child waving their arms around could be doing so because they enjoy the physical sensation (functional) or because they are pretending to be a windmill (socio-dramatic). As a result, it may be difficult for these types of play to be differentiated reliably via observation without including the child's perspective (see also Takhvar & Smith, 1990 for a review and critique of the cognitive play taxonomy). Mixed-methods approaches that include both observation and participatory methods can enrich and validate these more abstract taxonomies by bringing in the child's perspective.

Similarly, measurements of play quality, as well as those capturing other abstract constructs such as risk, often require inferences about the child's internal states to be made. Many of the measures of play quality captured children's own perspectives through self-report questionnaires or interviews and other participatory methods. Nonetheless, these

measures were also taken from the perspective of people other than the child themselves such as through observation, teacher interviews, or questionnaires completed by teachers and school staff. These kinds of measures may be open to bias, especially where only one perspective is taken (e.g. see Phillips & Lonigan, 2010). This is particularly important since measures of play quality are often used to assess the impact of play interventions, such as new playground designs and provision of equipment or staff training. One solution to this is to include multiple perspectives in research, perhaps combining observation with children's and teachers' perspectives. Relatively few of the studies included in the review collected data from more than one perspective.

We believe that the field would benefit from the development of a more critical stance on the methodologies employed, however, this is likely to involve applying different critical frameworks across methodological approaches and research questions. This kind of critical appraisal of the methods employed to measure play is an important next step for the field and we are hopeful that the broad overview of methods and aspects of play captured here gives some structure for moving towards this next stage in the improvement of the rigour of the field.

Quality of reporting

It has been stated that psychology has largely "ignored" children's play (Pellegrini, 2009, p. 137), except for pretend play. One reason for this might be the perception that play research lacks scientific rigour. Indeed, we note that in many cases, markers of high quality and rigorous research were sparse or missing entirely. We assessed the reporting of psychometric properties in quantitative studies included in the review focusing on validity and reliability. We found that only a quarter of included studies explicitly stated how the validity of the measures of play was established. Reliability was reported more frequently, especially for those studies using structured observation, but less so for studies using questionnaire methods. Nonetheless, consistent and transparent evaluation and reporting of the psychometric properties of the methods employed would increase the perceived rigour of play research.

We also assessed the transparency of reporting, considering guidelines for both quantitative and qualitative research (Bennett et al., 2011; Klein et al., 2018; O'Brien et al., 2014; Tracy, 2010). Robust and interpretable research requires methods to be described in sufficient detail to permit replication or to be open to scrutiny from the research community and, for quantitative research, measurement tools to be openly available. The majority of studies included in the review did not provide this level of detail and it was rare for important elements of the measures, such as detailed procedures, coding schemes, interview guides and questionnaire items to be easily accessible for other researchers and stakeholders. Furthermore, many studies did not report the sampling strategy employed, making it difficult for a reader to assess the risk of bias, the generalisability and the appropriateness of the sample to the research question.

This limitation reflects both quantitative and qualitative research. Whilst qualitative and ethnographic approaches are not by their nature designed to lead to reproducible findings, the methods should be transparent and open to scrutiny (Tracy, 2010); studies should be described in enough detail that it is clear what happened and why and how the researchers approached the analysis of their qualitative data (see O'Brien et al.,

2014 for guidance on reporting qualitative research). In quantitative research, full replication of study methods should be possible, but it was rare that sufficient detail was provided (along with access to measures) to permit this. We believe that this lack of transparency in reporting is largely responsible for the proliferation of measures for capturing play. We note that where transparency was high and protocols and materials were made available for other researchers (e.g. SOPLAY; McKenzie et al., 2000; and SOCARP; Ridgers et al., 2010), these measures were more likely to be used across multiple studies.

It is a goal of many researchers in this field for play to be taken seriously and protected by stakeholders such as teachers and policy-makers (London, 2019). We believe that improving the quality and robustness of study methods and reporting across methodological approaches will help researchers make the case for the importance of play. For all research, both qualitative and quantitative, we must provide detailed information about methods and approaches to analysis. This should include coding schemes, topic guides and questionnaires as well as detailed information about observation timing, participant selection and randomisation, identification of children, and analytical approach. Whilst there is often not enough space in journal articles to include full study details, and historic constraints relating to print publication may explain some of the sparse reporting in the reviewed articles, supplementary materials can typically be provided now, either via the journal or an online platform (see Munafò et al., 2017 for further discussion on reproducible research). In addition, quantitative researchers should work to establish a range of well-evaluated instruments that can be used across studies and have evidenced validity and reliability, and they should report validity and reliability of the measures used.

Gaps in the literature

The field of play research has a long history but is rapidly developing, and new methodological approaches are likely already being employed in ongoing research. Ongoing work is necessarily not included in the review, but the authors know of several new, creative approaches to measuring play that continue to be developed and trialled. For example, during the process of conducting the review we came across a new observational tool, the Tool for Observing Play Outdoors (TOPO; Loebach & Cox, 2020), which takes a behaviour mapping approach to observing play, allowing the user to map out the locations of different play types over time. To date, this has not been used for research on school playgrounds.

Whilst a range of taxonomies were used for characterising play across studies, it was surprising that Hughes's & Melville (1996) influential taxonomy of "play types" had not been directly referred to in the methods of any of the studies in the review given its prominence in playwork and early childhood education practice. Hughes's work is frequently cited in playwork literature, underpins playwork training (King & Newstead, 2017), and is commonly referred to in resources for early childhood educators. This taxonomy is also the basis for the coding scheme developed for the new TOPO tool (Loebach & Cox, 2020). We were also surprised to see that risky or adventurous play was only evaluated in two studies of school breaktimes given that there is increasing interest in this area of research (Dodd & Lester, 2021; Sandseter & Kennair, 2011; Tremblay et al., 2015).

In addition to these gaps in the existing literature, there is a dearth of research conducted in the global south or with minority groups, meaning that our existing understanding of children's play is currently dominated by a western, predominantly white, perspective. Play development, experiences, norms and traditions differ across cultural contexts (Edwards, 2000; Shimpi & Nicholson, 2014) and we should not assume that any of the measures developed in western contexts would capture the full range of play expressed in other cultural contexts. Very few of the included papers studied play in adolescents or in high school contexts – however, this may be due to the inclusion criteria, since breaktime activity in this age-group may not be referred to as play. It is likely that specific measures would need to be developed to accurately capture the play of adolescents and this is also a notable gap for future research. Similarly, studies examining the play of neurodiverse and disabled children were limited, with no measures developed to serve this purpose directly. This being said, we note that the Playground Observation of Peer Play (POPE; Kasari et al., 2011) tool, was used exclusively to examine the social play of autistic children in the four studies that used this tool and were captured in this review.

Strengths and limitations

The review has several strengths, in particular the inclusive approach which has allowed us to bring together methods from across this diverse field of research rather than limiting it to a single approach. This necessarily led to some challenges and limitations. For example, the reliability for deciding whether papers should be included or not was relatively low. As outlined in the method, we believe this was a result of the area of research being somewhat noisy in terms of how methods and study aims were defined and described.

Another limitation is that we were not able to describe in detail the wide range of qualitative methods used across studies. Given the idiosyncrasies of qualitative research, any collapsing together would have removed important nuances of specific approaches, and describing each in detail is beyond the scope of this review. We therefore hope that interested readers will refer to the methods sections of specific papers for full details of the various qualitative methods used.

Across both quantitative and qualitative methods, we considered the quality of reporting of the methods. However, we did not assess the quality of the research itself. The diversity of approaches and research questions prohibited the development of a single critical framework for assessing the quality of play research during school breaktimes. Finally, due to the number of included papers and the complexities of combining across this diverse literature we decided not to include grey literature beyond PhD theses but we acknowledge that other resources for observing play are likely to exist beyond the published research.

Conclusions

In summary, we have reviewed the methods used to measure, describe and understand play during school breaktimes. There was considerable variation in how play was measured, reflecting both variation in the aims of individual research teams and fields of research, and a dearth of well-established or standardised tools that are openly

accessible for use in the new research. By synthesising this diverse literature in a single review, we anticipate that the review will facilitate play research in educational contexts and support, where appropriate, more consistent use of methods and measures, allowing studies to be more easily compared, and evidence from play research to be translated into policy.

We hope that by highlighting the areas where reporting could be improved – in particular, openness and transparency around study materials and sampling strategies, we will inspire researchers to report their methods more transparently. Additionally, we encourage those researchers conducting quantitative research to interrogate the validity and reliability of the measures they decide to use or to develop. Finally, we hope that all researchers studying play in schools will consider the use of mixed methods and multivocality to get a richer picture of children's play than can be afforded by any single measure.

Notes

1. Global North and Global South are terms used to group countries according to their level of economic advantage. Here we use the Finance Center for South-South Cooperation from the United Nations' list (http://www.fc-ssc.org/en/partnership_program/south_south_countries) to determine which countries are classified as being part of the "Global South".
2. This construct was termed "appropriateness" in several studies but we have chosen not to adopt this term because some forms of rough or negative play are developmentally appropriate.

Acknowledgements

We would like to thank Sajida Parveen for their assistance in the abstract screening phase of this review.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is funded by a URKI Future Leaders Fellowship awarded to Professor Helen Dodd (MR/S017909/1).

Data availability statement

Extracted data, the complete list of included reports and metadata files are available from the following Open Science Framework repository https://osf.io/8rga6/?view_only=859bc5b0037746c4a8e6aa8885a072d0.

References

Andersen, M. M., Kiverstein, J., Miller, M., & Roepstorff, A. (2023). Play in predictive minds: A cognitive theory of play. *Psychological Review*, *130*(2), 462–479. https://doi.org/10.1037/rev0000369

Baines, E., & Blatchford, P. (2019). *School break and lunch times and young people's social lives: A follow-up national study* (p. 115). https://www.nuffieldfoundation.org/wp-content/uploads/2019/11/Baines204240220BreaktimeSurvey20-20Main20public20report20May19-Final1.pdf.

Bennett, C., Khangura, S., Brehaut, J. C., Graham, I. D., Moher, D., Potter, B. K., & Grimshaw, J. M. (2011). Reporting guidelines for survey research: An analysis of published guidance and reporting practices. *PLOS Medicine*, *8*(8), e1001069. https://doi.org/10.1371/journal.pmed.1001069

Bernardo, M. D. (1995). *'Laro tayo.': Parent-child and peer play activities of Filipino children and related variables*. https://www.proquest.com/docview/304120340?pq-origsite=gscholar&fromopenview=true.

Brussoni, M., Ishikawa, T., Brunelle, S., & Herrington, S. (2017). Landscapes for play: Effects of an intervention to promote nature-based risky play in early childhood centres. *Journal of Environmental Psychology*, *54*, 139–150. https://doi.org/10.1016/j.jenvp.2017.11.001

Bundy, A., Engelen, L., Wyver, S., Tranter, P., Ragen, J., Bauman, A., Baur, L., Schiller, W., Simpson, J. M., Niehues, A. N., Perry, G., Jessup, G., & Naughton, G. (2017). Sydney playground project: A cluster-randomized trial to increase physical activity. *Play, and Social Skills. Journal of School Health*, *87*10), 751–759. https://doi.org/10.1111/josh.12550

Bundy, A., Nelson, L., Metzger, M., & Bingaman, K. (2001). Validity and reliability of a test of playfulness. *The Occupational Therapy Journal of Research*, *21*(4), 276–292. https://doi.org/10.1177/153944920102100405

Charlton, T., Gunter, B., & Coles, D. (1998). Broadcast television as a cause of aggression? Recent findings from a naturalistic study. *Emotional & Behavioural Difficulties*, *3*(2), 5–13.

Children and Young Persons, Wales. (2012). *The play sufficiency assessment (Wales) regulations 2012*. Queen's Printer of Acts of Parliament. https://www.legislation.gov.uk/wsi/2012/2555/made.

Clevenger, K. A., Lowry, M., Perna, F. M., & Berrigan, D. (2022). Cross-sectional association of state recess laws with district-level policy and school recess provision in the United States. *Journal of School Health*, *92*(10), 996–1004. https://doi.org/10.1111/josh.13189

Covidence Systematic Review Software. (n.d.). [Computer software]. Veritas Health Innovation. http://www.covidence.org/.

Dodd, H. F., & Lester, K. J. (2021). Adventurous play as a mechanism for reducing risk for childhood anxiety: A conceptual model. *Clinical Child and Family Psychology Review*, *24*(1), 164–181. https://doi.org/10.1007/s10567-020-00338-w

Dodd, H. F., Nesbit, R. J., & FitzGibbon, L. (2022). Child's play: Examining the association between time spent playing and child mental health. *Child Psychiatry & Human Development*, 1–9.

Eberle, S. G. (2014). The elements of play: Toward a philosophy and a definition of play. *American Journal of Play*, *6*(2), 214–233.

Edwards, C. P. (2000). Children's play in cross-cultural perspective: A new look at the six cultures study. *Cross-Cultural Research*, *34*(4), 318–338. https://doi.org/10.1177/106939710003400402

Engelen, L., Wyver, S., Perry, G., Bundy, A., Chan, T. K. Y., Ragen, J., Bauman, A., & Naughton, G. (2018). Spying on children during a school playground intervention using a novel method for direct observation of activities during outdoor play. *Journal of Adventure Education and Outdoor Learning*, *18*(1), 86–95.

Evans, J. (1996). Children's attitudes to recess and the changes taking place in Australian primary schools. *Research in Education*, *56*(1), 49–61. https://doi.org/10.1177/003452379605600104

Fantuzzo, J., Sutton-Smith, B., Coolahan, K. C., Manz, P. H., Canning, S., & Debnam, D. (1995). Assessment of preschool play interaction behaviors in young low-income children: Penn Interactive Peer Play Scale. *Early Childhood Research Quarterly*, *10*(1), 105–120. https://doi.org/10.1016/0885-2006(95)90028-4

Fjørtoft, I., Kristoffersen, B., & Sageie, J. (2009). Children in schoolyards: Tracking movement patterns and physical activity in schoolyards using global positioning system and heart rate monitoring. *Landscape and Urban Planning*, *93*(3), 210–217. https://doi.org/10.1016/j.landurbplan.2009.07.008

Greer, D. L., & Stewart, M. J. (1989). Children's attitudes toward play: An investigation of their context specificity and relationship to organized sport experiences. *Journal of Sport and Exercise Psychology*, *11*(3), 336–342. https://doi.org/10.1123/jsep.11.3.336

Henley, J., McBride, J., Milligan, J., & Nichols, J. (2007). Robbing elementary students of their childhood: The perils of no child left behind. *Education*, *128*(1), 56–63.

Heravi, B. M., Gibson, J. L., Hailes, S., & Skuse, D. (2018). Playground social interaction analysis using bespoke wearable sensors for tracking and motion capture. *Proceedings of the 5th International Conference on Movement and Computing*, 1–8. https://doi.org/10.1145/3212721.3212818

Herrington, S., & Brussoni, M. (2015). Beyond physical activity: The importance of play and nature-based play spaces for children's health and development. *Current Obesity Reports*, *4*(4), 477–483. https://doi.org/10.1007/s13679-015-0179-2

Hughes, B., & Melville, S. (1996). *A playworker's taxonomy of play types*. PLAYLINK.

Hyndman, B. (2015). *Protocol for the lunchtime enjoyment of activity and play (LEAP) Questionnaire*^{TM}. https://doi.org/10.13140/RG.2.1.4032.9126

Ingram, D. H., Mayes, S. D., Troxell, L. B., & Calhoun, S. L. (2007). Assessing children with autism, mental retardation, and typical development using the playground observation checklist. *Autism*, *11*(4), 311–319. https://doi.org/10.1177/1362361307078129

Jamison, K. R. (2010). *Effects of a social communication intervention for promoting social competence through play in young children with disabilities* (854983784) [University of Virginia]. https://search.proquest.com/dissertations-theses/effects-social-communication-intervention/docview/854983784/se-2?accountid=13460.

Jarrett, O. S., Maxwell, D. M., Dickerson, C., Hoge, P., Davies, G., & Yetley, A. (1998). Impact of recess on classroom behavior: Group effects and individual differences. *The Journal of Educational Research*, *92*(2), 121–126. https://doi.org/10.1080/00220679809597584

Kasari, C., Locke, J., Gulsrud, A., & Rotheram-Fuller, E. (2011). Social networks and friendships at school: Comparing children with and without ASD. *Journal of Autism and Developmental Disorders*, *41*(5), 533–544. https://doi.org/10.1007/s10803-010-1076-x

Khorana, P. (2017). *The effects of wellness in the schools (WITS) on physical activity and play engagements during recess in New York City public schools* (1928585732) [Teachers College, Columbia University]. https://search.proquest.com/dissertations-theses/effects-wellness-schools-wits-on-physical/docview/1928585732/se-2?accountid=13460.

King, P., & Newstead, S. (Eds.). (2017). *Researching play from a playwork perspective* (1st ed.). Routlege. https://doi.org/10.4324/9781315622149

Klein, O., Hardwicke, T. E., Aust, F., Breuer, J., Danielsson, H., Mohr, A. H., IJzerman, H., Nilsonne, G., Vanpaemel, W., & Frank, M. C. (2018). A practical guide for transparency in psychological science. *Collabra: Psychology*, *4*(1), 20. https://doi.org/10.1525/collabra.158

Leff, S. S., & Lakin, R. (2005). Playground-based observational systems: A review and implications for practitioners and researchers. *School Psychology Review*, *34*(4), 475–489.

Lester, S., & Russell, W. (2010). Children's right to play: An examination of the importance of play in the lives of children worldwide. Working Papers in Early Childhood Development, No. 57. In *Bernard van Leer Foundation (NJ1)*. Bernard van Leer Foundation. https://eric.ed.gov/?id=ED522537.

Loebach, J., & Cox, A. (2020). Tool for observing play outdoors (TOPO): A new typology for capturing children's play behaviors in outdoor environments. *International Journal of Environmental Research and Public Health*, *17*(15), Article 15. https://doi.org/10.3390/ijerph17155611

London, R. A. (2019). The right to play: Eliminating the opportunity gap in elementary school recess. *Phi Delta Kappan*, *101*(3), 48–52. https://doi.org/10.1177/0031721719885921

Massey, W. V., Ku, B., & Stellino, M. B. (2018). Observations of playground play during elementary school recess. *Bmc Research Notes*, *11*(1), 755. https://doi.org/10.1186/s13104-018-3861-0

Massey, W. V., Thalken, J., Szarabajko, A., Neilson, L., & Geldhof, J. (2021). Recess quality and social and behavioral health in elementary school students. *Journal of School Health*, *91*(9), 730–740. https://doi.org/10.1111/josh.13065

McKenzie, T. L., Marshall, S. J., Sallis, J. F., & Conway, T. L. (2000). Leisure-time physical activity in school environments: An observational study using SOPLAY. *Journal Devoted to Practice and Theory*, *30*(1), 70–77.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, P., Simonsohn, N., Wagenmakers, U., Ware, E.-J., & Ioannidis, J. J., & A, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), Article 1. https://doi.org/10.1038/s41562-016-0021

Nijhof, S. L., Vinkers, C. H., van Geelen, S. M., Duijff, S. N., Achterberg, E. J. M., van der Net, J., Veltkamp, R. C., Grootenhuis, M. A., van de Putte, E. M., Hillegers, M. H. J., van der Brug, A. W., Wierenga, C. J., Benders, M. J. N. L., Engels, R. C. M. E., van der Ent, C. K., Vanderschuren, L. J. M. J., & Lesscher, H. M. B. (2018). Healthy play, better coping: The importance of play for the development of children in health and disease. *Neuroscience & Biobehavioral Reviews*, *95*, 421–429. https://doi.org/10.1016/j.neubiorev.2018.09.024

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, *89*(9), 1245. https://doi.org/10.1097/ACM.0000000000000388

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, *19*. https://doi.org/10.1177/1609406919899220

OFSTED School inspection handbook. (2019). 97. https://www.gov.uk/government/publications/school-inspection-handbook-eif

Parrish, A.-M., Chong, K. H., Moriarty, A. L., Batterham, M., & Ridgers, N. D. (2020). Interventions to change school recess activity levels in children and adolescents: A systematic review and meta-analysis. *Sports Medicine*, *50*(12), 2145–2173. https://doi.org/10.1007/s40279-020-01347-z

Parten, M. B. (1932). Social participation among pre-school children. *The Journal of Abnormal and Social Psychology*, *27*(3), 243–269. https://doi.org/10.1037/h0074524

Pawlowski, C. S., Ergler, C., Tjørnhøj-Thomsen, T., Schipperijn, J., & Troelsen, J. (2015). 'Like a soccer camp for boys': A qualitative exploration of gendered activity patterns in children's self-organized play during school recess. *European Physical Education Review*, *21*(3), 275–291.

Pellegrini, A. D. (2009). *The role of play in human development* (pp. ix, 278). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195367324.001.0001

Pellegrini, A. D., & Bjorklund, D. F. (1997). The role of recess in children's cognitive performance. *Educational Psychologist*, *32*(1), 35–40. https://doi.org/10.1207/s15326985ep3201_3

Pellegrini, A. D., & Bohn, C. M. (2005). The role of recess in children's cognitive performance and school adjustment. *Educational Researcher*, *34*(1), 13–19.

Pepler, D. J., Craig, W. M., & Roberts, W. L. (1998). Observations of aggressive and nonaggressive children on the school playground. *Merrill-Palmer Quarterly*, *44*(1), 55–76.

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation*, *13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Phillips, B. M., & Lonigan, C. J. (2010). Child and informant influences on behavioral ratings of preschool children. *Psychology in the Schools*, *47*(4), 374–390. https://doi.org/10.1002/pits.20476

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

Potter, J., & Cowan, K. (2020). Playground as meaning-making space: Multimodal making and remaking of meaning in the (virtual) playground. *Global Studies of Childhood*, *10*(3), 248–263. https://doi.org/10.1177/2043610620941527

Ridgers, N. D., Salmon, J., Parrish, A.-M., Stanley, R. M., & Okely, A. D. (2012). Physical activity during school recess: A systematic review. *American Journal of Preventive Medicine*, *43*(3), 320–328. https://doi.org/10.1016/j.amepre.2012.05.019

Ridgers, N. D., Stratton, G., & McKenzie, T. L. (2010). Reliability and validity of the system for observing children's activity and relationships during play (SOCARP). *Journal of Physical Activity & Health*, *7*(1), 17–25. https://doi.org/10.1123/jpah.7.1.17

Rolfe, G. (2006). Validity, trustworthiness and rigour: Quality and the idea of qualitative research. *Journal of Advanced Nursing*, *53*(3), 304–310. https://doi.org/10.1111/j.1365-2648.2006.03727.x

Rubin, K. H. (1989). *Play Observation Scale (POS)* [Unpublished training manual].

Rubin, K. H. (2001). *Play Observation Scale (POS, revised)* [Unpublished training manual].

Rusby, J. C., & Dishion, T. J. (1990). *The playground code (PGC)—Observing school children at play* [Unpublished training manual].

Sandseter, E. B. H., & Kennair, L. E. O. (2011). Children's risky play from an evolutionary perspective: The anti-phobic effects of thrilling experiences. *Evolutionary Psychology*, *9*(2), 257–284.

Shimpi, P. M., & Nicholson, J. (2014). Using cross-cultural, intergenerational play narratives to explore issues of social justice and equity in discourse on children's play. *Early Child Development and Care*, *184*(5), 719–732. https://doi.org/10.1080/03004430.2013.813847

Singer, D. G., Golinkoff, R. M., & Hirsh-Pasek, K. (2006). *Play = learning: how play motivates and enhances children's cognitive and social-emotional growth*. Oxford University Press.

Singer, D. G., & Singer, J. L. (1981). Television and the developing imagination of the child. *Journal of Broadcasting*, *25*(4), 373–387. https://doi.org/10.1080/08838158109386461

Smilansky, S. (1968). *The effects of sociodramatic play on disadvantaged preschool children*. https://eric.ed.gov/?id=ed033761.

Smith, P. K. (2009). *Children and play: Understanding children's worlds*. John Wiley & Sons, Incorporated. http://ebookcentral.proquest.com/lib/exeter/detail.action?docID=7104580

Stellino, M. B., & Sinclair, C. (2014). Examination of children's recess physical activity patterns using the activities for daily living-playground participation (ADL-PP) instrument. *Journal of Teaching in Physical Education*, *33*(2), 282–296. https://doi.org/10.1123/jtpe.2013-0156

Stellino, M. B., & Sinclair, C. D. (2013). Psychological predictors of children's recess physical activity motivation and behavior. *Research Quarterly for Exercise and Sport*, *84*(2), 167–176. https://doi.org/10.1080/02701367.2013.786159

Takhvar, M., & Smith, P. K. (1990). A review and critique of Smilansky's classification scheme and the "nested hierarchy" of play categories. *Journal of Research in Childhood Education*, *4*(2), 112–122. https://doi.org/10.1080/02568549009594792

Tracy, S. J. (2010). Qualitative quality: Eight "Big-TENT" criteria for excellent qualitative research. *Qualitative Inquiry*, *16*(10), 837–851. https://doi.org/10.1177/1077800410383121

Tremblay, M. S., Gray, C., Babcock, S., Barnes, J., Bradstreet, C. C., Carr, D., Chabot, G., Choquette, L., Chorney, D., Collyer, C., Herrington, S., Janson, K., Janssen, I., Larouche, R., Pickett, W., Power, M., Sandseter, E. B. H., Simon, B., & Brussoni, M. (2015). Position statement on active outdoor play. *International Journal of Environmental Research and Public Health*, *12*(6), 6475–6505. https://doi.org/10.3390/ijerph120606475

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

UN General Assembly. (1989). *Convention on the rights of the child*. UN General Assembly. https://treaties.un.org/doc/Treaties/1990/09/19900902%2003-14%20AM/Ch_IV_11p.pdf.

Webb, H. (1969). Professionalization of attitudes toward play among adolescents. *Aspects of Contemporary Sport Sociology*, 161–187.

Whitebread, D. (2017). Free play and children's mental health. *The Lancet Child & Adolescent Health*, *1*(3), 167–169. https://doi.org/10.1016/S2352-4642(17)30092-5

Yogman, M., Garner, A., Hutchinson, J., Hirsh-Pasek, K., Golinkoff, R. M., Committee on Psychosocial Aspects of Child and Family Health, Council on Communications and Media ON COMMUNICATIONS AND MEDIA, Baum, R., Gambon, T., Lavin, A., Mattson, G., Wissow, L., Hill, D. L., Ameenuddin, N., Chassiakos, Y. (Linda) R., Cross, C., Boyd, R., Mendelson, R., Moreno, M. A. … Smith, J. (2018). The power of play: A pediatric role in enhancing development in young children. *Pediatrics*, *142*(3), e20182058. https://doi.org/10.1542/peds.2018-2058