



Development and Validation of a Measure of Quality in Playgroups: Playgroups Environment Rating Scale

Vanessa Russo^{1*}, M. Clara Barata², Joana Alexandre¹, Catarina Leitão² and Bruno de Sousa³

¹ Iscte - Instituto Universitário de Lisboa, Center for Psychological Research and Social Intervention (CIS-Iscte), Lisboa, Portugal, ² Institute of Interdisciplinary Research, Center of Interdisciplinary Studies (CEIS20), University of Coimbra, Coimbra, Portugal, ³ Faculty of Psychology and Educational Sciences, Center for Research in Neuropsychology and Cognitive and Behavioral Intervention (CINEICC), University of Coimbra, Coimbra, Portugal

Playgroups are community-based services that bring together young children and their caregivers for the purpose of play and social activities. Preliminary evidence shows that playgroup impacts may be dependent on the quality of the playgroup. However, to date, there is no reliable and valid measure of playgroup quality. In this paper we describe the development and validation of the Playgroup Environment Rating Scale (PERS), a standardized observation measure designed to assess the quality of playgroups. PERS builds on traditional measures used to evaluate the quality of formal settings of early childhood education and care, while proposing to assess dimensions of quality specific to the nature of playgroups, namely complex interactions between several types of participants. After developing and testing the observation measure on 24 playgroup videos, we analyzed the psychometric properties. Results showed that the PERS had good interrater reliability, was internally consistent and shows a good preliminary factor structure. Tests for convergent and criterion-related validity also presented promising results. The process of design guaranteed that the PERS can be applied to different contexts of playgroups and may also be useful for informing service planning and practice. Further national and international validation will help replicate the validity of the scale.

Keywords: playgroups, validity, reliability, quality assessment, process quality, early childhood education and care (ECEC), supported playgroups, play

OPEN ACCESS

Edited by:

Mats Granlund,
Jönköping University, Sweden

Reviewed by:

Madeleine Sjöman,
Malmö University, Sweden
Lin-Ju Kang,
Chang Gung University, Taiwan

*Correspondence:

Vanessa Russo
vsfcrusso@gmail.com

Specialty section:

This article was submitted to
Assessment, Testing and Applied
Measurement,
a section of the journal
Frontiers in Education

Received: 15 February 2022

Accepted: 02 May 2022

Published: 26 May 2022

Citation:

Russo V, Barata MC, Alexandre J,
Leitão C and de Sousa B (2022)
Development and Validation of a
Measure of Quality in Playgroups:
Playgroups Environment Rating Scale.
Front. Educ. 7:876367.
doi: 10.3389/feduc.2022.876367

INTRODUCTION

Playgroups are community-based groups that bring together young children (prior to school age) and their parents or caregivers for the purpose of play and social activities (Dadich and Spooner, 2008). Playgroups generally meet in a semiweekly schedule for sessions of 2 h, during the school year, in diverse settings such as community spaces, public services or at the caregivers' home. Playgroups sessions are generally centered in the caregiver-child interaction, and are low-cost or free of cost (Williams et al., 2015).

In Portugal, playgroups were implemented at a national level in the pilot project Playgroups for Inclusion. This project was targeted to families with children up to 4 years old, not participating in any of the currently available Early Childhood Education and Care (ECEC) services. Recruitment was focused on families belonging to minority groups and families whose caregivers were unemployed and underemployed (Barata et al., 2017).

Playgroups for Inclusion were designed as supported playgroups with paid and continuously trained facilitators that provided semiweekly sessions during 10 months, and were supervised by a hired early childhood educator. The role of the supervisors was to support the facilitators in their work with the families and children, and to promote a time for reflection with each group of facilitators (Freitas-Luís et al., 2005). Activities included music and singing, imaginative, outdoor and free play, art and craft activities, and were designed with the purpose of creating opportunities to learn, socialize, develop and increase skills, while aiming to meet the needs and interests of the participants, in a climate of interaction, sharing and cooperation with peers.

The Playgroups for Inclusion project included an experimental study and a study of program implementation. A Theory of Change underpinning playgroups and the core intended outcomes of playgroups was constructed, in order to select a set of primary and secondary outcomes for measurement. A theory of change is a conceptual tool that allows teams to examine the congruence between the object of study, and the proposed research design(s), evaluation measures, analysis plan, etc. (Weiss, 1995; Connell and Kubisch, 1998; Anderson, 2005). Regarding the expected impact results of the project, the planned intervention aimed to affect interrelated outcomes in three domains: caregiver and caregiving, children's development, and community. In the domain of children's development, the project aimed to reduce developmental gaps in cognitive and social domains, the likelihood of future school failure and social exclusion during compulsory schooling. The study of program implementation aimed to describe playgroup development over 10 months, examining the nature and extent of implementation in key areas such as quality of playgroups (Barata et al., 2017).

Assessing Quality in Playgroups

Playgroups are implemented in several countries (e.g., Italy, Germany, United Kingdom, Australia) and are very important in the family support gap between maternity services and children's school entry (Dadich and Spooner, 2008). For example, in England, over 6% of preschool children up to 4 years participate in playgroups (Department for Education, 2018).

Research evidence about playgroups provides indication that such services improve a range of outcomes for children, such as language, cognition, behavioral skills, and on children's developing ability to reason through manual and visuospatial problems, including speed of working and precision (Deutscher et al., 2006; Page et al., 2021).

A growing body of research suggests that the magnitude of the benefits for children will depend on the level of quality of ECEC services, and that low-quality ECEC can be associated with no benefits or even with detrimental effects on children's development and learning (Howes et al., 2008; Britto et al., 2011). This evidence is especially strong in the case of children from disadvantaged families (Garces et al., 2002; Gormley et al., 2005) and from contexts of war and displacement (Wuermli et al., 2021). It is therefore essential that the quality of all

ECEC services is monitored with reliable and valid instruments, including instruments to assess playgroup quality.

Quality can be seen as encompassing all the features of children's environments and experiences that are assumed to benefit their wellbeing (Litjens and Makowiecki, 2014). Definitions of ECEC quality often distinguish between structural characteristics and process features (for a review see Slot, 2018).

Structural characteristics are conceptualized as more distal indicators of ECEC quality, such as child-staff ratio, group size and staff training or education (Howes et al., 2008; Thomason and La Paro, 2009; Slot et al., 2015; Barros et al., 2016). Structural quality has been perceived as providing the preconditions for process quality (Cryer et al., 1999; Melhuish and Gardiner, 2019).

Process quality concerns the more proximal processes of children's everyday experience and involves the social, emotional and physical aspects of their interactions with staff and other children while being involved in play, activities or routines (Howes et al., 2008; Ghazvini and Mullis, 2010; Anders, 2015; Slot et al., 2015; Barros et al., 2016). Process quality has been seen as the primary driver of children's development and learning through ECEC (Howes et al., 2008; Mashburn et al., 2008; Weiland et al., 2013; OECD, 2018). Several studies with preschool children have found that sensitive, well-organized, and cognitively stimulating interactions foster children's development in domains such as language, mathematics, self-regulation, and reduction of behavior problems (Howes et al., 2008; Mashburn et al., 2008; Weiland et al., 2013; OECD, 2018).

In addition to interactions, one other core domain of process quality has been identified in a robust study for infants and toddlers: use of space and materials (Barros et al., 2016). This domain describes infant interactions with materials and within activities that are intrinsically linked to caregivers' ongoing decisions and actions. Indicators that assess the quality of the experiences that infants have with space and materials have been linked with learning and development (Vandell, 2004; Helmerhorst et al., 2014; Berti et al., 2019).

The Environment Rating Scales (ERS) measures, for example, the Infant Toddler Environmental Rating Scale-Revised (ITERS-R; Harms et al., 2006) or the Early Childhood Environmental Rating Scale-Revised (ECERS-R; Harms et al., 1998), are the most commonly used observational instruments to evaluate the process quality in formal ECEC (Slot, 2018). These measures include a wide range of dimensions of environmental quality, such as furnishing and materials, the provision of variety of activities, aspects of the interactions and program structure.

Some studies also use the ERS to assess process quality in playgroups. The ERS, while valid and useful for assessing the quality in formal daycare, present severe limitations for assessing quality in playgroups. For example, playgroups usually take place in spaces available for other purposes (e.g., libraries) and so some indicators of quality by the ERS may be inappropriate (e.g., nap time, personal care; see Melhuish, 1994; Lera et al., 1996).

The ERS also do not consider parental involvement and participation in sessions, which are central to playgroup dynamics (Statham and Brophy, 1992). The multiplicity of roles that occur in a playgroup imply a careful look to all ongoing interactions, including between facilitators and

children, facilitators and caregivers, among caregivers, and among children. The same limitations apply to measures used in formal settings that were slightly adapted to playgroups, such as Preschool Program Quality Assessment of the program High/Scope (PQA, High/Scope Educational Research Foundation, 2001, see French, 2005), the Adult Style Observation Scale (ASOS, Laevers, 1994, see Ramsden, 2007) and the Quality Learning Instrument (QLI, Walsh and Gardner, 2005, see Cunningham et al., 2004). These limitations demonstrate the need to develop an observational scale that considers the specificities of playgroup environment and interactions.

Recently, Commerford and Hunter (2017) were pioneers by identifying core components of quality that are specific to playgroups. These include space, activities and play experiences, interactions taking place and the presence of skilled facilitators to engage families.

The space of the playgroup needs to be welcoming and warm, easily accessed, adequately resourced, and adaptive to the needs of different cultural groups (Williams et al., 2015). The group size is also very important. While a few studies have recommended group sizes of 4–12 families or 6–8 for playgroups (Social Entrepreneurs Inc., 2011; McArthur and Butler, 2012), the ideal group size needs to allow the caregiver and child to receive adequate attention, where problems can be identified more readily and a familiar and safe group environment can be fostered (Salinger, 2009).

The activities and play experiences for the families and children need to be fun, support child development, and allow caregivers to participate and further develop their own skills (Commerford and Hunter, 2017). Play provides children with many opportunities to learn (Department of Education and Training, 2009) and is associated with the development of language and literacy, sociability, and mathematical ability in children (Hancock et al., 2012). Also, the benefits of play can be introduced to the caregivers who, through their diverse or disadvantaged backgrounds, have little personal experiences of play (Commerford and Hunter, 2017).

Playgroup quality assessments also must provide an indication of the interactions taking place. Research demonstrates that one of the main reasons caregivers join playgroups is to develop a sense of belonging (Harman et al., 2014), develop friendships and finding emotional and social support (Gibson et al., 2015; Hancock et al., 2015). Also, young children learn through relationships (OECD, 2018), therefore it is essential that warm, welcoming and inclusive interactions that facilitate positive relationships be present in playgroups. Finally, facilitators need to have the training, knowledge, and skills to provide the support needed.

Designing the Playgroups Environment Rating Scale

The Playgroups Environment Rating Scale (PERS) was based on the core domains of process quality for infants and toddlers identified by Barros et al. (2016) and the core components of quality in playgroups identified by Commerford and Hunter (2017). The PERS focused on the three main goals for

the playgroups: to promote children and caregiver's natural learning through play; to promote wellbeing and socialization environments between all participants; and to ensure a space and time for exploration, discovery, sharing, and positive interactions between adult(s) and children, among adults and among children (for a detailed description see Freitas-Luís et al., 2005).

To maximize the alignment between the stipulated goals for the playgroups and the knowledge of existing measures, we decided that the PERS should follow the structure of the ERS for two reasons. First, the ERS seemed better suited to match stipulated goals, namely the focus on play-based learning. Second, the ERS were found to be the most used observational measures in the literature (Vermeer et al., 2016) and in 33% of the studies that assess quality in playgroups, making them easier to adapt and more relevant to the PERS target audience.

The goal of the present study is to examine the psychometric properties of a new measure to assess the quality of playgroups. The five specific aims are: (1) To test the reliability of the PERS, analyzing the internal consistency of the scale and the interrater reliability; (2) to assess the scale sensitivity to changes in quality over the 10 months of implementation; (3) to explore the factor structure of the PERS; (4) to test the convergent validity with the Adult Style Observation Schedule (ASOS; Laevers, 2000) and (5) to test the criterion-related validity of the PERS with structural characteristics of quality (concurrent validity) and with the outcomes for children that participated in playgroups (predictive validity). We hypothesized that significantly higher playgroup quality would be observed in playgroups with fewer dyads (1 dyad of a caregiver and one child), rather than more and facilitators with more years of experience, rather than fewer. Finally, we also hypothesized higher playgroup quality would be predictive of significantly higher outcomes in children's cognitive development. Based on our Theory of Change, initially based on extant literature on playgroups and ECEC services, and then reviewed in accordance to the intentionality discussed by the intervention team in common meetings, in this study we focus only on three domains of children's cognitive development: language, performance and practical reasoning.

MATERIALS AND METHODS

Development of the Scale

Following the Spector (1992) guide for scale construction, the development of the PERS began with the Exploratory phase, in which the research team conducted a literature review of the playgroups literature, focusing on existing ECEC quality measures, other observational measures, and gray literature on playgroup implementation (e.g., Berthelsen et al., 2012; McArthur and Butler, 2012).

In the second phase, conceptualization, we specified the domains of process quality that would compose the indicators of quality of the PERS. We decided to include indicators that assess the quality of the experiences that infants have with space and materials because of its reported importance for learning and development (Vandell, 2004; Helmerhorst et al., 2014). We also include indicators that were specific to the nature of playgroups

(Commerford and Hunter, 2017), namely the playgroup routine, the presence of different adults interacting with the child and with each other, and also indicators related with the contact with diversity because of its importance in the inclusion of culturally and linguistically diverse families (Strange et al., 2014).

In the third phase, item generation, the conceptualized dimensions of the PERS were operationalized into 18 subdimensions. The first indicators that we created for one item guided the development of subsequent indicators for the same item. For example, the item Interaction between Facilitators and Children included one indicator on the description for “inadequate” quality assessing if the facilitators promoted or not the autonomy of the children. On the description for “minimal,” “good” and “excellent” quality, this indicator became more specific, measuring how many times the facilitator promoted the autonomy of children and how.

In the fourth phase, expert review, five early childhood educators and 10 trained facilitators of playgroups made suggestions for improvements. Experts suggested the addition of new items as well as additional indicators based on their practice and experience. The items and suggestion were related to the organization of the space of the playgroup. Playgroups are frequently installed in spaces provided by local public or private organizations, and the quality of the playgroup depends on the potential of the space to be organized to accommodate children, adults and playgroup activities. Experts also requested further clarification of concepts such as “accessing materials,” “facilitator’s initiative” and “free play,” which were implemented in clarification notes where these concepts first appeared. Regarding the structure of the measure, experts made specific comments about the allocation of certain indicators to expected levels of quality. For example, in the item “materials” an indicator that was first considered to be on the level “minimal” quality (“switch materials to provide variety”) was recommended to be changed to the level “good”.

In the fifth phase, piloting, we collected 24 playgroup video sessions that corresponded to 12 playgroups observed approximately 10 months apart. One researcher scored all videos using the preliminary version of the PERS. Interrater reliability between the researcher and a reliability coder was carried out on a randomly selected sample constituted by 33% of the videos (i.e., 8 videos). Results from the preliminary psychometric study of the properties of the PERS revealed that the subscale items were not very consistent, suggesting that the subscale items were not very related. Results also noted that independent coders trained in using the preliminary version of the PERS were not reliable in their ratings, suggesting that further clarification of the items and indicators could enhance interrater reliability. Such results determined a revision of the measure that was grounded in an in-depth expert review.

In the sixth and final phase, final review, 10 facilitators were asked about which quality characteristics they valued the most about playgroups and how. The examples of adequate and inadequate quality characteristics that facilitators gave helped us add more clarification notes and examples to all the items and indicators. The facilitators also mentioned the importance of two aspects of the quality of playgroups that we did not have

included in the preliminary version: the interaction between the facilitators and the diversity of materials. We decided to include these aspects on the PERS for the reasons followed. Regarding the first aspect, the interactions between facilitators have found to determine playgroup quality by modeling positive interactions between parents (Social Entrepreneurs Inc., 2011). Regarding the second aspect, published research on the ECERS-E (ECERS-Extension, Sylva et al., 2003) supports the importance of the subscale “Diversity” for predictive validity of child outcomes (Sylva et al., 2006). Facilitators also highlighted the importance of playgroup supervision and the role of supervisors, an item which had been included on the PERS but does not apply to all playgroup models implemented outside of Portugal. The importance of a knowledgeable supervisor in playgroups has been reported to provide a sounding board and support for facilitators (Social Entrepreneurs Inc., 2011; Commerford and Hunter, 2017). Supervisors provide opportunities to reflect, problem solve, and even role-play difficult situations and relationships that ultimately can improve the quality of playgroups. Therefore, we decided to maintain the items regarding the supervisors’ role with the option to code “not applicable.”

Design, Participants, and Procedure

This study used data from 13 (out of 25) randomly selected Playgroups for Inclusion, as well as data from the 103 families, 14 facilitators and five supervisors in these playgroups. The 13 playgroups were located across five districts of Portugal, and were randomly selected (stratified by district) because of logistic constraints. Selected families were assessed at pretest ($N = 103$), and posttest (73%, $N = 75$).

Data collection of children’s development took place in August 2015 and posttest took place 1 year after. The participation in both phases implied a home visit (2 h maximum) by one or two trained psychologists.

The mean age of the participating children at pretest was 16 months, ranging from 70 days to 46 months and 50% were younger than 16 months. The mean age of the caregivers at posttest was 35 years old, ranging from 15 to 68 years old ($SD = 11.13$), 63% have completed secondary education, and they were mostly mothers (85%). Approximately 86% of the households’ income was above the minimum wage per employed adult and 74% of the households did not receive social welfare.

Data collection for quality evaluation of the playgroups took place 1 month after the beginning of playgroup implementation (December 2015, T1), and then 1 month before the proposed end (July 2016, T2). We measured playgroup quality at T1 and T2 through direct observation of playgroups. At T1 and T2, we recorded one full session from 12 out of the 13 selected playgroups. At T1, one playgroup video was not collected due to the fact that, on the day of our visit, there were no families present (i.e., all the participants failed to attend). At T2, one playgroup video was not collected due to the fact that the playgroup had already closed. In the end, 24 playgroup video sessions were recorded (see **Table 1**). At T1, 39 (42%) families were present, and the number of dyads in each playgroup ranged from 1 to 5 ($M = 3.33$, $SD = 1.25$). At T2, 34 (41%) families were present, and the number of

dyads in each playgroup ranged from 1 to 6 ($M = 3.25$, $SD = 2.00$). In total, 50 families participated (overlapping). Levels of attendance during the two rounds were a little below of the previous literature on playgroups (57%, Berthelsen et al., 2012).

The mean age of the participating children was 18 months, ranging from 1 to 37 months and 50% of them were younger than 16 months. The mean age of the caregivers was 37 years old, ranging from 15 to 68 years old ($SD = 12.46$), and they were mostly mothers (82%).

The average professional experience of the facilitators that delivered the semiweekly sessions at T1 was 3 years ($M = 3.29$; $SD = 3.12$), all but one had at least 1 year of professional experience working with children and two had 10 years (maximum).

To establish interrater reliability for the PERS, two Ph.D. students with previous experience in using classroom quality measures were trained over a 2-day workshop. In the end, the coders completed a test using one playgroup video (excluded from the final sample), in which the mean interrater percent agreement was 82% based on the consensus scoring within one rating point. The two trained observers coded the 24 videos using the PERS. The observers were blind to the time of monitoring (T1 or T2).

To check the convergent validity of the PERS, the 24 playgroup videos were scored with other measure of process quality: Adult Style Observation Schedule (ASOS, Laevers, 2000). Formal training and certification in the ASOS were provided to the main researcher and another observer by a certified trainer of ASOS.

Ethical Consideration

This study followed the strictest ethical procedures. Prior to data collection, and in accordance to national law prior to the implementation of the GDPR, the Portuguese Data Protection Authority approved all data collections procedures. All research assistants signed a confidentiality agreement. This tool reminded all researchers involved about the ethical limits of their action, the participants' rights and the precautions with data access and sharing. The research assistants also followed scripted guidelines on how to approach families, specifically regarding the timing of the moments of contact and type of discourse that was considered more appropriate in interacting with eligible families. Before the administration of the instruments, caregivers and/or legal representatives of the participating child (when not the same person), facilitators of the playgroups and supervisors were asked to sign the informed consent form, following the ethical guidelines of the European Commission. We collected data only after this consent.

The participation in the study was completely voluntary. The participants were told about the confidentiality of their data and that they could dropout at any time. We did not expect nor found any associated risks or costs. Research assistants also asked the child's assent from all children who could be assumed to understand

the question and provide an answer (i.e., generally all children above age 3) concerning their participation in the child assessments.

In the collection, analysis, access and dissemination of results, we continuously ensure total confidentiality and anonymity of the participants. The participants are not identified in any report or publication. Their privacy is protected using non-identifiable codes and access passwords to the files where the data is stored. The data were collected through paper forms, which are archived in a locked cabinet, accessible only to the evaluation team. The videotapes made are stored in a hard-drive which belongs to the coordinator of the monitoring of implementation.

In terms of data access to participants, ethical requirements, namely the ethical principal of reciprocity, states that all participants should have usable access to the data collected pertaining to their own person, as well as their legal dependent. To maximize this principle, families were informed verbally and in the informed consent form that they were entitled to a report reflecting a summarized profile of their child's development based on the standardized measure used. The request for such report had to be submitted in writing to the evaluation team. Twenty-six requests were received, and the same number of reports was sent to families.

Measures

Griffiths Mental Development Scales

To measure children's development, we used the Portuguese version of Griffiths Mental Development Scales (GMDS; Griffiths, 1954; 0–2 years old: Huntley, 1996; 2–8 years old: Luiz et al., 2007). Based on the Theory of Change for the project Playgroups for Inclusion, we applied only the following subscales: Language, Performance, and Practical Reasoning.

GMDS subscale raw scores were computed by adding the total number of correct items. This computation followed the GMDS 0–2 years Manual (Huntley, 1996) but it was carried out for children of all ages in order to allow comparison of results across the two age groups. **Table 2** presents the descriptive statistics for the subscales. In a previous study, the internal consistency of the GMDS subscales was excellent ($\alpha > 0.90$) for all the three subscales at pretest and posttest (Barata et al., 2017). In the present study, at pretest and posttest the internal consistency was excellent for the subscale language and performance ($\alpha > 0.90$) and good for the subscale practical reasoning ($\alpha \geq 0.80$).

Playgroup Environment Rating Scale

The PERS included 18 items organized under four conceptually defined subscales: Space and Materials, Activities and Routines, Contact with Diversity and Climate and Interactions (**Table 3**). For each item, the observer responded to a series of yes/no indicators that were anchored on a 7-point item. Then the observer applied rules to the pattern of yes/no indicators to determine a score, which were labeled as inadequate (1), minimal (3), good (5), and excellent (7). Because the PERS was designed to be applicable to a range of playgroup practices

TABLE 1 | Data collection for quality evaluation.

Selected playgroups	T1 observation		T2 observation	
	Yes/No	Dyads (n)	Yes/No	Dyads (n)
Playgroup A	Yes	4	Yes	3
Playgroup B	Yes	1	Yes	1
Playgroup C	Yes	2	Yes	2
Playgroup D	Yes	3	Yes	1
Playgroup E	Yes	2	No	0
Playgroup F	Yes	3	Yes	3
Playgroup G	Yes	5	Yes	2
Playgroup H	Yes	3	Yes	2
Playgroup I	Yes	5	Yes	6
Playgroup J	Yes	4	Yes	5
Playgroup K	Yes	3	Yes	3
Playgroup L	Yes	4	Yes	3
Playgroup M	No	0	Yes	3
Total	12	39	12	34

TABLE 2 | Griffiths Mental Development Scales (GMDS) scores at T1 and T2.

Variables	N	Mean (SD)	Range	Skew	Kurt	Alpha
T1 Language	92	36.57 (17.69)	3–73	–0.12	–0.77	0.98
T2 Language	62	51.74 (14.63)	24–78	0.05	–0.96	0.97
T1 Performance	95	37.61 (17.55)	3–66	–0.36	–0.85	0.98
T2 Performance	69	52.20 (11.25)	28–72	–0.40	–0.72	0.96
T1 Practical reasoning	22	4.14 (3.14)	0–12	–0.83	0.53	0.80
T1 Practical reasoning	32	7.78 (5.44)	0–25	0.75	1.70	0.89

We did not compute a global reliability score because we did not use all the subscales in the GMDS.

TABLE 3 | PERS dimensions and items.

PERS dimensions (items)	Description
Space and materials (Indoor space, space organization, materials)	Assesses whether the space complied with the minimum requirements of security, hygiene, and accessibility; whether the playgroup space can be easily arranged in order to accommodate children and adults and playgroup activities and the type and quality of the materials that is used during the sessions.
2. Activities and routines (general routine, activities of the session, free play)	Assesses the associated structure of the playgroups in practice, i.e., greetings and farewells, transitions. Whether the playgroup sessions take into account facilitators' proposals of activities that meet the needs and interests of the participants and the privilege of free exploitation of space and materials by the participating families.
3. Contact with diversity (diversity of dialogues, diversity of practices, diversity of materials)	Assesses the respect and introduction of cultural elements from different countries in sessions and planning of individual needs. Whether the playgroups session include activities and materials that reflect the specific cultures of the playgroup families.
4. Climate and interactions (general climate, interaction children, interaction facilitators, and children, interaction facilitators, and caregivers, interaction caregivers and their children, interaction caregivers and other children, interaction caregivers, interaction supervisors and caregivers, interaction supervisors and facilitators)	Assesses the emotional connection between the participating families, children, and facilitators. Interactions encompass demonstration of care, interest and warmth between all the participants in the playgroups. Whether there are stimulating dialogues, sensibility, and given autonomy to the children, play, and active participation in the sessions and trust, proximity, and collaboration between adults.

(e.g., with or without exterior space, for older and younger children, facilitated playgroups and self-managed playgroups, with or without supervisor) some items and indicators may not be applicable in some contexts. In that case, these items and indicators were coded “Not applicable.”

The scoring of the PERS was based on the observation of one full playgroup session, backed up by information collected from facilitators whenever extra information was needed to

score an item, and a form with basic characterization of the playgroup space.

Adult Style Observation Schedule

The ASOS (Laevers, 2000) assesses the quality of an adult's interactions with a child. ASOS has three dimensions for the quality of interactions: stimulation, sensitivity and giving autonomy. The score of the ASOS was rated on a 7 point-scale,

in which the values 1 and 2 corresponded to predominantly negative behaviors, the values 3, 4, and 5 corresponded to the neutrality, and the values 6 and 7 to predominantly positive behaviors.

ASOS is normally used in formal educational settings. In the context of playgroups, an older version of this scale (Laevers, 1994, with 5 points) was successfully used with slight differences on the methodology (for example, in the days of observation, see Ramsden, 2007). We decided to use the updated version because the score criteria is similar to the PERS.

ASOS coding involves observing each facilitator in two separated days in 4 periods of 10 min. Because this was not possible with our sample, adaptations to the ASOS coding methodology were discussed with and approved by the certified trainer in order to code 40 min of observation in total per facilitator. The mean weighted kappa between the two trained raters on 30% of the videos (7 videos) was moderate ($\kappa = 0.70$) (Fleiss et al., 2003).

Table 4 presents the descriptive statistics for the subscales and overall score of ASOS. The ASOS has demonstrated good internal consistency of the items in the three subscales (Van Heddegem et al., 2004). In the present study, the internal consistency was good for the subscale stimulation ($\alpha = 0.88$) and excellent for the subscale sensitivity ($\alpha = 0.98$) subscale autonomy ($\alpha = 0.91$) and overall score ($\alpha = 0.95$).

Data Analysis

To test the internal consistency of the PERS we calculated Cronbach's alpha for each of the subscales and the total scale. Two items of the PERS (Interaction between Supervisors and Caregivers and Interaction between Supervisors and Facilitators) scored "not applicable" in 80% of our sample, because the supervisors were not present in the sessions. Therefore, ratings on these items were not included in the analyses.

To assess the interrater reliability of the PERS we calculated the linear weighted kappa (Fleiss et al., 2003). The weighted kappa is commonly reported in other studies using the ITERS-R (see Barros and Aguiar, 2010; Barros et al., 2016) or using other measures of preschool quality (ICP; Soukakou, 2012).

To test PERS sensitivity, we conducted a Wilcoxon's signed rank test to analyze whether there were differences between T1 and T2 at subscale level. Because only 12 playgroups were monitored at T1 and T2, we present results for 12 playgroups. Based on similar studies (e.g., Smith-Donald et al., 2007; Barros and Aguiar, 2010), we also presented correlations to test convergent and criterion-related validity analyses.

A preliminary exploratory factor analysis (EFA) was conducted to explore the factor structure of the PERS. Our data yield mean loadings of 0.83 for four factors accounting for 14 variables. According to Winter et al. (2009), our sample size of 24 is sufficient for factor recovery. Following Smith-Donald et al. (2007), we used principal component extraction for standardized version of the 16 items of the PERS. Standardization is recommended when dealing with variables that vary widely

with respect to the standard deviation values of the raw data, which is the case (DiStefano et al., 2009). Resulting components were rotated obliquely using Direct Oblimin to allow correlation between factors. Cronbach's alpha was calculated for each emerging construct and provides an index of internal consistency based on the average of the items scores in the construct. We used IBM SPSS, Version 25.0 for the analyses (IBM Corp, 2017).

RESULTS

Reliability of the Playgroups Environment Rating Scale

Internal Consistency and Interrater Reliability of the PERS

Table 5 shows Cronbach's alpha and correlations for PERS subscales and overall score. The internal consistency for the overall score and for the subscale Climate and Interactions was good ($\alpha = 0.81$ and $\alpha = 0.84$, respectively). The alphas for the subscale Space and Materials and Activities and Routines were questionable ($\alpha = 0.63$ and $\alpha = 0.67$) and for the subscale Contact with Diversity was unacceptable ($\alpha = 0.41$).

Moderate to high correlations were found between the subscales and the overall score (ranged from 0.65 to 0.88), except for the subscale Space and Materials ($r = 0.32$).

Table 5 also presents the mean weighted kappa scores for each of the subscales and the overall score. Interrater exact percent agreement ranged from 35 to 100% ($M = 64$, $SD = 0.20$) and interrater within one scale point percent agreement ranged from 61 to 100% ($M = 85$, $SD = 0.13$). Levels of interrater agreement within one scale point were considered good (Fleiss et al., 2003), and matched expectations from other common quality scales (Barros and Aguiar, 2010; Cadima et al., 2018).

Sensitivity of the Playgroups Environment Rating Scale

Descriptive Statistics and Overall Score

Table 6 presents descriptive information for the items, subscales, and overall score. Overall mean results on PERS ranged from 4.09 to 6.52 ($M = 5.10$; $SD = 0.68$). Subscale means ranged from 3.63 to 6.15 with the lowest scoring subscale being Contact with Diversity, and the highest subscale average occurring in Space and Materials.

Mean results at the item level ranged from 2.21 to 6.80 with the lowest scoring items being Diversity of Materials (item 9), and the highest item averages occurring in Interaction between Supervisors and Facilitators (item 18). All items presented mean scores that indicated presence of minimal quality, except item 9. Five of the 18 items were scored between 1 and 7. Of the remaining 13 items, all presented minimal ratings (between 1 and 3), except for item 1 (Indoor Space), item 12 (Interaction between Facilitators and Children) and item 18, which scored, respectively, 6.00, 5.00, and 6.00 as minimums. All items presented maximum scores (between 6 and 7).

TABLE 4 | Adult style observation schedule subscales and overall score.

Variables	N	Mean (SD)	Range	Skew	Kurt	Alpha
ASOS sensitivity	20	5.97 (0.78)	3.94–7.00	–0.90	0.35	0.94
ASOS stimulation	20	5.24 (0.84)	3.20–6.42	–0.50	0.70	0.88
ASOS autonomy	20	6.27 (0.77)	3.53–7.00	–2.22	6.19	0.91
ASOS overall score	20	5.83 (0.67)	3.74–6.63	–1.36	2.74	0.95

TABLE 5 | Cronbach's alpha, correlations for PERS subscales and overall score and inter-rater agreement of the measure's dimensions and overall score.

	Number of items	Alpha	Correlations	Mean weighted kappa (κ) scores (N = 8)
PERS space and materials	3	0.67	0.32	0.77
PERS activities and routines	3	0.63	0.65*	0.65
PERS contact with diversity	3	0.41	0.65*	0.76
PERS climate and interactions	7	0.84	0.88*	0.69
PERS overall score	16	0.81		0.85

For estimating Cronbach's coefficient alpha and Pearson's correlations, two items (item 16 and 17) were excluded, because these two items were "not applicable" for 80% of our sample.

$p < 0.050$; * $p < 0.010$.

TABLE 6 | Descriptive statistics for the PERS items, subscales, and overall score.

	Sample size	Mean	SD	Minimum-maximum
I. Space and materials	24	6.15	0.66	3.83–6.83
1. Indoor space	24	6.79	0.51	6–7
2. Space organization	24	6.29	0.99	3–7
3. Materials	24	5.38	0.98	3–7
II. Activities and routines	24	5.27	1.07	2.17–6.67
4. General routine	24	4.43	1.38	1–7
5. Activities of the session	24	5.22	1.65	2–7
6. Free play	24	6.15	1.17	3–7
III. Contact with diversity	24	3.63	0.93	1.83–5.67
7. Diversity of dialogues	24	3.54	1.51	1–7
8. Diversity of practices	24	5.13	1.36	1–7
9. Diversity of materials	24	2.21	1.22	1–6
IV. Climate and interactions	24	5.20	1.04	3.04–6.75
10. General climate	24	5.69	1.35	3–7
11. Interaction children	24	4.23	1.35	1–7
12. Interaction facilitators and children	24	6.73	0.53	5–7
13. Interaction facilitators and caregivers	24	5	1.15	3–7
14. Interaction caregivers and their children	24	6.22	1.28	3–7
15. Interaction caregivers and other children	24	4.83	2.19	1–7
16. Interaction caregivers	24	3.69	1.47	1–6
17. Interaction supervisors and caregivers	5	3.80	1.92	2–7
18. Interaction supervisors and facilitators	5	6.80	0.45	6–7
Overall score	24	5.10	0.68	4.09–6.52

Kurtosis and skewness values for item 1, 3, 12, 14, and 18 presented data skewed to the right and a little peaked, which reflected a concentration of scores on higher scale points.

Sensitivity of the Playgroups Environment Rating Scale to Change Over Time

Figure 1 illustrates PERS scores at T1 and T2. PERS overall score at T1 was on average 4.90 indicating that the playgroups were at a moderately high level of quality in the beginning of the playgroup implementation. Averages of the subscales at T1 ranged from 4.83 to 5.92, except for Contact with Diversity that was lower

($M = 3.43$; $SD = 0.94$). The subscale with the highest score was Space and Materials ($M = 5.92$; $SD = 0.80$).

At T2, three subscales and overall score had higher averages than at T1, but differences over time were only significant for Space and Materials at the subscale level. The Wilcoxon's signed rank test indicated that the median T2 scores for this subscale were statistically significantly higher than the median T1 scores ($Z = -1,826$; $p = 0.048$). Effects sizes were equal to 0.088 at the subscale level and ranged from 0.215 to 0.811 at item level. This statistically significant difference reflected a higher score on the item Materials at trend level ($d = 0.811$; $p = 0.098$). The higher score at the subscale Activities and Routines reflected a

statistically significant higher score on the item Free Play ($Z = -1,902$; $p = 0.035$) with effect size equal to 0.979. There were no other significant differences in median scores from T1 to T2 in the other subscales or items.

Validity of the Playgroups Environment Rating Scale

Factor Structure of the Playgroups Environment Rating Scale

Factor analysis on the 16 items of the PERS indicated five components with eigenvalue > 1 (Kaiser rule) reflecting playgroup interactions, use of space and materials, interactions with children and two undefined components. However, the items Indoor Space, Diversity of Materials and Diversity of Dialogues loaded highly onto more than two components, and were thus discarded. Factor analysis was redone excluding these items and only including factor loadings above or equal to 0.32 (Tabachnick and Fidell, 2013). The final structure consisted of four dimensions: Playgroup Interactions, Playgroup Experiences, Play-based Learning and Facilitator Practices and Interactions and explained 77.3% of the variation (Table 7). In the final solution, the item Interaction between Facilitator and Children loaded well into more than one component. Cronbach's alpha if item deleted was calculated for these two components and the item was retained in the component where the alpha was less impaired. The reliability of the final constructs, as determined by Cronbach's alpha, yielded reasonable-to-excellent reliabilities. Most of the items contributed to the good reliability of the dimensions as can be seen from the corrected item-total correlations presented in Table 7. This was especially the case for the items in the Facilitator Practices and Interactions dimension. In this dimension, the obtained Cronbach's alpha was low, but all items showed reasonable correlations with the total score, which attested to the internal consistency of this dimension.

Convergent Validity of the Playgroups Environment Rating Scale

Convergent validity was examined by correlating the PERS dimensions (derived from the PCA) with the ASOS subscales and overall score. As shown in Table 8, scores on the Playgroup Experiences dimension were positively associated with the ASOS Stimulation subscale ($r = 0.62$, $p = 0.001$). Similar patterns of a high degree of correspondence were observed for the Facilitator Practices and Interactions dimension, where correlations were positively associated with the Stimulation subscale ($r = 0.40$, $p = 0.052$). No other relevant correlations for the convergent validity reached significance.

Criterion-Related Validity of the Playgroups Environment Rating Scale

Concurrent validity was examined by correlating the PERS dimensions (derived from the PCA) with group size and facilitators' experience (measured in years). As shown in Table 9, scores on the Playgroup Interactions dimension were significantly and positively associated with the group size

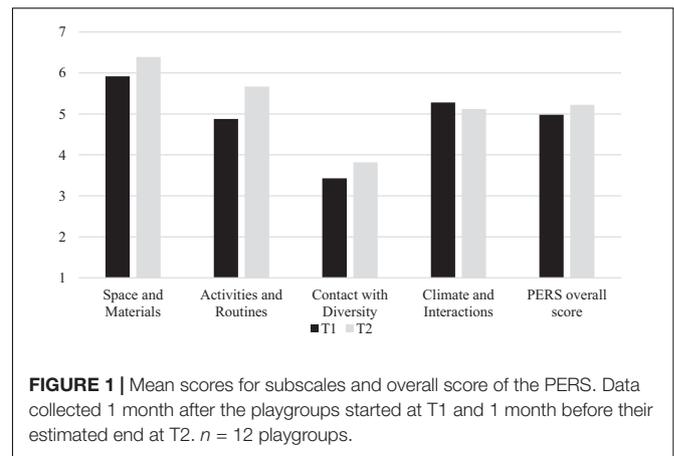


FIGURE 1 | Mean scores for subscales and overall score of the PERS. Data collected 1 month after the playgroups started at T1 and 1 month before their estimated end at T2. $n = 12$ playgroups.

($r = 0.61$, $p = 0.002$). We found no association between the PERS dimensions and facilitator's years of experience.

Predictive validity was examined by correlating the PERS-R dimensions (average of T1 and T2 scores) with the Griffith's subscale scores (average of T1 and T2 scores). As shown in Table 9 scores on the Play-based Learning dimension were significantly and positively associated with the Griffith's subscales of Language ($r = 0.28$, $p = 0.005$) and Practical Reasoning ($r = 0.50$, $p = 0.002$). No other relevant correlations for the predictive validity reached significance.

DISCUSSION

The development of the Playgroups Environment Rating Scale (PERS) involved an interactive process following a classical guide for scale construction (Spector, 1992). The results of this process led to a measure that follows the structure of the ERS and was conceptualized to incorporate four main dimensions: Space and Materials, Activities and Routines, Contact with Diversity, and Climate and Interactions.

In the present study, firstly we tested the reliability and sensitivity of the PERS and secondly we tested the validity of the scale. Results showed a normally distributed overall score for the 24 playgroup observations. Results from the internal consistency are consistent with those reported by Barros and Aguiar (2010) relatively to the Portuguese translation of the ITERS-R (Harms et al., 2012), where the alpha coefficient for the subscale Interactions was higher than the others subscales. These results suggest caution should be used when conducting analyses at the subscale level.

Correlations between PERS overall score and PERS subscales were moderate to high, except for the Space and Materials subscale, which is similar to the correlations found in other study using ECERS-R (Li et al., 2014). Results also noted that trained independent coders were quite consistent in their ratings, suggesting that the PERS can be applied reliably by multiple observers. This is an important finding given that we designed the PERS as an easy-to-use measure that can be administered by professionals and by parents managing

TABLE 7 | Principal components factor analysis of the PERS, corrected item-total correlation, and internal reliability of the measure's dimensions.

Dimensions and items	Loadings	Corrected item-total correlation
Playgroup interactions (eigenvalue = 4.24; α = 0.92)		
10. General climate	0.96	0.92
16. Interaction caregivers	0.93	0.80
11. Interaction children	0.90	0.82
15. Interaction caregivers and other children	0.90	0.87
Playgroup experiences (eigenvalue = 2.84; α = 0.75)		
14. Interaction caregivers and their children	0.85	0.58
5. Activities of the session	0.79	0.63
4. General routine	0.76	0.54
Play-based learning (eigenvalue = 1.72; α = 0.77)		
6. Free play	0.87	0.70
2. Space organization	0.82	0.56
3. Materials	0.79	0.58
Facilitator practices and interactions (eigenvalue = 1.26; α = 0.62)		
12. Interaction facilitators and children	0.60	0.57
13. Interaction facilitators and caregivers	0.89	0.45
8. Diversity of practices	0.68	0.45

Rotation Method: Oblimin Rotation with Kaiser normalization.

TABLE 8 | Bivariate correlations between PERS dimensions with ASOS subscales and overall score.

	ASOS sensitivity	ASOS stimulation	ASOS autonomy	ASOS overall score
1. PERS playgroup interactions	0.24	0.15	0.24	0.25
2. PERS playgroup experiences	0.03	0.62*	-0.14	0.22
3. PERS play-based learning	-0.16	0.18	-0.14	-0.04
4. PERS facilitators practices and interactions	0.11	0.40†	0.13	0.26

† $p < 0.10$; $p < 0.050$; * $p < 0.010$.

TABLE 9 | Bivariate correlations for the criterion-related validity of PERS dimensions.

	Concurrent (N = 24)		Predictive		
	Group size	Facilitators years of experience (M)	Language ^a	Performance ^b	Practical reasoning ^c
1. PERS playgroup interactions	0.61*	0.17	-0.01	0.03	0.05
2. PERS playgroup experiences	0.10	0.18	-0.09	-0.01	-0.10
3. PERS play-based learning	0.04	-0.07	0.28*	0.19	0.50*
4. PERS facilitators practices and interactions	-0.11	-0.05	0.04	0.02	-0.10

Four the predictive analysis we used the means of T1 and T2 of the PERS dimensions and Griffith's subscales.

^a(N = 99), ^b(N = 102), ^c(N = 36).

$\rho < 0.050$; * $p < 0.01$.

community playgroups with families and children with minimal, but adequate, additional training.

As expected, results from the sensitivity of the PERS indicated that playgroup quality was higher at T2 specifically considering the materials used and the opportunities for free play. Playgroup sessions significantly improved the opportunities of contact with everyday materials, disposable materials, and contact with nature, and there were accessible materials that promoted curiosity, discovery, and challenge. Also, playgroup sessions significantly improved the opportunities for children to play freely, and the present adults (caregivers or facilitators) stimulated interest, supported and challenged children in the course of the play. These results provide preliminary evidence that the PERS is sensitive to the continuous training of the facilitators, and changes in playgroup practice.

Results from assessing the PERS structural validity revealed a good factor structure with four distinct but interrelated dimensions: Playgroup Interactions, Playgroup Experiences, Play-based Learning, and Facilitators Practices, and Interactions. However, scores for three items of the PERS were eliminated in this final structure because of their insufficient variability. In our sample, the item Indoor Space was found to be highly skewed, with high averages. This may have been a function of program requirements, i.e., all spaces were carefully assessed before the playgroups began to determine minimal quality standards. Low scores on the items Diversity of Dialogues and Diversity of Materials may have derived from the somewhat low diversity of participating families. Therefore, because of contextual characteristics, we acknowledge that these three items may require further

validation, and support the inclusion of these items in future studies that use the PERS.

The pattern of moderate correlations with the ASOS, another measure of process quality, provides initial support for construct validity. Interestingly, results showed that the PERS dimension of Playgroup Experiences was positively associated with the Stimulation subscale of the ASOS. This dimension includes the routines and activities of the playgroup sessions where the facilitator role is very important to foster stimulating and interesting activities for the families. Results also show a positive association, albeit not statistically significant, between the PERS dimension of Facilitators Practices and Interactions and the Stimulation subscale of the ASOS. This dimension includes the interactions and practices of the facilitator with children, but it also includes interactions and practices with the caregivers, which may explain why it did not reach significance. These positive associations of PERS dimensions with a measure of process quality focused on interactions provide an indication that PERS is a valid assessment of one of the central and specific features of playgroup quality, i.e., interactions between adults and the children but also among adults in different roles (facilitators and caregivers).

Results from the criterion-related validity revealed unexpected and expected results. Results suggested that playgroups with more families were positively associated with more play-based activities, which is contrary to our initial hypothesis. This indicates potential challenges in the implementation of small playgroups, limiting their quality. It is also important to note that our sample was composed by very young children, with a mean age of 18 months. This suggests that playgroups ideally need to be within a range of 4–10 families to allow for more play-based activities to occur for the youngest children, which is aligned with the recommendation by Social Entrepreneurs Inc. (2011). The presence of at least four families in playgroups may generate a greater variety of expressed views and ideas about play materials or may increase the chances of a child playing with a peer of the same age. We did not find a relation between the years of experience of the facilitators and PERS scores, which can be due to the small size of the sample or the lack of variability of the years of experience of the facilitators.

Finally, as expected, the factor dimension of Play-based Learning was positively associated with higher outcomes for children in language and in practical reasoning. This finding is supported by the literature that states that play is associated with the development of language and mathematical ability in children (Department of Education and Training, 2009; Hancock et al., 2012). Effects on the child's ability to use language for comprehension also seem to confirm the playgroup's focus on the promotion of development and the precursors of learning. This focus on development and learning seems to have been paired with an intentionality to promote socialization between the children. All combined, these results add evidence that the PERS is assessing what it has meant to assess.

Some limitations of this study should be acknowledged when interpreting our findings. Although a random sampling procedure was used to select 13 playgroups, as a consequence of low attendance, only data from 12 playgroups was collected.

This condition should be considered when generalizing the results. Also, the restricted variability of facilitators experience may have limited the power to detect statistically significant associations between this variable and the factor dimensions of the PERS. In addition, the internal consistency for the Contact with Diversity dimension was low, which may be due to the lack of variability of the families or the facilitators were less prepared to assure those requirements. We also acknowledged that the three selected subscales of the GMDS do not cover all aspects of children development, but only the related to precursors of learning. This limitation should be considered when generalizing the results. The focus of this paper was to detailed the development of PERS and present the first reliability and validity results. In order to continue to study reliability and validity of PERS, we recommended to have larger, more diverse samples, with careful attention paid to testing their equivalence across different contexts of playgroups, different countries, and in playgroups with children differing by age and by socioeconomical background.

CONCLUSION

Our study provides support for the Playgroups Environment Rating Scale (PERS) as a short, reliable, multidimensional measure of playgroups quality that is relatively inexpensive to administer in field settings. In addition, our findings suggest that this instrument yields data that are psychometrically valid.

A major advantage of this measure lies in its flexibility and sensitivity. The PERS can be applied to different contexts of playgroups and is also useful for informing service planning and practice. The careful selection of indicators can help improve the quality of playgroups by improving practices, and ultimately child outcomes, as outlined in this study.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors if requested.

ETHICS STATEMENT

The study was reviewed and approved by the Portuguese CNPD–Comissão Nacional de Proteção de Dados. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

VR, MCB, and JA contributed to conception, design of the study, performed the statistical analysis, and wrote sections of the manuscript. VR, CL, and BS organized the database. VR wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This study used data of the project “Playgroups for Inclusion” that was supported by the European Union Programme for Employment and Social Solidarity—PROGRESS (2007–2013) under the call VP/2013/012: Call for Proposals for Social Policy Experimentations Supporting Social Investments (Grant Agreement: VS/2014/0418). This study was also supported by

a grant to the VR by Fundação para a Ciência e Tecnologia (PD/BD/128242/2016).

REFERENCES

- Anders, Y. (2015). *Literature review on pedagogy in OECD countries*. Paris: OECD Publishing. Available online at: [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/EDPC/ECEC\(2015\)7&doclang=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/EDPC/ECEC(2015)7&doclang=en) (assessed March 11, 2019)
- Anderson, A. (2005). *The community builder's approach to theory of change. A practical guide to theory development*. New York, NY: The Aspen Institute. Available online at: https://www.theoryofchange.org/pdf/TOC_fac_guide.pdf (assessed April, 2022)
- Barata, M. C., Alexandre, J., de Sousa, B., Leitão, C., and Russo, V. (2017). *Playgroups for Inclusion: Experimental Evaluation and Study of Implementation, Final Report*. Lisboa: University of Coimbra & ISCTE-IUL.
- Barros, S., and Aguiar, C. (2010). Assessing the quality of Portuguese child care programs for toddlers. *Early Childhood Res. Q.* 25, 527–535. doi: 10.1016/j.ecresq.2009.12.003
- Barros, S., Cadima, J., Bryant, D. M., Coelho, V., Pinto, A. I., Pessanha, M., et al. (2016). Infant child care quality in Portugal: associations with structural characteristics. *Early Childhood Res. Q.* 37, 118–130. doi: 10.1016/j.ecresq.2016.05.003
- Berthelsen, D., Williams, K., Abad, V., Vogel, L., and Nicholson, J. (2012). *The Parents at Playgroup Research Report: Engaging Families in Supported Playgroups*. Brisbane, QLD: Queensland University of Technology. Available online at: <https://eprints.qut.edu.au/50875/> (assessed October 15, 2017).
- Berti, S., Cigala, A., and Sharmahd, N. (2019). Early Childhood Education and Care Physical Environment and Child Development: State of the art and Reflections on Future Orientations and Methodologies. *Educ. Psychol. Rev.* 31, 991–1021. doi: 10.1007/s10648-019-09486-0
- Britto, P., Yoshikawa, H., and Boller, K. (2011). Quality of early childhood development programs and policies in global contexts: rationale for investment, conceptual framework and implications for equity. *Soc. Policy Rep.* 25, 1–31. doi: 10.1002/j.2379-3988.2011.tb00067.x
- Cadima, J., Aguiar, C., and Barata, C. M. (2018). Process quality in Portuguese preschool serving children at-risk of poverty and social exclusion and children with disabilities. *Early Childh. Res. Q.* 45, 93–105. doi: 10.1016/j.ecresq.2018.06.007
- Commerford, J., and Hunter, C. (2017). *Principles for high quality playgroups: Examples from research and practice*. Child Family Community Australia Practitioner Resource. Australian Institute of Family Studies. Available online at: <https://aifs.gov.au/cfca/publications/principles-high-quality-playgroups> (assessed March 28, 2019).
- Connell, J. P., and Kubisch, A. C. (1998). “Applying a theory of change approach to the evaluation of comprehensive community initiatives: Progress, prospects, and problems,” in *New approaches to evaluating community initiatives. Volume 2: Theory, measurement, and analysis*, eds K. Fulbright-Anderson, A. C. Kubisch, and J. P. Connell (Washington DC: The Aspen Institute), 15–44.
- Cryer, D., Tietze, W., Burchinal, M., Leal, T., and Palacios, J. (1999). Predicting process quality from structural quality in preschool programs: a cross-country comparison. *Early Childh. Res. Q.* 14, 339–361. doi: 10.1016/S0885-2006(99)00017-4
- Cunningham, J., Walsh, G., Dunn, J., Mitchell, D., and Mcalister, M. (2004). *Giving children a voice: Accessing the views and interests of three-four year old children in playgroup*. Belfast: Stranmillis University College. Available online at: https://www.researchgate.net/publication/228476114_Giving_children_a_voice_Accessing_the_views_and_interests_of_three_four_year_old_children_in_playgroup/citations (assessed June 21, 2019)
- Dadich, A., and Spooner, C. (2008). Evaluating playgroups: an examination of issues and options. *Austr. Comm. Psychol.* 20, 95–104.
- Department for Education (2018). *Official statistics: Childcare and Early Years Survey of Parents in England, 2018*. Early Years Analysis and Research. Available online at: <https://www.gov.uk/government/statistics/childcare-and-early-years-survey-of-parents-2018> (assessed date June 21, 2019)
- Department of Education and Training (2009). *Belonging, being and becoming: The early years learning framework for Australia*. Retrieved from <https://www.dese.gov.au/national-quality-framework-early-childhood-education-and-care/resources/belonging-being-becoming-early-years-learning-framework-australia> (assessed date October 10, 2019)
- Deutscher, B., Fewell, R. R., and Gross, M. (2006). Enhancing the Interactions of Teenage Mothers and Their At-Risk Children: effectiveness of a Maternal-Focused Intervention. *Top. Early Childh. Spec. Educ.* 26, 194–205. doi: 10.1177/02711214060260040101
- DiStefano, C., Zhu, M., and Mindrilă, D. (2009). Understanding and Using Factor Scores: considerations for the Applied Researcher. *Prac. Assess. Res. Eval.* 14, 1–11.
- Fleiss, J., Levin, B., and Paik, M. (2003). “The Measurement of Interrater Agreement,” in *Wiley Series in Probability and Statistics: Statistical Methods for Rates and Proportions*, eds W. A. Shewart and S. S. Wilks (Hoboken, NJ: Wiley Online Library), 2598–2626. doi: 10.1002/0471445428.ch18
- French, G. M. (2005). *Valuing Community Playgroups: Lessons for Practice and Policy, Social Sciences*. Irland: The Katharine Howard Foundation, doi: 10.21427/D7J78M
- Freitas-Luís, J., Santos, L., and Marques, L. (2017). *Playgroups for Inclusion Workpackage 1: Policy Design Final Report*. Coimbra: Bissaya Barreto Foundation & Ministry of Education.
- Garces, E., Thomas, D., and Currie, J. (2002). Longer-Term Effects of Head Start. *Am. Econ. Rev.* 92, 999–1012. doi: 10.1257/00028280260344560
- Ghazvini, A., and Mullis, R. (2010). Center-based care for young children: Examining predictors of quality. *J. Gen. Psychol.* 163, 112–125. doi: 10.1080/00221320209597972
- Gibson, H., Harman, B., and Guilfoyle, A. (2015). Social capital in metropolitan playgroups: A qualitative analysis of early parental interactions. *Austral. J. Early Childh.* 40, 4–11. doi: 10.1177/183693911504000202
- Gormley, W., Gayer, T., Phillips, D., and Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Dev. Psychol.* 41, 872–884. doi: 10.1037/0012-1649.41.6.872
- Griffiths, R. (1954). *The abilities of babies: A study in mental measurement*. New York, NY: McGraw-Hill.
- Hancock, K., Cunningham, N., Lawrence, D., Zarb, D., and Zubrick, R. (2015). Playgroup participation and social support outcomes for mothers of young children: a longitudinal cohort study. *PLoS One* 10:7. doi: 10.1371/journal.pone.0133007
- Hancock, K., Lawrence, D., Mitrou, F., Zarb, D., Berthelsen, D., Nicholson, J., et al. (2012). The association between playgroup participation, learning competence and social-emotional wellbeing for children aged four-five years in Australia. *Austral. J. Early Childh.* 37, 72–81. doi: 10.1177/183693911203700211
- Harman, B., Guilfoyle, A., and O'Connor, M. (2014). Why mothers attend playgroup? *Austral. J. Early Childh.* 39, 131–137. doi: 10.1177/183693911403900417

- Harms, T., Clifford, R. M., and Cryer, D. (1998). *Early childhood environment rating scale: revised edition (ECERS-R)*. New York, NY: Teachers College Press.
- Harms, T., Cryer, D., and Clifford, R. (2006). *Infant/toddler environment rating scale: revised edition (ITERS-R)*. New York, NY: Teachers College Press.
- Harms, T., Cryer, D., and Clifford, R. (2012). *Infant/toddler environment rating scale: revised edition (ITERS-R)*. New York, NY: Teachers College Press.
- Helmerhorst, K., Riksen-Walraven, M., Vermeer, H., Fukkink, R., and Tavecchio, L. (2014). Measuring the Interactive Skills of Caregivers in Child Care Centers: Development and Validation of the Caregiver Interaction Profile Scales. *Early Educ. Dev.* 25, 770–790. doi: 10.1080/10409289.2014.840482
- High/Scope Educational Research Foundation (2001). *High/Scope Program Quality Assessment, PQA-Preschool Version, Assessment Form*. Michigan: High/Scope Educational Research Foundation.
- Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D. M., Clifford, R. M., et al. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childh. Res. Q.* 23, 27–50. doi: 10.1016/j.ecresq.2007.05.002
- Huntley, M. (1996). *The Griffiths Mental Developmental Scales from Birth to Two Years. Manual (Revision)*. Amersham: Hogrefe: Association for Research in Infant and Child Development (ARICD).
- IBM Corp (2017). *IBM SPSS Statistics for Windows, Version 25.0*. Armonk, NY: IBM Corp.
- Laevers, F. (1994). "The Innovative Project Experiential Education and the Definition of Quality in Education," in *Defining and Assessing Quality in Early Childhood Education*, ed. F. Laevers (Leuven: Leuven University Press). 159–172.
- Laevers, F. (2000). *Adult Style Observation Schedule - Form C*. KU Leuven: CEGO.
- Lera, M.-J., Owen, C., and Moss, P. (1996). Quality of educational settings for four-year-old children in England. *Eur. Early Childh. Educ. Res. J.* 4, 21–32. doi: 10.1080/13502939685207901
- Li, K., Hu, B. Y., Pan, Y., Qin, J., and Fan, X. (2014). Chinese Early Childhood Environment Rating Scale (trial) (CECERS): a validity study. *Early Childh. Res. Q.* 29, 268–282. doi: 10.1016/j.ecresq.2014.02.007
- Litjens, I., and Makowiecki, K. (2014). *Literature review on monitoring quality in early childhood education and care*. Paris: OECD Publishing. Available online https://www.europe-kbf.eu/~media/Europe/TFIEY/TFIEY-4_InputPaper/Monitoring-Quality-in-ECEC.pdf (assessed March 11, 2019)
- Luiz, D. M., Barnard, A., Knoesen, M. P., Kotras, N., Horrocks, S., McAlinden, P., et al. (2007). *Escala de Desenvolvimento Mental de Griffiths - Extensão Revista (Revisão de 2006) dos 2 aos 8 anos. Manual de Administração*. Lisboa: Cegoc-Tea.
- Mashburn, A., Pianta, R., Hamre, B., Downer, J., Barbarin, O., Bryant, D., et al. (2008). Measures of Classroom Quality in Prekindergarten and Children's Development of Academic, Language, and Social Skills. *Child Dev.* 79, 732–749. doi: 10.1111/j.1467-8624.2008.01154.x
- McArthur, M., and Butler, K. (2012). *Supported Playgroups and Parent Groups Initiative (SPPI): outcomes evaluation*. Melbourne: Department of education and early childhood development. Available online at: <https://acuresearchbank.acu.edu.au/item/890w9/supported-playgroups-and-parent-groups-initiative-sppi-outcomes-evaluation> (assessed March 28, 2019).
- Melhuish, E. C. (1994). What influences the Quality of care in English playgroups. *Early Dev. Parent.* 3, 135–143. doi: 10.1002/edp.2430030302
- Melhuish, E. C., and Gardiner, J. (2019). Structural Factors and Policy Change as Related to the Quality of Early Childhood Education and Care for 3–4 Year Olds in the UK. *Front. Educ.* 4, 1–15. doi: 10.3389/educ.2019.00035
- OECD (2018). *Engaging Young Children: Lessons from Research about Quality in Early Childhood Education and Care*. Starting Strong. Paris: OECD Publishing, doi: 10.1787/9789264085145-en
- Page, J., Murray, L., Niklas, F., Eadie, P., Cock, M. L., Scull, J., et al. (2021). Parent Mastery of Conversational Reading at Playgroup in Two Remote Northern Territory Communities. *Early Childhood Educ. J.* 2021:1148. doi: 10.1007/s10643-020-01148-z
- Ramsden, F. (2007). The impact of the effective early learning 'quality evaluation and development' process upon a voluntary sector playgroup. *Eur. Early Childh. Educ. Res. J.* 2007, 37–41. doi: 10.1080/13502939785208051
- Salinger, J. (2009). *Supported playgroup evaluation: the Springvale and St. Albans Playgroups*. Richmond, Vic: Odyssey House Victoria & Mary of the Cross Centre.
- Slot, P. (2018). *Structural characteristics and process quality in early childhood education and care: A literature review*. OECD Education Working Paper No. 176. Netherlands: Utrecht University, doi: 10.1787/edaf3793-en
- Slot, P., Leseman, P., Verhagen, J., and Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early childhood education and care settings. *Early Childh. Res. Q.* 33, 64–76. doi: 10.1016/j.ecresq.2015.06.001
- Smith-Donald, R., Raver, C. C., Hayes, T., and Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Childh. Res. Q.* 22, 173–187. doi: 10.1016/j.ecresq.2007.01.002
- Social Entrepreneurs Inc., (2011). *Best Practices in Playgroups Research review and quality enhancement framework. For First 5 Monterey County (F5MC) Playgroups Serving Children 0-3 Years Old*. Reno, Nevada, USA. Available online at: <https://www.first5monterey.org/userfiles/file/F5MCBestPracticesinPlaygroups.pdf> (assessed March 28, 2019)
- Soukaku, E. (2012). Measuring quality in inclusive preschool classrooms: Development and validation of the Inclusive Classroom Profile (ICP). *Early Childh. Res. Q.* 27, 478–488. doi: 10.1016/j.ecresq.2011.12.003
- Spector, P. E. (1992). *Sage university papers series: Quantitative applications in the social sciences, No. 82. Summated rating scale construction: An introduction*. Thousand Oaks, CA: Sage. doi: 10.4135/9781412986038
- Statham, J., and Brophy, J. (1992). Using the 'Early Childhood Environment Rating Scale' in playgroups. *Educ. Res.* 34, 141–148. doi: 10.1080/0013188920340205
- Strange, C., Fisher, C., Howat, P., and Wood, L. (2014). Fostering supportive community connections through mothers' groups and playgroups. *J. Adv. Nurs.* 70, 2835–2846. doi: 10.1111/jan.12435
- Sylva, K., Siraj-Blatchford, I., and Taggart, B. (2003). *Assessing quality in the early years: Early Childhood Environment Rating Scale-Extension (ECERS-E): Four curricular subscales*. Stoke-on Trent. Sterling, VA: Trentham Books.
- Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliot, K., et al. (2006). Capturing quality in early childhood through environmental rating scales. *Early Childh. Res. Q.* 21, 76–92. doi: 10.1016/j.ecresq.2006.01.003
- Tabachnick, B. G., and Fidell, L. S. (2013). *Using Multivariate Statistics*. Boston: Pearson.
- Thomason, A., and La Paro, K. (2009). Measuring the quality of teacher-child interactions in toddler child care. *Early Educ. Dev.* 20, 285–304. doi: 10.1080/10409280902773351
- Vandell, D. L. (2004). Early child care: the known and the unknown. *Merrill-Palmer Q.* 50, 387–414. doi: 10.1353/mpq.2004.0027
- Van Heddegem, I., Gadeyne, E., Vandenberghe, N., Laevers, F., and Van Damme, J. (2004). *Longitudinaal Onderzoek in het Basisonderwijs. Observatie-Instrument Schooljaar 2002-2003*. Leuven: Steunpunt SSL. Available online https://informatieportaalssl.be/archieflooplanen/rapporten/LOA-rapport_20.pdf (assessed May 11, 2022).
- Vermeer, H., van IJzendoorn, M. H., Cárcamo, R. A., and Harrison, L. J. (2016). Quality of Child Care Using the Environment Rating Scales: a Meta-Analysis of International Studies. *Internat. J. Early Childh.* 48, 33–60. doi: 10.1007/s13158-015-0154-9
- Walsh, G., and Gardner, J. (2005). Assessing the quality of early years learning environments. *Early Childh. Res. Prac.* 7, 1–17. doi: 10.1080/09575146.2017.1342223
- Weiland, C., Ulvestad, K., Sachs, J., and Yoshikawa, H. (2013). Associations between classroom quality and children's vocabulary and executive function skills in an urban public prekindergarten program. *Early Childh. Res. Q.* 28, 199–209. doi: 10.1016/j.ecresq.2012.12.002
- Weiss, C. H. (1995). "Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families," in *New approaches to evaluating community initiatives: Concepts, methods and contexts*, eds J. P. Connell, A. C. Kubisch, L. B. Schorr, and C. H. Weiss (Washington, DC: The Aspen Institute), 65–92.
- Williams, K. E., Berthelsen, D., Nicholson, J. M., and Viviani, M. (2015). *Systematic literature review: Research on supported playgroups*. Queensland University

- of Technology, Australia. Available online: <https://eprints.qut.edu.au/91439/1/91439.pdf> (assessed October 10, 2017).
- Winter, J., Dodou, D., and Wieringa, P. (2009). Exploratory Factor Analysis With Small Sample Sizes. *Multiv. Behav. Res.* 44, 147–181. doi: 10.1080/00273170902794206
- Wuermli, A., Yoshikawa, H., and Hastings, P. (2021). A bioecocultural approach to supporting adolescent mothers and their young children in conflict-affected contexts. *Dev. Psychopathol.* 2021, 1–13. doi: 10.1017/S095457942000156

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Russo, Barata, Alexandre, Leitão and de Sousa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.