Early Childhood Research Quarterly 55 (2021) 79–96

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

# Preschool practices in Sweden, Portugal, and the United States

Vera Coelho[a,*], Frida Åström[b], Kimberly Nesbitt[c], Madelene Sjöman[b], Dale Farran[d], Eva Björck-Åkesson[b], Caroline Christopher[d], Mats Granlund[b], Lena Almqvist[b,e], Catarina Grande[a], Ana Isabel Pinto[a,†]

[a] Faculty of Psychology and Educational Sciences, Porto University, Rua Alfredo Allen, 4200-135 Porto, Portugal
[b] CHILD, Swedish Institute for Disability Research, Jönköping University, Box 1026, 551 11 Jönköping, Sweden
[c] University of New Hampshire, Department of Human Development & Family Studies Pettee Hall, 55 College Road, Durham, NH 03824, United States
[d] Vanderbilt University, Peabody Research Institute, PMB 181, Nashville, TN 37203, United States
[e] School of Health, Care, and Welfare, Mälardalen University, Box 883, 721 23 Västerås, Sweden

ARTICLE INFO

*Article history:*
Received 29 March 2019
Received in revised form 8 September 2020
Accepted 9 November 2020

*Keywords:*
International
Early Childhood Education and Care (ECEC)
Preschool practices
International early childhood policies
Behavioral count measures
Observation

ABSTRACT

Across countries, there are important differences related to the goals, organization, and educational philosophies of care provided to young children prior to formal schooling. Those differences are likely reflected in the classroom practices and teacher-child interactions within a country's early childhood education and care (ECEC) classrooms. This study aims to evaluate the within-country relevance of two classroom observation measures primarily based on a behavioral count approach focused on teacher and child behaviors; and to examine preschool practices in Sweden, Portugal, and the U.S., as they reflect each country's ECEC goals, organization, and educational philosophies. Participants are 78 preschool settings in Sweden, 42 in Portugal, and 168 in the U.S. Results show that the measures targeted culturally-relevant behaviors and provided inter-rater reliability for the behavior count variables in the three countries. Future collaborations may address additional culturally-specific variables. The behavioral descriptions yielded by combining behavioral counts of the measures are analyzed by researchers from the relevant country for insights to the country's values related to early childhood as well as current debates regarding care for children. Measures that provide comprehensive descriptions of classroom settings and apply minimal external or comparative value judgments on the behaviors observed are of practical utility for collaborative international work.



## 1. Preschool practices in Sweden, Portugal, and the United States

The 21st century has seen steady growth in early childhood education and care (ECEC) driven in large part, at least in Europe, by the need to support parents with young children who increasingly combine employment with caring responsibilities (Cohen & Korintus, 2017). International comparisons of participation in and policies regarding ECEC have been published by UNICEF (2013), the World Bank (2013), the Organization for Economic Cooperation and Development (OECD, 2017), and the International Association for the Evaluation of Educational Achievement (Bertram & Pascal, 2016). In Europe, an ambitious study of ECEC (the CARE project) was begun in 2014 and resulted in several important reviews of policies and practices (e.g., Moser, Leseman, Melhuish, Borekhuizen, & Slot, 2017; Broekhulzen, Leseman, Moser, & van Trijp, 2015; Melhuish et al., 2015). Over the past decades in the United States (U.S.), early childhood education core elements of classroom practice have also been widely studied (e.g., Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Farran, Meador, Christopher, Nesbitt, & Bilbrey, 2017; Keys et al., 2013).

In European countries, as well as in countries like Australia and Canada (Bradbury, Corak, Waldfogel, & Washbrook, 2015), the provisions for young children tend to be different from the U.S. For instance, the proportion of children served by ECEC is lower in the U.S. than in other countries with comparable levels of prosperity. Among U.S. states, public funding is limited and eligibility criteria differ, prompting Barnett to assert each state in the U.S. is like a European country with its own individual policies and practices (quoted in Jacobson, 2019). If a child does not meet stipulated eligibility criteria for publicly funded programs, a family may not have options for access to other providers of ECEC. In contrast, many European countries, such as Portugal and Sweden, provide

* Corresponding author.
*E-mail address:* veracoelho@fpce.up.pt (V. Coelho).
† Died in May 2019.

https://doi.org/10.1016/j.ecresq.2020.11.004
0885-2006/

universal access to ECEC, with more than 90% of children aged 4 years and over attending out-of-home care (European Commission, 2015). Such structural differences across countries/contexts introduce challenges to understanding the importance of preschool practices in each country, and confound the identification of a single, common measure of quality (Fischer & Poortinga, 2018; Sheridan, Giota, Han, & Kwon, 2009).

### 1.1. Observing ECEC: rating scales

Rating scales to measure ECEC quality have been developed since the 1980s, almost all developed in the U.S. Applying such context-specific measures to classrooms in other countries has been termed "ethically dubious" (Kline, Shamsudheen, & Broesch, 2018). In the U.S., rating scales have been used in an evaluative manner, with funding for public programs tied to scores (Farran & Nesbitt, 2019). Rating scales consider the type of materials available, the way the day is structured and, to different extents, micro-level interactions (Farran & Nesbitt, 2019). Internationally, the Early Childhood Environment Rating Scale (ECERS; Harms, Clifford, & Cryer, 2004) and the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008) are among the most frequently used measures (Brunsek et al., 2017), with both strengths and weaknesses highlighted in the literature (Perlman et al., 2016).

Several studies describe the utility, adaptations and psychometric characteristics of these rating scales in different countries (e.g., Cadima, Aguiar, & Barata, 2018; Garvis, Sheridan, Williams, & Mellgren, 2017; Gordon, Fujimoto, Kaestner, Korenman, & Abner, 2013; Vermeer, van IJzendoorn, Cárcamo, & Harrison, 2016). The most frequently mentioned weaknesses for rating scale systems are reliability and validity, as observational rating scales can be subjective and therefore associated with rater error and bias problems (Grammatikopoulos, Gregoriadis, & Zachopoulou, 2015; Hoyt & Kerns, 1999; Mashburn, 2017). General classroom rating scales are likely more sensitive to variations in raters' values, biases and specific experiences of preschool practices (e.g., Cash, Hamre, Pianta, & Myers, 2012; Mashburn, 2017); these differences are likely to vary among raters in different countries (Boer, Hanke, & He, 2018). Mashburn, Meyer, Allen, and Pianta (2014) found that differences among observers' ratings, when using rating scales, were the largest source of error variance for the CLASS as did Cash et al. (2012). In general, the more inference required to rate an item, the greater the bias and unreliability of the raters (Li & Karakowsky, 2001)

Additionally, general rating scales often reflect an underlying value system, thus increasing the challenge of comparing ECEC practices across countries with differing goals and strategies for preschool educational settings (Pastori & Pagani, 2017; Rentzou, 2017). Moreover, general rating scales do not identify differences in practices at a behavioral level, which can be examined through the additional lenses of a specific country-level ECEC system organization, philosophy, and focus.

### 1.2. Observing ECEC: behavioral counts

Although rating scales have been adopted by European researchers and others, a more objective and behaviorally-oriented method of measurement, such as counting the frequency of specific and identifiable behaviors, might be less susceptible to observer bias (Briesch, Chafouleas, & Riley-Tillman, 2010; Hoyt & Kerns, 1999). Compared to rating scales, behavioral counts or systematic direct observations are more objective as they indicate only the frequency of certain behaviors (Briesch et al., 2010). In behavioral count measures, the task of the observer is to determine whether a certain behavior is present, as opposed to determining how well a preschool performs on general, predetermined aspects. As such, behavior counts provide more concrete micro-level behavior information that can then be interpreted within the scope of a country's specific ECEC system and challenges. It must be stressed, however, that the selection of which behaviors to observe and count is still subject to values.

While behavioral count measures may be less vulnerable to observer drift and cultural biases, the use of behavior counts requires adequate training and exact inter-rater agreement standards (as opposed to within one-point agreement usually reported for general classroom rating scales; Mashburn, 2017). Measures must include concrete operational definitions to ensure the reliability and validity of the behavioral counts. Moreover, collecting counts continually over a period of time may be cognitively demanding, requiring the total focus of the observer. Low frequency behaviors are also of concern, with lower reliability for behaviors that occur less often (Dowsett, Huston, Imes, & Gennetian, 2008).

Despite those concerns, behavior counts have the potential to facilitate highly reliable measures. Behavior counts are flexible enough to be used in different ECEC settings, focus on both individual child and teacher behaviors, and provide important starting points for discussion about practices of ECEC within and across different countries.

### 1.3. International research considerations

Campbell-Barr and Bogatić (2017) describe the recent global interest in ECEC and its growing "thirst for evidence that demonstrates 'effective' practice" (p. 1461), yet acknowledge no clear common conception for how to define "effective". In the U.S., preschool is often defined as an intervention, compensatory for children presumed to need remedial help before school entry. As such, U.S. decisions about quality indicators have been based primarily on how well classroom practices predict gains in cognitive skills and to a lesser extent social-emotional growth (Farran & Nesbitt, 2019). ECEC has been a priority area for the EU since 1992 (European Agency for Special Needs & Inclusive Education, 2017), but a council recommendation on high-quality ECEC emerged only recently from the European Council (2019).

In March 2020, OECD released the first report from its International Early Learning Study (IEL, OECD, 2020), which has been termed the "baby PISA", in reference to the OECD's Programme for International Student Assessment (PISA) measuring 15-year-olds' reading, mathematics and science knowledge and skills to meet real-life challenges in the 36 member states/countries. As PISA has come to be a common metric for international comparisons of high-school student achievement, the IEL study raised concern among researchers and early childhood advocates (Jacobson, 2019). The initial IEL work compares five-year-olds in the U.S., England, and Estonia on emergent literacy and numeracy outcomes, emotional knowledge, and ratings from parents and teachers of social skills. OECD collected data on the childcare experiences of all the children in the study and compared the test scores of those who had and had not experienced early education. Of concern is that this kind of work could set the foundation for later evaluations of the quality of preschool programs internationally, based on how much they affect growth in these areas (e.g., Moss et al., 2016).

In a 1989 study of preschool in three different cultures, Tobin and colleagues discussed "cultural relativism" (Tobin, Wu, & Davidson, 1989). In this work and subsequent publications, Tobin has argued that how children are cared for, especially in early childhood, is deeply rooted in a country's history and culture. "At its core, cultural relativism is an epistemological rather than a moral concept which argues that it is intellectually and methodologically unsound to attempt to understand another people's cultural practices using the assumptions and categories of one's own culture" (Tobin, 2005, p. 425). The concept of cultural relativism challenges

the creation of decontextualized quality standards in early childhood education.

Our study addresses the tension between cultural relativism and possible cross-cultural quality assessments of ECEC. For three countries, we provide detailed behavioral descriptions of the types of classroom practices and interactions among children and between teachers and children. The classroom practices were identified by the three countries as a first step in our collaboration, having previously been related to child outcomes valued in some U.S. preschool settings (Farran et al., 2017). Importantly, researchers in each of the countries have been intimately involved in the interpretation of the observations within the particular context and values of their early childhood education systems. The three preschool contexts are not presented in a comparative framework; rather we demonstrate the value of clear behavioral observations to assist individual countries in determining whether preschools reflect the values – and perhaps identify the controversies – of ECEC within the country.

Because each of the countries will be highlighted separately, this paper is organized somewhat differently from most. First, we present the measures used for all three countries, cultural adaptations performed for use in Sweden and Portugal, and general procedures. Then, information about each country is presented separately, including descriptions of the country's ECEC context, the country-specific sample, results (including inter-rater reliability), and discussion. Structural differences among the three countries included the group of children served (i.e., at-risk populations, mainstream preschool for all children), the proportion of time spent in teacher-led versus child-initiated activities, the amount of time children interact with each other, and the percentage of the day occupied with academic-focused instruction. Finally, a general discussion on the utility and challenges of using an observational approach in different countries is provided.

## 2. General method

### 2.1. Measures

The *Teacher Observation in Preschool* (TOP; Bilbrey, Vorhaus, & Farran, 2007) and the *Child Observation in Preschool* (COP; Farran & Son-Yarbrough, 2001) were originally developed in the U.S. to assess preschool practices, with some studies showing potential for use in other countries with minor adaptations (e.g., Coelho & Pinto, 2018; Luttropp & Granlund, 2010). Study-specific adaptations will be presented later.

The measures are primarily behavioral count systems indexing behaviors shown to capture unique and important aspects of preschool quality associated with children's gains across and beyond preschool (Farran et al., 2017; Fuhs, Farran, & Nesbitt, 2013; Spivak & Farran, 2016). The COP and TOP are used in tandem to collect data on each individual in a setting through a 3-second "sweeping" method. The COP allows the observer to capture children's behavior in preschool settings throughout the day (see Appendix A for definitions of main COP/TOP categories). A systematic behavior-sampling procedure collects information on (a) children's listening and verbal behaviors (including to whom the behavior was directed), (b) schedule of classroom activities, (c) children's proximity to others, (d) level of social interaction, (e) type of task (e.g., passive instruction, pretend play, etc.), (f) learning focus of activities (e.g., literacy or math), and (g) level of involvement (e.g., assess how focused and engaged the child is). Each child is located, observed, and then after a count of 3 seconds, the observer immediately codes all categories. Similarly, the TOP also uses a systematic procedure to capture teachers' behaviors, including (a) teacher's listening and verbal behaviors (including to whom), (b) schedule of classroom activities, (c) teacher proximity to others (d) teacher tasks (e.g., instruction, administrative, personal/care, etc.), (e) materials, (f) learning focus of activities, (g) level of instruction (e.g., no instruction, high inferential instruction), and (h) teacher tone and affect. While the measures are primarily behavioral count systems, rating scales are used to index child level of involvement, teacher tone and level of instruction.

In both COP and TOP, each category is coded independently of the others, meaning that behaviors can be examined individually or as they co-occur. (One reason, for example, to code "schedule" in both COP and TOP is that the type of activity, such as level of involvement in small groups facilitated by a teacher, can then be used as a context for examining its effects on other behaviors).

The observer starts by coding each teacher's behaviors and then moves to each child, coding his or her behaviors one-by-one until each person has been observed. Then, the observer starts the process over, continuing to cycle through each person across a daylong observation. Thus, each participant in the classroom, each teacher and each child, is observed an equal number of times. These data can then be aggregated to the classroom level to summarize the teachers' and children's experiences during the day. The system allows individuals to be observed in a variety of settings with enough behavioral episodes to be aggregated with some reliability (Hoyt & Kerns, 1999).

In the present study we have chosen to use data from the following COP categories: schedule, listening and verbal behaviors, interaction state (captures social learning, i.e., social parameters of the child's learning experience including associative and cooperative activities providing opportunities for children to interact with one another and to learn to take another's point of view; Farran & Son-Yarbrough, 2001), type of task (to capture demands of the learning task), learning focus of activities, and the level of involvement (overall, during each type of schedule, and in learning opportunities); and TOP categories: listening and verbal behaviors, teacher tasks, level of instruction, and teacher tone and affect. Narrowing the behaviors examined allows each country to examine the same set of behaviors within the particular context.

Coding options for each category are mutually exclusive (e.g., the observer had to choose from among the options within each of the categories rather than endorsing more than one) and most scores were computed based on summed and aggregated behavioral counts. One rating scale item is included in COP: Level of Involvement (1 = low involvement, to 5 = high involvement); and two are included in TOP: Tone (1 = extremely negative, to 5 = vibrant and enthusiastic) and Level of Instruction (0 = no instruction to 4 = high inferential instruction).

### 2.2. Adaptations for Portuguese and Swedish Contexts

Due to the variability in preschool delivery across the different contexts, pilot observations and discussions among the international researchers prior to data collection in Portugal and Sweden ensured that the measures were as meaningful, yet as precise and equivalent as possible (Hui & Triandis, 1985). Though some language adaptations were necessary, English was kept as the instrument language. All observers in both Portugal and Sweden were fluent in English.

While major setting differences in Portugal and Sweden required changes in COP/TOP to allow for contextualization, few of the modifications or additions were included in the analysis presented in this study. Measures were adjusted to consider: (a) greater and different use of outdoor space, (b) settings including children with disabilities, and (c) mixed-age group classrooms including toddlers. For instance, the Swedish preschool philosophy emphasizes the outdoors as an important learning space where teachers organize activities and facilitate learning. To consider the outdoors as a component of the measure (as it was not coded in

the measures in the U.S.), a set of examples to guide observers in coding COP/TOP during outdoor activities was created. Examples include (1) considering that outdoor spaces are usually larger than indoor classrooms, the distance proposed by the COP/TOP authors to code proximity was increased for outdoors; (2) different interaction states for the outdoors coded parallel play as when two children were sitting in the sandbox and pouring sand in a bucket without talking; or if two or more children rode bicycles on the playground but were not engaged with each other, such as by racing or chasing each other.

The Portuguese and the Swedish systems also required adaptation of the measure for preschool settings serving children with disabilities or in need of support; specific examples were added to the coding manuals for this population, particularly regarding the level of involvement. Similarly, although U.S. and Portuguese preschool classrooms do not include toddlers, Swedish preschools commonly do. Therefore, examples were created for observing the level of involvement for toddlers. For instance, high levels of involvement were coded when a child tried to gain the attention of the teacher on a specific material or activity by grabbing clothing or a hand. A medium level of involvement was coded when a child attended to a task that another child or teacher was carrying out and tried to imitate the behavior of others within the activity. Portugal and Sweden decided not to use the automatic low code for involvement when the interaction state was coded as social because in these countries' curriculum guidelines, social interactions are considered important processes that drive development and learning.

For this paper, only the codes for behaviors connected to the preschool practices known to be related to quality in U.S. settings are examined. The composite practices are comprised of 11 behavior count variables (transitions, instruction, behavior approving, behavior disapproving, teacher listening to a specific child, teacher listening to children, non-sequential activities, sequential activities, associative interactions, cooperative interactions, math and literacy learning focus) and three variables derived from rating scales (Child Level of Involvement, Teacher Tone, and Teacher Level of Instruction).

Pilot observations in Sweden and Portugal were conducted to test the validity and reliability of the adapted measures. As the COP/TOP measures observable behaviors, not scales of latent constructs, the validity of the adapted components was examined by the frequency of occurrence of all coding options during training. Researchers agreed that the code captured relevant behaviors in each country and therefore the preschool reality. During training, inter-rater reliability achieved adequate overall exact agreement values (over 80% for the rating scale; over 90% for behavior counts in both countries).

One adaptation made in the data from Portugal and the U.S. was to average the data from the "lead" teacher and assistant in each classroom. Sweden does not distinguish roles for the multiple adults in the classroom and could not separate the data from that of a head teacher versus an assistant.

### 2.3. Procedures

#### 2.3.1. Training

Prior to data collection, all observers received theoretical and practical training and achieved acceptable reliability. Observers were trained to use both measures with materials provided by the U.S. researchers. In addition, one or more lead researchers from Sweden and Portugal received individual training in person from the U.S. team in a train-the-trainer model. The one-week intensive training included theoretical sessions, video coding, and discussions, as well as two days of observations in preschool settings. Additionally, both Portuguese and Swedish data collectors received training together (in Portugal) from U.S. trainers. Following this coordination with U.S. researchers, each country trained its local observers and assessed inter-rater reliability.

#### 2.3.2. Data collection

All observers from all countries coded the study variables using the same operational definitions. Double coding of observations for inter-rater reliability was conducted in the field, during data collection, for each country. Each pair of local observers coded the exact same time period for the target children and teachers during the same observations. Two preschool settings, not included in the study sample, were double coded for inter-rater reliability in Sweden; twenty-three percent of data collection observations were double-coded for Portugal; and double coding was obtained for 15.4% of the observations in the U.S. Exact percent agreement and Cohen's κ were computed for inter-rater reliability in all countries (provided in each country's Results and Discussion section). For the continuous rating scale items, intraclass correlations (ICCs) were also computed.

Although similar data collection procedures were planned, and minimum standards were defined, there were some differences between the countries' data collection procedures due to each country's preschool organization, funding for this specific project, and other logistical issues (e.g., parental consent). A minimum number of "observation sweeps" per child was defined (12) to account for the different amount of hours of observation collected in each country. All countries started observations at the beginning of the preschool day, according to a set of common data collection procedure requirements. Data from all children and teachers observed in each country were aggregated at the preschool classroom level for data analysis and further consideration within each country.

Ethical procedures were followed according to each country's national regulations, with informed consents and legal authorizations collected prior to the study.

### 2.4. Analytic approach

COP and TOP are formatted in a matrix with "sweeps" in rows and behaviors in independent columns. This format allows behaviors to be summed and compared, and the creation of conditional probabilities across behaviors (e.g., the level of child involvement during whole group). Variables were computed as the sum of individual scores across the observations and then aggregated to the preschool setting level to yield an understanding of practices. Behavioral counts (11 variables) were further computed as proportions of sweeps in which the target behavior occurred out of the total number of sweeps observed, while the variables derived from rating scales (three variables) were computed as averages across all sweeps observed.

To create variables that captured the proportion of sweeps for a particular code, we used conditional probability looping syntax. This method instructs the statistical software to search through a group of variables of the same category (e.g., Schedule codes) and counts/sums the number of instances in which a certain code was used (e.g., transitions). After that count variable was created, we calculated a proportion. The count variable was treated as the numerator while the total number of times any schedule code was recorded was treated as the denominator (e.g., sum of transition sweeps / sum of all schedule sweeps). Some of the preschool practices were captured using a combination of different variable categories and thus required slightly more complex calculations. For example, examining the percentage of time teachers were listening to a child specifically required that we create syntax to simultaneously search two groups of variables, namely Verbal and To Whom.

The third type of conditional variables drew from a combination of counts and rating scales. To create the Level of Involvement (rating, 1 = low involvement to 5 = high involvement) by Schedule (count), looping syntax indicated that all schedule sweeps should be searched for a specific code representing the type of schedule. For example, to calculate level of involvement during whole group activities, the looping syntax would identify sweeps coded as whole group and then create a sum of the involvement ratings during all the whole group sweeps. An average of the involvement during whole group was then created by dividing the sum of all whole group involvement ratings by the number of whole group sweeps.

## 3. Sweden

### 3.1. Preschool context in Sweden

The Swedish universal ECEC system is regulated by the Swedish Education Act (2020) includes all children aged 1–5 years, and is the first part of lifelong learning within the Swedish educational system. Over 85% of Swedish children between one to five years are enrolled in preschool, and over 95% of children between four and five. The curriculum of preschool is decided by the government (Swedish National Agency for Education, 2019b) and is based on holistic, inclusive, and ecological principles. Designed for care, development, and learning to form a whole, Swedish preschool is part of the welfare state, family policies focused on dual-earning families, and the ambition to provide a good start in life for all children. Democracy is stressed as a fundamental pillar, and the development and learning of all children is promoted. Children's participation and influence on their education are emphasized. The curriculum states preschool should offer children a good environment that is accessible for all children, and a balanced daily rhythm with both rest and activities adapted according to the children's needs and length of stay. The environment is intended to inspire children to play together and explore the world around them. Children are to be offered varied activities in different contexts, both indoors and outdoors.

Play is considered the foundation for development, learning, and well-being, yet over time academic learning and teaching have become more emphasized. Pramling Samuelsson, Williams, Sheridan, and Hellman (2016) summarized the pedagogical approach as being "one whereby both the foundation of academic knowledge and the tradition of a wholeness with play, care and learning should be integrated." (p. 446). It is the responsibility of the preschool teacher to organize pedagogical activities to promote the principles stated in the curriculum (Einarsdottir, Purola, Johansson, Broström, & Emilson, 2015), yet specific pedagogical methods are not identified. The National Agency for Education seeks to ensure that Swedish education maintains a standard of quality through national school development programs and training programs. The Swedish Schools Inspectorate evaluates ECEC to improve quality and outcomes.

Among preschool staff – including both teachers and assistants – 40% have a 3.5-year academic preschool teacher education (bachelor's degree). Assistants, also called child-minders, have backgrounds that span upper secondary education and academic degrees. Usually, there is more than one preschool teacher in each unit ("unit" being the equivalent of a U.S. "classroom"), although this can vary between and within preschools. There is a strong tradition of teamwork, and roles are often shared among the team members.

The physical environment of Swedish preschools has a functional design intended to provide a stimulating indoor and outdoor environment and to foster children's interaction with peers and teachers (Westberg, 2019). A typical preschool has three units, with each unit usually targeting younger children (1–3 years) or older children (4–5 years). On average, the group size is 16.1 for children ages 4–5 years, and 12.6 for the younger children. Teachers-child ratio in general is 1:5, with fewer children per adult for units with younger children (Swedish National Agency for Education, 2019a). Many young children spend a large part of everyday life in preschool and have individualized schedules based on their caregivers' working hours. A summary of Swedish preschools' characteristics is presented in Table 1.

The Swedish preschool system tends to score well in international quality assessments (OECD, 2017) and has also the highest expenditure per child in all OECD countries (OECD, 2015). Ratings are mainly based on structural conditions, such as workforce, subsidized services, group-size, teacher-child ratios, and universal access. A recent report (Swedish Schools Inspectorate, 2018) indicated quality differences in Swedish preschools contrary to the fundamental value of equivalent education stated in the curriculum. Quality differences were also found in a previous study (Pramling Samuelsson & Sheridan, 2009) assessing preschool environments for the youngest children (1–3 years), using an adapted version of ECERS (Harms & Clifford, 1980). To describe the diverse environments and activities in preschool is complex, and structural elements are not sufficient. This quantitative observational study aims to assess how preschool practices look from an empirical perspective, focusing on the processes and activities children and teachers are exposed to and experience across the day in preschool.

### 3.2. Swedish sample

The participating Swedish sample includes 78 preschool units from two different projects, Early Detection Early Intervention (TUTI) and Participation and Engagement in Preschool International (PEPI). Preschool units from PEPI ($n$ = 39) were selected by a combination of purposeful and convenience sampling to increase the probability that enough children with special support needs were included. Preschool units from TUTI ($n$ = 39) were selected based on convenience sampling with the purposefully selected municipalities representing communities of varying sizes and population density. The total sample was in the mid-south to southeast region of Sweden, with two units in large cities (>200,000 inhabitants), 23 units in smaller or rural municipalities, and 53 units in, or close to, mid-size cities. Compared to Sweden as a whole, the sample underrepresented units from large cities, and overrepresented units from mid-size municipalities. The sample consisted of both public and private non-profit settings: 24% of the units were private, comparable to current Swedish national data (28%; Swedish National Agency for Education, 2019a). Most of the units had mixed ages ($n$ = 70), while six units served single age groups (3-, 4-, or 5-year-olds). On average, units had 20.70 children ($SD$ = 6.59) enrolled, with 4.09 ($SD$ = 5.68) second language learners of Swedish, and 0–1 child with disabilities ($M$ = 0.73, $SD$ = 0.98). Teacher-child ratios varied between 1:3 and 1:8, with an average of 1:5. Across the units, a total of 302 teachers and 925 children were observed, with child ages ranging from 16 to 72 months ($M$ = 52.55, $SD$ = 11.02). On average, 82% ($SD$ = 22) of the children in the unit were observed.

### 3.3. Swedish procedures

Observations were conducted from September to December of 2014 (TUTI) and 2015 (PEPI). All observers ($N$ = 3) were project employees, females, with education ranging from a bachelor's degree in behavioral science to a Ph.D. Each unit was observed once, for about 7 h, starting at approximately 8.00 a.m., with an average of 16.52 ($SD$ = 6.65) sweeps per child being collected.

**Table 1**
Characteristics of early childhood education and care in Sweden.

| Characteristic | Description |
|---|---|
| Responsible | Ministry of Education and Research (SFS 2010:800). Preschool is the first part of the school system and is open for all children from the ages of 1–5 years. The municipality is responsible for offering a place in a preschool as close to the child's home as possible. |
| Main goal | All children shall learn and develop knowledge and values and enhance each child's development and a life-long desire to learn. It shall also convey and secure respect for human rights and basic democratic values (Swedish National Agency for Education, 2019b). |
| Curriculum | The national curriculum (Swedish National Agency for Education, 2019b) focuses on what the preschool shall provide to all children in terms of care, development and learning. |
| Main pedagogical principles | The curriculum (Swedish National Agency for Education, 2019b) states that a) the education should contribute to children developing an understanding of themselves and their environment, b) exploration, curiosity and a desire to play should be the foundation of the education, c) the education should take its starting point in the curriculum and in children's needs, experiences and what they show an interest in, and the flow of children's thoughts and ideas should be utilized to create diversity in learning, and d) give special support to those children who for various reasons need it in their development. It is the responsibility of the preschool teacher to organize pedagogical activities to promote the principles. |
| Inclusion | The Swedish Education Act (2020) states that the preschool is for all children. Inclusion is not mentioned in the curriculum since it is clearly outspoken that preschool is universal. |
| Child age | 1–5 years. The characteristics of the group differ, from all children 1–5 years in the same group to smaller groups with only toddlers and larger groups with 3–5 years. |
| Type | Public 72% (municipality) and independent 28% (non-profit) (Swedish National Agency for Education, 2019a). National financing via the municipalities and the curriculum cover both types. |
| Enrollment rate & opening hours | 2018 data: 1–5 years 85%, 4–5 years 95% (Swedish National Agency for Education, 2019a). Preschools are open throughout the year and have no general opening hours; they are usually open Monday to Friday between 6.30 a.m. and 6.30 p.m. |
| Training | Preschool teachers: 3.5-year academic education bachelor's degree; Child-minders/preschool assistants: from no education to upper secondary education and academic degrees. |
| Ratios | Group sizes can vary based on the children's needs and the conditions of the units. On average, the group size is 16.1 for children 4–5 years, and 12.6 for the younger children.<br>Staff-child ratio in general is 1:5, with fewer children per adult for units with younger children (Swedish National Agency for Education, 2019a). The number of teachers in the unit varies during different periods of the day. |
| Physical design | The design of the preschool environment varies, but usually consists of a hall, a play hall/larger room, a kitchen and dining room, several smaller activity rooms, and napping areas. Much time is spent outdoors year around. |

**Table 2**
Descriptive statistics of classroom schedule in a sample of preschool classrooms from Southern Sweden.

| Schedule | *M* (*SD*) | Range |
|---|---|---|
| Whole-Group | .08 (.06) | .00–.25 |
| Small-Group | .04 (.05) | .00–.22 |
| Centers[a] | .57 (.13) | .19–.85 |
| Small-Group/s & Centers[b] | .04 (.06) | .00–.23 |
| Transition | .13 (.06) | .03–.27 |
| Meal-time | .13 (.04) | .02–.23 |
| Other[c] | .01 (.02) | .00–.11 |

*Note.* $N = 78$ classrooms. Variables represent the proportion (range 0–1) of observational sweeps a given practice was observed.

[a] Centers (free play) refer to relatively high child choice in activities and include free play activities both inside the unit and outdoors.

[b] Small Group/s & Centers refer to simultaneous occurrence of Small Group/s and Centers.

[c] Other included gym, nap, and specials.

### 3.4. Swedish results and discussion

All variables presented adequate reliability estimates, except for Level of Involvement in COP and for Instruction and Tone in TOP. For COP, the average exact agreement was 80.3%, with values ranging between 60.8% (Level of Involvement) and 92.6% (Schedule), and Cohen's κ ranging between .55 (Level of Involvement), and .90 (Schedule). For TOP, the average exact agreement was 84.9 %, ranging between 72.6% (Task) and 93.6% (Schedule), and Cohen's κ ranged between .11 (Tone) and .92 (Schedule). Considering the continuous nature of the rating scale items, ICCs were also computed to estimate inter-rater reliability. Reliability was lower in the Swedish sample for the rating scale items of Level of Involvement (ICC = .003), Tone (.39) and Level of Instruction (.39).

Summaries of the observed preschool schedule are presented in Table 2, where centers or free play was the most frequent activity in the Swedish preschool units. Whole group and small group activities, and simultaneous use of small group/s and centers were relatively rare. Perhaps related to the extended periods in centers, and the circulation of individual children in and out of activities, group-level transitions occurred for only 13% of the observational day. The relatively high use of centers, indicating high child choice in activities, aligns with the Swedish curriculum, which emphasizes the importance of children's play and exploration and their influence over their education (Swedish National Agency for Education, 2019b). A similar pattern of free play has been observed in Norway (Storli & Hansen Sandseter, 2019), a country with comparable curriculum values and ideas (Einarsdottir et al., 2015).

Observed classroom practices are presented in Table 3. Preschool teachers were observed instructing for less than a fifth of the observational day, with an average level of instruction below basic skills ($M = 1.55$). The average level of instruction varied across preschool units, however, potentially a result of teachers' freedom in focusing on specific content, or the composition of the child-group (e.g., Vallberg Roth, 2020). In other words, teachers might have different views on the value of teacher-led activities and plan activities accordingly. Relatedly, the philosophy of the Swedish preschool curriculum states that children's learning and development should be based on their interest and exploration of the environment (Swedish National Agency for Education, 2019b), indicating a more passive role by the teachers to not interfere with children's exploration of materials or activities with other children. The tendency by Swedish teachers to step back from children's play and interaction and merely set the scene for play, thereby adopting a managing role, may lead to lower levels of instruction in general.

The role of the teacher in teaching or instruction is highly debated in Swedish preschool research and practice (e.g., Sheridan & Williams, 2018). Instruction could be considered "education," which Swedish teachers may find a challenging task, not to be

**Table 3**
Descriptive statistics of classroom practices in a sample of preschool classrooms from Southern Sweden.

| Variable | *n* | *M* (*SD*) | Range |
|---|---|---|---|
| Instruction provided by the teachers | | | |
| Instruction | 78 | .17 (.09) | .00–.40 |
| Level of instruction (while instructing)[a] | 78 | 1.55 (0.37) | 0.75 – 2.17 |
| Emotional climate in the classroom | | | |
| Behavior approving | 78 | .06 (.07) | .00–.33 |
| Behavior disapproving | 78 | .04 (.04) | .00–.19 |
| Teacher tone[b] | 78 | 3.25 (0.15) | 2.89 – 3.80 |
| Teachers and children talking and listening | | | |
| Children talking (overall) | 78 | .29 (.07) | .12–.44 |
| Children talking to teacher | 78 | .05 (.02) | .01–.12 |
| Teacher talking (overall) | 78 | .51 (.11) | .24–.78 |
| Teacher listening to a specific child | 78 | .04 (.03) | .00–.19 |
| Teacher listening to children (overall) | 78 | .05 (.04) | .00–.22 |
| Type of activity | | | |
| Non-sequential activities | 78 | .34 (.10) | .14–.54 |
| Sequential activities | 78 | .11 (.06) | .00–.28 |
| Social learning (associative and cooperative interactions) | | | |
| Associative interactions | 78 | .19 (.11) | .00–.45 |
| Cooperative interactions | 78 | .03 (.03) | .00–.11 |
| Levels of child involvement[c] | | | |
| Overall involvement | 78 | 2.39 (0.29) | 1.86 – 3.12 |
| Involvement in whole-group | 70 | 2.90 (0.49) | 1.80 – 4.00 |
| Involvement in small-group | 49 | 2.91 (0.68) | 1.00–4.50 |
| Involvement in centers | 78 | 2.79 (0.35) | 2.09 – 3.58 |
| Involvement in small-group/s & centers | 45 | 2.72 (0.69) | 1.00–4.00 |
| Involvement in learning opportunities | 78 | 2.80 (0.34) | 2.15 – 3.60 |
| Academic focus | | | |
| Math focus | 78 | .03 (.03) | .00–.12 |
| Literacy focus | 78 | .07 (.04) | .00–.20 |

*Note.* Sample size = 78 classrooms. Reported *n*s reflect the number of classrooms a given practice was observed. Unless otherwise indicated, variables represent the proportion (range 0–1) of observational sweeps a given practice was observed.
[a] Range: 0 (no instruction) to 4 (high inferential instruction).
[b] Range: 1 (extremely negative) to 5 (vibrant/enthusiastic).
[c] Range: 1 (low involvement) to 5 (high involvement).

emphasized in preschool. Rather, teachers may endeavour to engage children in learning activities without formally educating them (Jonsson, Williams, & Pramling Samuelsson, 2017). On the other hand, the definition of "instruction" in COP/TOP is a wide concept, including any activity where the teacher interacts with a child or children on a learning topic, and not necessarily related to formal education. The topic can be typically academic (e.g., math or literacy) or the focus can be art, music, puzzles, or blocks. Swedish preschools include a relatively high frequency of free play activities; Goble et al. (2016) reported that free play correlates with less instruction, in comparison to teacher-led activities such as whole group contexts.

As for the emotional climate, teachers provided more behavior approving than disapproving, and the emotional tone was on average neutral towards slightly positive ($M$ = 3.25). Although teachers provide more behavior approving, their tone tended toward neutral. A possible explanation for the neutral tone might be that teachers are fairly passive towards the children, adopting a managing role setting the scene for child activities, similarly indicated by teachers' relatively low amounts of instruction and listening behaviors, as described below.

Teachers spoke in half of the observations but were seen listening to children in only five percent of the observations. Children talked less frequently than teachers and quite seldom to a teacher. The results indicate that relatively little of the communication in preschool occurs between teachers and children. A recent study (Vallberg Roth, 2020) stressed that teachers seem to trust children's own choices of activities and competence, and merely pay attention to what children are learning on their own. However, the low amount of teachers' listening to children was unexpected considering that the curriculum states the importance of letting children's voices be heard (Swedish National Agency for Education, 2019b).

Concerning math and literacy, children were observed interacting with a literacy focus twice as much as they were with a math focus. To address curriculum goals, Swedish pedagogical scholars stress the importance of combining informal learning, when children play together without the involvement of teachers, with specific teacher-led literacy and communication activities (Pramling Samuelsson, 2010). But the low incidence of observed math learning is surprising since the curriculum explicitly emphasizes the importance of children's "understanding of space, time and form, and the basic properties of sets, patterns, quantities, order, numbers, measurement and change, and to reason mathematically" (Swedish National Agency for Education, 2019b, p.19).

Regarding social learning, associative interactions occurred in about a fifth of the observations, while cooperative interactions were infrequently observed. Across the observational day, the average level of child involvement was between medium and medium-low. The highest level of involvement was noted in small- and whole-group activities, whereas the lowest was during free play (small-group/s & centers). This is surprising as a recent Norwegian study (Storli & Hansen Sandseter, 2019) found a strong positive relation between the extent of play and children's involvement. A potential reason could be that teacher-led activities are rare in Swedish preschools and therefore more easily elicit children's interest. In other preschool environments where free play seldom occurs, the free play might elicit more interest because it is rare and, therefore, more attractive.

The results from the Swedish study indicate that behavioral count observation instruments developed in the U.S. can be implemented, if adapted, with an acceptable degree of reliability. The use of behavioral counts provided a unique examination of practices in Swedish preschool units, which have a focus on child-directed collaboration and exploration to support holistic development.

Information such as the variation in the use of teacher-led activities, the low amount of listening behaviors evidenced by the teachers, and the relatively low level of child involvement in free play activities can provide Swedish authorities an empirically based understanding of actual interactions in preschool and inform teacher preparation and Sweden's national curriculum.

## 4. Portugal

### 4.1. Preschool context in Portugal

The Portuguese ECEC system is organized in two main services, namely childcare (for children aged 0–3) and preschool education (for children aged 3–6), with a common framework and directions for teachers' practices (e.g., Silva, Marques, Mata, & Rosa, 2016). Although preschool attendance is optional, there is a high rate of enrollment (European Commission/EACEA/Eurydice, 2019), with approximately 95% of 5-year-olds attending preschool education programs (European Commission, 2015).

As described in Table 4, the main goal for preschool education is to provide opportunities for children's development, autonomy, and socialization, while contributing to success at formal schooling. National guidelines aim to ensure national congruence across all preschool institutions (Silva et al., 2016). All Portuguese preschools are inclusive, meaning that all classrooms are open to receive and support all children, including those with disabilities. Moreover, children at risk are given priority in entering preschool education, and specific recent legislation emphasizes providing early intervention strategies to meet the needs of all children (DL 281/2009; DL54/2018). The legislation defines children "at-risk" as those who have any biological and/or environmental risk factors. The eligibility criteria for additional support for at-risk children is defined in terms of the number of biological or social risk factors to which the child is exposed (DL 281/2009). Regarding staff training and qualifications, it is mandatory to have one lead teacher with a degree in preschool education for each classroom (four-year, higher education degree), responsible for planning classroom activities, organizing the classroom-learning environment and materials, and interacting in a positive and responsive manner.

In Portugal, previous studies using measures such as the ECERS-R (Harms et al., 2004) and the Quality of Inclusive Preschool Experience Measures (QIEM; Wolery, Paucca, Brashers, & Grant, 2000), reported that both inclusive and non-inclusive preschool settings presented mediocre levels of general quality (e.g., Abreu-Lima et al., 2014; Abreu-Lima, Leal, Cadima, & Gamelas, 2013; Aguiar, Moiteiro, & Pimentel, 2010). Many studies of preschool quality in Portugal assess the classroom level, not analyzing individual children's behaviors or identifying contingencies between specific behaviors of children and teachers. Thus, assessing specific aspects of classroom processes that occur contingently using a behavior count measure to document teacher and child behaviors, represented a chance to obtain in-depth knowledge about actual preschool experiences.

High variability in pedagogical models used by teachers in Portuguese preschools exists, as national curriculum guidelines provide no mandatory model, but rather state an overall direction and broad principals and goals for ECEC. Therefore, this study can contribute to understanding how daily practices in preschools are aligned with the national recommendations on (a) organizing children's schedules; (b) promoting children's agency in learning and developmental processes; (c) considering child involvement as pivotal for learning and development; (d) focusing on different pre-academic domains; and (e) ensuring the teacher's role as a facilitator of children's learning (Silva et al., 2016).

### 4.2. Portuguese sample

Participants include 42 inclusive preschool settings from the north of Portugal, including both rural (33%) and urban settings. Portuguese classrooms were selected from the list of preschools registered at the Ministry of Education website from a district in the north of Portugal. Following a random list of numbers, 80 ECEC school directors were contacted; 26 directors did not agree to participate; 12 institutions were excluded given that they did not meet the inclusion criteria (the center needed to have at least one child eligible for early childhood intervention or special education support services in the classroom).

The remaining 42 ECEC classrooms, serving children 3–6 years old, met the criteria and agreed to be part of the study. Of the 42, 38 were public institutions, while only four were private non-profit. No private for-profit institutions participated. Most classrooms were mixed-age ($n$ = 37), a common approach in Portuguese public preschool (81.4% mixed-age classrooms according to the national survey; Abreu-Lima et al., 2014). Participant classrooms had on average 20.42 children ($SD$ = 2.49), and 1–4 children with disabilities ($M$ = 1.58, $SD$ = 0.68). Overall, 247 children were observed across settings, with ages ranging from 31 to 66 months ($M$ = 49.92, $SD$ = 6.71). Teacher-child ratios varied between 1:5 and 1:26, with an average of 1:13. Teachers were all female with ages ranging between 27 and 59 years old ($M$ = 50.04, $SD$ = 6.58). All teachers had at least a 4-year degree in preschool education.

### 4.3. Portuguese procedures

Observations were conducted between January and March 2016 by trained researchers with a degree in Psychology ($N$ = 4). An average of 29% of children per classroom had parental consent to be included in the study. Observations started at approximately 9 a.m., accommodating the Portuguese preschool schedules and routines, and lasted for an average of 3 h, with an average of 20.87 ($SD$ = 2.17) sweeps per child being collected. Each classroom was observed once.

### 4.4. Portuguese results and discussion

Considering the lack of information about current practice in Portuguese preschools and how teachers' behaviors are aligned with national guidelines, this study aimed to portray teachers' and children's ECEC experiences. To our knowledge, this was the first study using COP/TOP in Portugal and although more research is needed, we believe this constitutes a first piece of evidence on the relevance of using these measures, orienting research to micro-level processes and events across the preschool day.

Appropriate inter-observer agreement values were found using COP/TOP in Portugal. The average exact agreement for COP was 94.7%; values across categories ranged between 78.6% (Level of Involvement) and 98.0% (Schedule). The average Cohen's κ was .91; values across categories ranged between .65 (Level of Involvement) and .99 (Interaction State). The average exact agreement for TOP was 97.2%; values across categories ranged between 90.9% (Tone) and 99.2% (Level of Instruction, Focus, and Materials). The average Cohen's κ for TOP was .94; values across categories ranged between .77 (Tone) and .99 (for Tasks). ICC for Level of Involvement were also acceptable at .92, .91 for Tone, and .99 for Level of Instruction.

Portuguese observed preschool schedules (see Table 5) showed that the highest proportion of the day was spent in whole group activities (nearly 50%), while lower proportions of time were observed in centers (12%) or small-group (3%). Transition time occurred an average of 10% of the observational sweeps, and other activities (i.e., gym, special activities, outdoors) were coded for 11% of the observation sweeps. This distribution of activities in

**Table 4**
Characteristics of early childhood education and care in Portugal.

| Characteristic | Description |
|---|---|
| Responsible | Ministry of Education (Law 5/1997). Preschool is the first part of the school system and is open for all children from the age of 3. Before age 3, parents have childcare services, with the Ministry of Labor and Social Welfare being the main responsible entity. |
| Main goal | To favor all children's development and potentialities, providing opportunities for the development of children's autonomy and socialization, while contributing to success at formal schooling. To provide important support to families in their children's educational processes (Silva et al., 2016). |
| Curriculum | Curriculum guidelines (Silva et al., 2016) focus on several content areas (e.g., math, literacy, social studies, motor skills, arts, socio-emotional development . . .).<br>Diverse learning formats are recommended, including whole-group, small-group, individual and free-play activities. |
| Main pedagogical principles | Main pedagogical principles state that: (a) learning and developmental processes are inseparable; (b) the child is an active agent in learning processes; (c) all children must receive adequate support (*based on the assumption that all children have the right to see their needs, interests and capacities adequately supported and valued, and the right to participate in the group activities*); and (d) construction of learning must be articulated. |
| Inclusion | All preschools are inclusive. Early Childhood Intervention in Infancy - Decree-Law 281/2009, 2009); Inclusive Education - Decree-Law 54/2018, 2018.[a] |
| Child age | 3 to 5 |
| Type | Public; Private non-profit; and Private for profit. |
| Enrollment rate & opening hours | 2016 data: 3-year-olds = 83%; 4-year-olds = 90%; and 5-year-olds >95%. Preschools are usually open between 8 a.m. and 7 p.m., with an educative component - teacher working hours - of 5 h per day. Remaining hours are mainly a service for families who need more time of care for their children. |
| Training | Lead teacher with at least a superior degree in preschool teaching; Assistants with no mandatory training. Teachers and assistant have different roles in the classroom (hierarchical role definition). |
| Ratios | 1 mandatory, full-time, lead teacher plus a part-time assistant for up to 25 children. |
| Physical design | Preschool design varies between schools. Sometimes preschools are a part of elementary school settings institutions.<br>Preschool is usually organized in small classrooms, with a recommended space per child of 2 square meters, and with shared spaces such as dining room, play hall, and outdoor space. |

[a] Replacing the previous decree-law (3/2008) that regulated supports for children with special education needs till September 2018.

**Table 5**
Descriptive statistics of classroom schedule for a sample of preschool classrooms in North of Portugal.

| Schedule | *M* (*SD*) | Range |
|---|---|---|
| Whole-Group | .47 (.19) | .06–.83 |
| Small-Group | .03 (.09) | .00–.38 |
| Centers[a] | .12 (.16) | .01–.69 |
| Small-Group/s &Centers[b] | .10 (.14) | .01–.46 |
| Transition | .10 (.07) | .01–.29 |
| Meal-time | .07 (.03) | .01–.17 |
| Other[c] | .11 (.10) | .00–.37 |

*Note.* $N = 42$ classrooms. Variables represent the proportion (range 0–1) of observational sweeps a given practice was observed.

[a] Centers or free play refer to relatively high child choice in activities. Centers occurred only inside the classroom and were coded only inside the classrooms.

[b] Small Group/s & Centers refer to the simultaneous occurrence of Small Group/s and Centers.

[c] Other included gym, specials, and the playground.

preschools aligns with previous national studies (e.g., Abreu-Lima et al., 2014; Coelho & Pinto, 2018). Although diverse learning formats are recommended in the Portuguese national guidelines, including whole group, small group, free play, and outdoors, results suggested that classrooms could improve the balance in the children's schedule. For instance, outdoor play offers children a wide range of opportunities positively linked to wellbeing, health, development, and learning (e.g., Kalpogianni, 2019). In Portugal, however, going outdoors is often considered "recess time". Teachers (not always present during such moments) or assistants usually only monitor and manage children's behavior rather than providing intentional learning opportunities, despite recommendations from the national guidelines.

Instruction time occurred frequently (37% of the day) but focused on basic skills or general interactions with materials, despite a high amount of time spent instructing children in whole group (Table 6). This finding may indicate that teachers are not fostering interactions that promote child agency and active roles in learning through a mix of open-ended questions that would support increased participation.

Emotional climate was documented through counts of behavior approval and disapproval and scores (using a rating scale) of teachers' tone and affect. While sums for approving and disapproving are important, so is the ratio between them. Counts of disapproval were twice that of approval (though both approval and disapproval were low). Regardless, teachers' tone was, on average, between neutral to positive.

Regarding verbal behaviors, teachers were coded talking more than half of the observation (55%) and listening for 11%. This high amount of teachers talking might be explained by the high proportion of time in whole group activities. Similarly, a high frequency of teachers talking may indicate that child agency and their active role in learning processes were not being considered, although this is a main pedagogical principle in the national guidelines (Silva et al., 2016).

Children, on the other hand, talked on average for 36% of the observation period, with most of their verbal behaviors directed to teachers (26% of observation time). These observations may indicate that children were in a rote responding mode (as a group) in whole-group activities that required singing songs or answering a teacher's questions collectively.

Children were observed in a higher proportion of sequential (e.g., activities following a sequence of steps) than non-sequential activities. Associative interactions (interactions in activities/tasks that do not have predetermined rules) were observed 11% of the time, while cooperative interactions (i.e., interactions activities/tasks with common goals, rules, and organization) were almost never observed (3%). Despite such low occurrence, the Portuguese ECEC goal related to children's autonomy and socialization skills contributing to future school success (Silva et al., 2016) stresses cooperative interactions as crucial.

Overall, Portuguese curriculum guidelines underline the need to follow children's interests and promote their involvement (Silva et al., 2016). Results showed that children's overall involvement level was below the mid-level on a five-point scale. A very similar

**Table 6**
Descriptive statistics of classroom practices for a sample of preschools in North of Portugal.

| Variable | *n* | *M* (*SD*) | Range |
|---|---|---|---|
| Instruction provided by the teachers | | | |
| Instruction | 42 | .37 (.12) | .00–.90 |
| Level of instruction (while instructing)[a] | 42 | 1.37 (0.29) | 1.00–3.00 |
| Emotional climate in the classroom | | | |
| Behavior approving | 42 | .02 (.02) | .00–.17 |
| Behavior disapproving | 42 | .05 (.05) | .00–.33 |
| Teacher tone[b] | 42 | 3.43 (0.37) | 0 – 4.32 |
| Teachers and children talking and listening | | | |
| Children talking (overall) | 42 | .36 (.12) | .15–.47 |
| Children talking to teacher | 42 | .26 (.11) | .07–.63 |
| Teacher talking (overall) | 42 | .55 (.12) | .00–1.00 |
| Teacher listening to a specific child | 42 | .04 (.05) | .00–.29 |
| Teacher listening to children (overall) | 42 | .11 (.08) | .00–.29 |
| Type of activity | | | |
| Non-sequential activities | 42 | .19 (.08) | .06–.42 |
| Sequential activities | 42 | .31 (.12) | .12–.58 |
| Social learning (associative and cooperative interactions) | | | |
| Associative interactions | 42 | .11 (.07) | .00–.26 |
| Cooperative interactions | 42 | .03 (.04) | .00–.17 |
| Levels of child involvement[c] | | | |
| Overall involvement | 42 | 2.90 (0.28) | 2.38–3.67 |
| Involvement in whole-group | 42 | 2.69 (0.30) | 2.10–3.47 |
| Involvement in small-group | 14 | 3.05 (0.47) | 2.40–4.00 |
| Involvement in centers | 22 | 3.29 (0.40) | 2.50–4.18 |
| Involvement in small-group/s & centers | 23 | 3.34 (0.47) | 2.75–5 |
| Involvement in learning opportunities | 42 | 2.90 (0.33) | 2.30–3.79 |
| Academic focus | | | |
| Math focus | 42 | .04 (.08) | .00–.40 |
| Literacy focus | 42 | .09 (0.1) | .00–.37 |

*Note.* Sample size = 42 classrooms. Reported *n*s reflect the number of classrooms a given practice was observed. Unless otherwise indicated, variables represent the proportion (range 0–1) of observational sweeps a given practice was observed.
[a] Range: 0 (no instruction) to 4 (high inferential instruction).
[b] Range: 1 (extremely negative) to 5 (vibrant/enthusiastic).
[c] Range: 1 (low involvement) to 5 (high involvement).

average was observed when analyzing involvement in schedules with learning opportunities (i.e., not observed in transition, mealtime, and other). The lowest involvement levels were coded during whole group, while higher involvement levels were coded during the simultaneous occurrence of small group/s and centers. While Portuguese curriculum guidelines identify child involvement and interests as pillars of child learning, yet as shown above, teachers devote more of the day to structured/whole group schedules. Teachers may be assuming that structured/whole group schedules are more effective in promoting children's learning, being unaware of the fact that such activities actually elicit lower involvement levels among children.

Portuguese curriculum guidelines (Silva et al., 2016) suggest that teachers should approach several content areas throughout the day, including math, literacy, social studies, gross and fine motor skills, arts, socio-emotional development, all recognized as important for children's development across preschool years (Bratsch-Hines, Burchinal, Peisner-Fineberg, & Franco, 2019; Duncan et al., 2007; Farran et al., 2017; Santos & Alves Martins, 2011; Silva et al., 2016; Ulferts, Anders, Leseman, & Melhuish, 2016). Results showed that opportunities for children to engage in math activities occurred for an average 4% of the observation time, with some classrooms spending no time in math activities. Literacy activities were observed twice as often as math, at 9%. However time in a focus on literacy was very uneven across classrooms with some classrooms spending no time in literacy-related activities (similar to results reported for math), while others spent about 37% of the observation (Table 6). The high amount of whole group activities and the stated curriculum areas of content (Table 4) might have anticipated a higher focus on academic content. Regardless, these low values are consistent with previous research in Portugal (e.g., Abreu-Lima et al., 2013, 2014).

These results highlight important considerations for the design of ECEC in Portugal. As teachers spent about half of their day in structured/whole group, activities that were associated with low child involvement levels, it is important to consider strategies to help teachers be aware of the frequency of their own behaviors and of children's behaviors. The efficacy of initial and in-service training to help teachers link the Portuguese guidelines with specific classroom practices must be considered in light of the possible disconnect between guidelines and observed practices revealed here. The identification of strategies to help facilitate alignment between national guidelines and children's ECEC experiences may facilitate more positive outcomes for children in Portugal.

## 5. United States

### 5.1. Preschool context in the United States

In 2017, about 69% of the four-year-old children in the United States were enrolled in some form of early childhood education and care (ECEC) center-based programs (Institute of Education Sciences, NCES, 2019). Center-based care in the U.S. includes publicly-funded programs such as Head Start (comprehensive early childhood education for low-income families), Title I (federal supplemental funds for high poverty school districts) and/or state-funded prekindergarten/preschools, as well as private non-profit and for-profit childcare centers (see Table 7 for a summary of U.S. characteristics of ECEC). While the number of children in these programs has remained stable, many states have recently shifted to educating a larger portion of four-year-olds in state-funded prekindergarten programs (Friedman-Krauss et al., 2018), from 14% in 2002 to 33% in 2018.

**Table 7**
Characteristics of early childhood education and care in the United States.

| Characteristic | Description |
|---|---|
| Responsible | Multiple federal and state agencies most often including Departments of Education and Departments of Health and Human Services or Welfare |
| Main goal | There is no national goal set for ECEC or pre-kindergarten programs. Head Start's purpose is to "promote the school readiness of low-income children by enhancing their cognitive, social and emotional development" (https://eclkc.ohs.acf.hhs.gov/policy/head-start-act/sec-636-statement-purpose). Individual states may have early learning standards regarding content in state-funded pre-k programs. |
| Curriculum | Similarly, there is no national curriculum. Programs are free to choose from published curricula. 75% of Head Start programs use Creative Curriculum, a global, general developmental curriculum (Moiduddin et al., 2017). |
| Main pedagogical principles | No agreed-upon pedagogical principles. The federal Preschool Development Grant competition supports States to (1) build or enhance a preschool program infrastructure that would enable the delivery of high-quality preschool services to children, and (2) expand high-quality preschool programs in targeted communities that would serve as models for expanding preschool to all 4-year-olds from low- and moderate-income families. |
| Inclusion | All federally supported programs (e.g., Head Start, Preschool Development Grant-Expansion) must be inclusive. |
| Child age | Head Start: 3–5. PDG-E: 4–5 |
| Type | Public; Private non-profit; and Private for-profit |
| Enrollment rate & opening hours | 2016 data: 3-year-olds: 41.6%, 4-year-olds: 65.9%<br>Federally funded programs are typically 5–6 hours a day. Private childcare can be the entire workday for families. |
| Training | Mixed requirements depending on the program.<br>For Head Start, about 50% of teachers have bachelor's degrees. For PDG-E funded programs, all teachers must be certified, licensed teachers. For state-funded prekindergarten programs, requirements vary by state. For private childcare programs, a high school diploma required. |
| Ratios | Varies by program: For programs following National Association for the Education of Young Children (NAEYC) guidelines, group size of 20 and adult-child ratio of 1–10 is expected. |
| Physical design | Design varies by program type. Generally, licensing requirements state that an early childhood program must have 35 square feet of usable indoor activity space per child, and 75 square feet of accessible outdoor space or in certain cases, indoor space, for active gross motor play. Classrooms in public schools, however, do not have to abide by these figures. |

Historically in the U.S. the purpose of caring for young children was to allow parents to work (Farran & Nesbitt, 2019; Scarr & Weinberg, 1986). The shift to publicly-funded preschool has emerged from a belief in the importance of intervening early to prevent the pernicious effects of high-poverty environments (Parker, Workman, & Atchison, 2016). In the U.S., ECEC varies widely by the source of its funding, and the state and local contexts.

Head Start, Title I, and most state-funded programs, including the prekindergarten programs included for this study, provide five to six hours of care a day, nine months of the year (i.e., a school year calendar), while non-profit and for-profit childcare centers tend to provide year-round programs with extended hours to match the U.S. 8-h workday. State-funded programs, in contrast to traditional childcare centers, tend to have admission requirements; most are means-tested with family income having to be at or below a certain income level for eligibility. Kindergarten readiness frequently appears as a primary objective of U.S. public prekindergarten.

In contrast to most European countries, U.S. regulations for preschools and childcare programs are the responsibility of separate state agencies. Even within similar funding streams, prekindergarten programs vary within and across states and territories in the U.S. Additionally, each state's department of education can set different requirements for teacher training, curricula, and definitions of good practices (Farran & Lipsey, 2016). Even within a given state, school districts may set different requirements in terms of formative and summative assessments, use of assessments, professional development provided for teachers, and which state-approved curricula to adopt. Differences between program requirements and expectations also depend on the presence of federal funding. The exception to this state control is the Head Start program, the closest to a federally controlled program for young children in the U.S.

In the U.S., there is increasing concern about moving beyond access and focusing on the quality of the prekindergarten programs (Hamre, 2014). The current evaluation systems used to assess the quality of U.S. prekindergarten classrooms tend to be global general ratings, to lack reliability and not to relate to children's development in either the short or long term (see Farran & Nesbitt, 2019, for a review). A behavioral count system provides a close examination of classroom behavioral interactions the results of which could become the basis for constructing a shared vision of quality in U.S. prekindergarten programs.

### 5.2. United States sample

The U.S. sample included 168 classrooms serving children whose families were considered low-income (at or below 200% of the poverty level in the U.S.). Classrooms were located in two urban areas in the southeastern region of the U.S. and were a part of the 2014 Preschool Development Grant-Enhancement (PDG-E) program funded by the U.S. Department of Education. The focus of the PDG-E initiative was to close the "opportunity gap" in achievement between children from low- and middle-/high-income families in the U.S. All classrooms were publicly-funded and housed in public schools. As mandated by the PDG-E, each classroom had at a minimum two teachers and up to 20 children (i.e., an adult to child ratio 1:10). In some classrooms, more than two teachers (e.g., special education teachers) were present for part of the day.

The average number of children in a classroom was 18.15 (*SD* = 1.79) and on average there were 1.24 (*SD* = 1.24) children per classroom identified with a developmental disability (data missing for 14 classrooms). Of the 168 study classrooms, 162 were homogeneous regarding age, serving children aged four years old by August 15. On average, the study's 2420 children were 55.36 months-old (*SD* = 3.65) and 49% male. Children were approximately 60% African American or Black, 21% Hispanic or Latinx, 12% White, 3% Asian, 3% Multiracial, and 1% other. The 168 lead teachers were on average 42.46 years-old (*SD* = 11.38, data missing for 22 teachers), had been teaching for 9.74 years (*SD* = 8.48, data missing for 45 teachers), and all but one was female. School districts did not provide comprehensive information about teacher ethnicity or educational attainment, but per the PDG-E funding, all lead teachers were required to have at a minimum a bachelor's degree and a teaching license. Information on educational assistants was also not provided by school districts

**Table 8**
Descriptive statistics of classroom schedule in Southeastern region of the United States.

| Schedule | *M* (*SD*) | Range |
|---|---|---|
| Whole-Group | .25 (.09) | .07–.49 |
| Small-Group | .02 (.06) | .00–.23 |
| Centers | .14 (.09) | .00–.49 |
| Small-Group/s & Centers[a] | .02 (.05) | .00–.24 |
| Transition | .34 (10) | .10–.68 |
| Meal-time | .16 (.04) | .07–.28 |
| Other[b] | .03 (.04) | .00–.23 |

*Note.* $N$ = 168 classrooms. Variables represent the proportion (range 0–1) of observational sweeps a given practice was observed. Coding was restricted to inside the school building, thus statistics do not capture the schedule occurring outdoors.

[a] Small Group/s & Centers refer to the simultaneous occurrence of Small Group/s and Centers.

[b] Other included gym, nap, and specials.

but all assistants were required to have a high school diploma or equivalent.

### 5.3. United States procedures

Each study classroom was observed one time between February and March 2017. All observers ($N$ = 18) had at least a bachelor's degree. Observations started at approximately 8 a.m. and lasted for an average of 6 h, with an average of 20.09 ($SD$ = 3.35) sweeps per child being collected. Observations did not occur during nap, meals or outside; thus most of the observational window was in the morning. All teachers and children present on the day of the observation were included in the study.

### 5.4. United States results and discussion

For the COP, the exact agreement across categories ranged between 77.7% (Level of Involvement) to 92.8% (Schedule) with an overall average of 85.9%, while Cohen's κ ranged between .66 (Level of Involvement) and .91 (Schedule) with an overall average of .80. For the TOP, the exact agreement across categories ranged between 77.3% (Task) to 92.3% (Schedule) with an overall average of 86.7%, while Cohen's κ ranged between .46 (Tone) and .99 (Schedule) with an overall average of .75. ICCs presented a value of .86 for Level of Involvement, .54 for Tone, and .87 for Level of Instruction. Estimates of interobserver reliability suggested that for the U.S. sample of low-income, urban, federally-funded classrooms, raters were able to adhere to operationalized definitions, and a behavioral count observation system can be implemented with an acceptable degree of reliability.

Examination of the descriptive statistics provides an informative picture of how urban prekindergarten classrooms housed in elementary schools organized their 6 -h day (see Table 8). On average, children were observed in transitions for approximately a third of the day (or 2 h), though with great variability among classrooms (ranging from 10% to 68%). Time in transitions does not include the time spent on naps and meals, which then cumulatively occupied more than 50% of the classroom day. The amount of time spent in transitions in these classrooms is comparable to that found by Early and colleagues in an evaluation of Georgia's Quality Rated Validation Project (Early et al., 2019). Transitions were created both by variables under the control of the teacher (e.g., transitions due to frequent and drawn-out periods of downtime in the classroom) as well as those dictated by the elementary school setting (e.g., meals and bathrooms outside of the home classroom).

Perhaps reflective of the goal of school readiness for these publicly funded classrooms, the predominant method of instruction was whole group (average = 25% of sweeps, range from 7% to 49%). Center time or free play, a defining characteristic of many preschool classrooms just a few years ago (Early et al., 2005), happened relatively seldom and did not last very long (average, 14%, range from 0% to 49%). Small group instruction was also a little utilized approach in these classrooms.

On average, teachers were observed providing instruction for approximately 24% of sweeps (Table 9), with the average level of instruction across the entire day equivalent to a focus on basic skills ($M$ = 1.80). In addition to instruction, teachers engaged in behaviors that contributed to the emotional climate of the classroom. Teachers were observed twice as often engaging in direct disapproval of children's behavior compared to approval of children's behavior, but again there was variability in teachers' use of these behaviors (e.g., the $SD$ for the variables was a large as the mean). Regarding teacher tone, on average it was neutral to positive in tone ($M$ = 3.26). These data reflect what has become perhaps an implicit emphasis in publicly funded prekindergarten classrooms with the emphasis on school readiness; defined as children developing basic skills and learning how to behave appropriately in a school environment.

Children had little opportunity to talk and spent little time interacting with each other. Averaged between the lead teacher and the assistant, teachers were observed talking for more than half of the sweeps (58%). In contrast, children were only observed talking in 23% of sweeps, and when they were talking, it was often not directed to the teacher. Consistent with this pattern, teachers were not observed listening to a specific child (6%) or a group of children (10%) very often. The percentage of sweeps that included an observation of a social learning interaction was quite low, with 5% of sweeps including associative interactions and 1% including cooperative interactions (i.e., interactions that are marked by a child engaging with another person in pursuit of a common goal). Little time was also spent in interacting with children in a way that promotes their active role in learning processes (i.e., posing a mix of open-ended and closed-ended questions and encouraging active participation); none in more extended inferential interactions (i.e., teachers consistently use open-ended questions that allow several rounds of turn taking and active participation), considered high-quality ways of instructing children in ECEC (e.g., Tompkins, Zucker, Justice, & Binici, 2013). Sequential activities (e.g., activities following a sequence of steps), which are more often associated with instruction, were observed twice as often as non-sequential activities. Sequential activities are the kinds of interactions that may well produce school readiness skills of the constrained variety, as Phillips et al. (2017) summarized, but they may not be linked to the long-term school achievement outcomes many anticipate in funding these programs.

In general, across all schedules, children's level of involvement was in the low to medium range on a 5-point scale ($M$ = 1.94), being about 1 point higher during schedules with clear learning opportunities (i.e., not observed in transition, mealtime, and other; $M$ = 2.80). Levels of involvement varied across schedules, with the highest level observed during centers (either co-occurring with small-groups or alone, $M$ = 2.94 and 2.93, respectively) and lowest during whole group instruction ($M$ = 2.75).

In terms of activities focused on math and literacy, children were more than twice as likely to be observed in literacy (15%) compared to math activities (6%). Substantial variability among classrooms was observed. Some classrooms were never observed providing children math activities while others were observed in math for 14% of sweeps. The similar distribution for literacy ranged from 4% to 32%.

To understand the generalizability of the observations in this study, future work must consider the variety of contexts in which young children are cared for or educated in the U.S. As the U.S. has no explicitly stated national goals, curriculum, or guiding principles for ECEC, classrooms providing care and education vary widely. The 168 classrooms we observed are representative of

**Table 9**
Descriptive statistics of classroom practices in Southeastern region of the United States.

| Variable | *n* | *M* (*SD*) | Range |
|---|---|---|---|
| Instruction provided by the teachers | | | |
| Instruction | 168 | .24 (.10) | .00–.50 |
| Level of instruction (while instructing)[a] | 167 | 1.80 (0.19) | 1.00–2.07 |
| Emotional climate in the classroom | | | |
| Behavior approving | 168 | .03 (.03) | .00–.23 |
| Behavior disapproving | 168 | .06 (.06) | .00–.42 |
| Teacher tone[b] | 168 | 3.26 (0.20) | 2.81 – 3.83 |
| Teachers and children talking and listening | | | |
| Children talking (overall) | 168 | .23 (.07) | .08–.50 |
| Children talking to teacher | 168 | .03 (.02) | .00–.08 |
| Teacher talking (overall) | 168 | .58 (.12) | .26–.93 |
| Teacher listening to a specific child | 168 | .06 (.06) | .00–.26 |
| Teacher listening to children (overall) | 168 | .10 (.08) | .00–.33 |
| Type of activity | | | |
| Non-sequential activities | 168 | .09 (.05) | .01–.27 |
| Sequential activities | 168 | .20 (.07) | .05–.39 |
| Social learning (associative and cooperative interactions) | | | |
| Associative interactions | 168 | .05 (.04) | .00–.19 |
| Cooperative interactions | 168 | .01 (.02) | .00–.12 |
| Levels of child involvement[c] | | | |
| Overall involvement | 168 | 1.94 (0.22) | 1.36 – 2.49 |
| Involvement in whole-group | 168 | 2.75 (0.29) | 1.65 – 3.96 |
| Involvement in small-group | 49 | 2.84 (0.63) | 1.00–5.00 |
| Involvement in centers | 160 | 2.93 (0.38) | 1.50 – 3.95 |
| Involvement in small-group/s & centers | 41 | 2.94 (0.35) | 2.22 – 3.59 |
| Involvement in learning opportunities | 168 | 2.80 (0.22) | 2.10 – 3.69 |
| Academic focus | | | |
| Math focus | 168 | .06 (.03) | .00–.14 |
| Literacy focus | 168 | .15 (.05) | .04–.32 |

*Note.* Sample size = 168 classrooms. Reported *n*s reflect the number of classrooms a given practice was observed. Unless otherwise indicated, variables represent the proportion (range 0–1) of observational sweeps a given practice was observed.
[a] Range: 0 (no instruction) to 4 (high inferential instruction).
[b] Range: 1 (extremely negative) to 5 (vibrant/enthusiastic).
[c] Range: 1 (low involvement) to 5 (high involvement).

U.S. public-school-based prekindergarten classrooms serving low-income urban minority children through state or federally funds. As in many states, but not all, teachers were required to be certified in early childhood education, meaning they had a bachelor's or master's degree, and they were paid on the same schedule as K-12 teachers. We believe the results of these observations could assist discussions as the U.S. ECEC field attempts to facilitate positive developmental outcomes for vulnerable children.

## 6. Overall discussion

This descriptive study used two primarily behavioral count measures for observing ECEC practices, focusing on both teacher and child behaviors, in three countries, Sweden, Portugal, and the U.S. The measure was developed in the U.S. Some adaptations to the tool were instituted in Sweden and Portugal, but variables summarized in this paper are the same within each country. This study had two main aims, namely to evaluate the cross-country relevance and inter-rater reliability of a measure of preschool practices primarily based on a behavior count approach, and to provide researchers within each country concrete descriptions of behaviors and interactions that could be evaluated against the ECEC approach and philosophy within the country.

In Sweden, the behavioral count descriptions indicated a rather high level of child agency and choice of activities (e.g., high amount of time in free-play), corresponding to national curriculum intentions. Still, the data suggest that the involvement in free play might not be achieving all that it could. Children were not observed to be highly involved in their free play activities. It is possible in large outdoor spaces for children to wander and for some to have difficulty finding an activity in which to be engaged. The curriculum stresses the need for children's voices to be heard, but preschool teachers seldom listened to children. The great freedom teachers have to implement the curriculum seemed most visible in the use of whole group activities, with large variation observed across preschool units. This variation could perhaps reflect the current controversy in Sweden concerning teacher instruction and how much instruction should occur. Still, the findings from some classrooms indicate that opportunities for academic learning can be integrated into child-managed contexts, in line with curriculum goals. Overall, the results illustrate the dilemmas and possibilities within a curriculum emphasizing child agency and play, while also attempting to move toward more teacher-led instruction.

In Portugal, using a behavior count measure to capture classroom practices brought to light several inconsistencies between the national guidelines and current practices in preschools. Preschool national guidelines value child rights and define infancy as a period where learning and development are interdependent, emphasizing children's involvement and agency. Regardless, preschools observed here seemed to be focused more on future formal schooling, with teacher-led activities, a large amount of time spent in whole group, and instruction focused on obtaining correct answers. Observing specific teacher and child behaviors clarified understandings of the preschool experience. This study adds to previous analyses of preschools, affirming that teachers' interpretation of Portuguese preschool guidelines needs to be examined and supported to fulfill the core principles of ECEC, with an increased focus on such aspects as considering the child as an active agent, fostering child involvement, and following child interests.

Data from the U.S. classrooms provide a picture of life within the classroom, updating the one from more than ten years ago (Phillips, Gormley, & Lowenstein, 2009). Phillips et al. (2009) provided an in-depth description of the classroom climate and instructional focus of prekindergarten classrooms in Tulsa, what they termed getting

inside the "black box" of prekindergarten instruction. We consider our more intensive descriptions as providing a current picture. The observations reveal a learning environment for young children in which teacher-led instruction is primary, with little time provided for child agency. There were numerous transitions as children in these classrooms moved frequently between activities, both academic and personal care routines. These frequent changes were associated with more "down" time, time spent in transitions, and perhaps with the need for more behavior control from the teacher, which was more often negative than positive. The low amount of time spent in small group activities likely affected the relatively low rating for instruction; having an inferential discussion with 20 children in a whole group setting would be quite difficult. The value of such data for the U.S. context is the opportunity for a more focused discussion about the purpose of prekindergarten, including how to define quality and the corresponding outcomes anticipated for children. While identifying quality of ECEC environments is of particular focus in the U.S., the concept is becoming more prevalent internationally. If related changes are recommended, these data provide actionable paths for teacher professional development.

Future studies examining teacher perspectives on instruction and how to focus on academic content (e.g., math and literacy) while maintaining a commitment to children's interests, could contribute to better understanding of our results, particularly in Portugal and Sweden. Using additional categories and combinations of categories from COP/TOP to capture a detailed picture of learning opportunities and peer interaction across the preschool day can also be informative for the three countries. The fact that COP/TOP allows category combinations gives researchers a wide range of possibilities to examine everyday interactions in preschools. For example, combining behavior counts can yield the extent of child unoccupied behaviors during centers or whole group (requiring a combination of the categories interaction state and schedule), or the extent of focus during centers (requiring focus codes and schedule). Another distinctive aspect of COP/TOP is the possibility to collect and analyze data at both the child and classroom levels. Future multilevel analyses can provide insights on ECEC practices.

### 6.1. Utility of COP/TOP cross-culturally

To examine the utility of the COP and TOP cross-culturally, we assess the data from the three countries, not in an effort to compare the *results* across the countries, rather to present issues that arose with the use of the measures. Although the COP/TOP are based on a behavioral count approach, three of the preschool practices analyzed involved a rating scale score: child level of involvement, teacher tone/affect, and teacher level of instruction (the latter is not, however, a Likert scale and was more reliably coded). Overall, the behavior count variables achieved high exact inter-observer agreement in the three countries in this study, suggesting the COP/TOP may be promising resources for observing ECEC practices, and addressing some concerns about reliability of rating scales as previously highlighted (Mashburn, 2017).

When the COP measure was created in the U.S., the acknowledged focus for children's level of involvement was within "learning activities," reflecting the U.S. perspective on the goals for early childhood education. It is perhaps easier to accurately observe children's level of involvement in organized and didactic activities. Since centers are more common in Swedish preschool settings than in the Portuguese and U.S. classrooms, a scale measuring "involvement in learning" may be particularly problematic to use there. Such differences reflect a possible construct bias in conducting cross-cultural research as cautioned by Boer et al. (2018). Similarly, ensuring that all children have access to quality ECEC and are included (United Nations, 2015) requires adaptations in coding classroom practices in ECEC. The adaptations to code behaviors of children in need of special support in the Portuguese and Swedish classrooms may also be useful in the U.S., extending the application of the measures to inclusive settings.

A similar potential bias emerges with the teacher tone rating. Cultures have dissimilar patterns of expressing emotion (e.g., Li & Karakowsky, 2001), and we found relatively low reliability across the countries in the rating scale of teacher tone. However, in addition to possible cultural variations in emotional tone, one serious problem with achieving reliability for a rating of tone is the lack of variability among teachers and sweeps. For most of the observations in both Sweden and the U.S. observers coded the teacher as "flat" or neutral. Rarer more positive or negative expressions contribute to the low reliability among observers to catch these fleeting instances and agree.

The lower reliabilities for ratings of child level of involvement and teacher tone (both scored trough ratings scales) in the Swedish context were discussed extensively among the authors. One particular difficulty of coding these variables reliably in Sweden was the high amount of time spent outdoors. When children were outdoors in Swedish preschools, they typically performed gross motor activities (e.g., bicycling, running, or sledding). They also engaged in dramatic play, often in relation to a shed, playhouse, or sheltered area. Additionally, the outdoor environment could be natural, including trees and bushes, and in the time of Swedish data collection (fall/winter), children wore overalls and caps. These outdoor conditions likely complicate observations in general, but specifically ratings of child level of involvement and teacher tone, because of the difficulty of seeing faces accurately which may lead to low inter-rater observations.

Overall, behavior count measures provide observers with specific and concrete definitions of what to observe and code, allowing observers to achieve high exact inter-rater agreement on most variables. With the exception of the two Likert rating scale variables, the measures used in the present study appeared to be suitable for use in different contexts, across preschool settings organized differently by schedule and environment.

### 6.2. Limitations

Some limitations of this study must be acknowledged. First, a random selection of preschools in each country was not possible. This means that we cannot claim that the samples from each country are representative of classroom practices in the participating countries, as there are many possible sources of variation in practices within countries, such as funding, group structure and number of children, teacher's education, and differences in consent rates for participating. Thus, interpretation of the data must be carefully contextualized.

Another related limitation is that, although COP and TOP were designed to observe all elements in each setting, only in the U.S. were all children in each classroom observed. In Portugal and Swedish classrooms, observations were limited to children whose parents gave consent. As a result, a smaller number of children in each setting were observed and may have reduced the effectiveness of the measure in capturing the experiences of all children. However, there are many published studies that use fewer children per classroom, some as few as four to six (e.g., Bratsch-Hines et al., 2019; Early et al., 2005).

The effect of raters is another potential limitation. Portuguese and Swedish data collectors received initial training together (in Portugal), with the opportunity to observe practices jointly. During data collection, however, observations were conducted only by local observers (in each country), and reliability across coders from different countries was not possible to check due to a lack of funding. Furthermore, the U.S. observers did not observe in Swedish

or Portuguese classrooms. Unless observers are multi-lingual, this may always be a potential issue; higher inter-rater reliability may be obtained within a country than might be achieved across countries (Heine, Lehman, Peng, & Greenholtz, 2002).

## 7. Conclusion

This study, involving multiple classrooms in three countries and used similar observation measures, from which the same variables were extracted for analyses for the following aims: to examine the within-country relevance of two classroom observation measures primarily based on a behavioral count approach, and to examine preschool practices in Sweden, Portugal, and the U.S., as they reflected each country's ECEC goals, organization, and educational philosophies.

We demonstrated that the measures targeted culturally relevant behaviors and provided inter-rater reliability for the behavior count variables in the three countries. We also note that some culturally specific variables needed to be added. The behavioral descriptions reflected each of the country's values, as well as current debates regarding education and care for young children within each.

In terms of further international work, we argue that measures such as these can provide comprehensive descriptions of classroom settings and apply minimal external or comparative value judgments on the behaviors observed. While these data are expensive to collect, the current analysis and possible future exploration of the data collected here could provide the basis for actionable steps as desired by ECEC experts.

## Author note

This paper is financed by FEDER funds through the Operational Competitiveness Program- COMPETE and by Portuguese national funds through FCT - Portuguese Foundation for Science and Technologyunder the reference SFRH/BD/111211/2015. The support of the Metro Nashville Public School System, contract # 2-218740-19 is gratefully acknowledged, as well as the support from The Swedish Research Council, contract #2016-05555, and FORTE, Sweden.

## APPENDIX A. : COP/TOP category definitions

| Category | Coding options | Definitions |
|---|---|---|
| | Whole Group | Whole groups occur when the entire group is meeting together, and some form of content is being discussed. |
| **COP/TOP** *Schedule* | Small Group | Children are working in small groups that are facilitated by a teacher and/or are assigned by a teacher, a teacher assigns specific children to participate, the activity is not optional. Does not include center time. |
| | Centers | Center time (sometimes called "Activity Time" or "Free Play") is characterized by children having relatively more freedom about what they do and where they go, although teachers may place limitations on children's choices. |
| | Small-Group Centers | Small group(s) and center time are simultaneously occurring in the classroom. |
| | Transitions | Transitions begin when an activity has clearly ended for the large majority (>75%) of the children (e.g., storytime) and the next activity has not yet begun. Includes lining up for bathroom or meals, walking in the hallway. |
| | Meal-Time | Mealtime is coded when children are scheduled to eat breakfast, lunch, or snacks. |
| | Passive Instruction | The child is merely receiving instruction, rather than being an active participant. |
| | Non-sequential | The child is involved with an activity or with materials but not following a predetermined set of steps. For instance, the child is doodling on paper, browsing through the books in the classroom library, pushing toy trucks around on the rug. |
| **COP** *Type of Task* | Sequential | The child is involved in activities or with materials that involve a sequence of steps. For instance, the child is reading (either actually reading or looking carefully at the pictures), conducting a science experiment, working on a puzzle. |
| | Fantasy/Drama | The child is engaged in a sequenced and predictable make believe and pretend play enacting familiar stories, plays or role activities. |
| | Disruptive | The child exhibits any behavior that draws other people off task, including aggressive behavior (physical or verbal), inappropriate use of materials, and inappropriate actions toward other children. |
| | None | The child is not directly engaged in an activity or with materials and is not engaged in social talk. |
| | Other | The child is engaged in expected classroom tasks with no learning focus. For instance, the child is waiting for the teacher to check materials, washing hands. |

| | *Coding options* | *Definitions* |
|---|---|---|
| | Time Out | The child is isolated from the group in variations of time out defined by the teacher as a punishment. |
| | Unoccupied | The child is not attending to any particular learning-related activity that is occurring. |
| | Alone | The child is working alone in an activity that is unique from the activities of the others in the classroom (i.e., not in parallel interaction). |
| **COP** *Interaction State*[1] | Non-academic | The child is in a non-learning-related activity about which has little choice. For instance, the child is transitioning from one activity to another, the child is participating in appropriate required classroom routine, the child is forced to wait and has nothing to do. |
| | Onlooker | The child is observing other children or a teacher doing a learning activity that is not part of the target child activity. |
| | Parallel | The child is working independently, not interacting with other children, but is using materials that are similar to those other children are using. For instance, the child is working independently on an activity that others in the classroom also are working on. |
| | Social | The child is interacting with one or more children and/or a teacher, but it does not appear to be a learning or pretend play topic in the interaction. |
| | Associative | The child is interacting in the context of an activity or task that does not have predetermined rules. For instance, the child is building something with blocks together with other children or with teachers, the child is sharing a book with other children or with teachers. |
| | Cooperative | The child is interacting with other children or with teacher, sharing common goals, rules, and organization. For instance, the child is participating in formal games, competitions aimed at winning something. |
| **COP** *Involvement* | Level of (overall) Involvement[2] | Level of overall involvement captures how focused and engaged the child is in whatever activity he or she is doing. This is coded in a 5- point rating scale, where 1 means low involvement (e.g., totally out of task, not paying attention to the activity, sitting quietly; fiddling with another child's hair or clothing, eyes not focused on ongoing activity), 2 means medium-low involvement (e.g., looking at teacher and/or material inconsistently, flat affect, looking bored, visible attention going in and out, visible lack of persistence), 3 means medium involvement (e.g., on task, maintaining eye contact with teacher, participating but may briefly look around, although immediately comes back to task), 4 means medium-high involvement (e.g., eager expression, relevant self-talk during tasks, volunteering response with positive affect, looking at material throughout the entire time; leaning forward, showing persistence), and 5 means high involvement (e.g., intense focus, serious persistence and pursuit of activity, very difficult to be distracted from the activity, seeming oblivious to noise and the behaviors of the other children that are not related to the task). |
| **COP/TOP** *Focus*[3] | Literacy | Literacy is coded when the learning focus is literacy related (e.g. invented writing or tracing, dictating or writing text with meaning, name writing or recognizing the names of other children in the class, conversations and communication for the purpose of language development). |
| | Math | Math is coded when the focus is related to math concepts (e.g., number, comparing number, operations, shape, comparing shapes, composing shapes, spatial reasoning, measurement, patterning, classification). |
| | Instruction | Instruction involves any learning activity during which the teacher is interacting with a child or children. It includes activities that are typically considered academic in content (e.g., math or literacy), as well as activities in which the focus is art, music, puzzles, or blocks. Instruction can occur with or without materials. |
| **TOP** *Type of Task* | Assessment | The teacher is administering an assessment or test, taking anecdotal notes about children, but is not directly interacting with children. |
| | Administrative | The teacher is engaged in an activity that is required by the school, for instance paperwork, speaking to a parent, talking on the phone for a school-related reason. |
| | Managerial | The teacher is actively engaged in an activity that is required to run a classroom. For instance, the teacher is lining children up, organizing children to move from one activity to another, passing out materials. |
| | Monitoring | The teacher is passively observing children, scanning the room to determine what children are doing. |
| | Behavior Approving | The teacher uses approving verbal comments, facial expressions, or a physical contact with children. |
| | Behavior Disapproving | The teacher uses disapproving facial expressions, verbal comments, tone of voice, and/or physical contact with children. |
| | Personal/Care | The teacher performs personal care tasks for children. For instance, the teacher/assistant is supervising children or helping to tie shoes, brush teeth, fix clothes, wash hands. |
| | Social | Social is coded when personal conversations or physical contact that has no learning content is occurring. |
| **TOP** *Level of Instruction* | Level of Instruction | Level of instruction describes the instruction that is occurring at the moment. The levels are coded from 0 (no instruction) to 4 (high inferential instruction). Low level of instruction (1) occurs when the observer cannot recognize the intent to teach a specific academic skill (e.g., fine- or gross motor skills, songs without learning content). Basic skills instruction (2) is coded when the focus is on learning a specific skill (e.g. counting, recognizing letters), the correct answer is predetermined, and the goal is to that children learn the correct response/answer. Level 3 - some inferential learning – is coded when teacher interacts with children using a mix of closed-ended and open-ended questions thus allowing children to be an active agent and participate actively in the learning process. Level 4 is coded when the teacher is involved in inferential interactions, i.e., the teacher is interacting with children using open-ended questions, consistently, in which the answer is not predetermined, allowing children to participate, takes several turns, and sharing information. |
| **TOP** *Tone* | Tone | Teacher's tone captures the tone/affect of the teacher, aiming to reflect the positive or negative feel of the classroom and the interaction of the teacher. This is coded in a rating scale ranging from 1 (extremely negative affect) to 5 (vibrant and enthusiastic positive affect). |

*Note.* Adapted from Farran and Son-Yarbrough (2001). Title I funded preschools as a developmental context for children's play and verbal behaviors. *Early Childhood Research Quarterly, 16*, 245–262. https://doi.org/10.1016/S0885-2006(01)00100-4; See manual for extended definitions and examples, available from D.C. Farran, Vanderbilt University. [1] Associative and cooperative interactions refer to social learning opportunities. [2] Besides overall involvement, the combination of the involvement category from COP with the schedule category) allows documentation of involvement by type of schedule, namely: involvement in whole group, involvement in small group, involvement in centers, involvement in small group/s & centers; and involvement in learning opportunities; Learning opportunities occur when a target child has the opportunity to be involved in a learning activity, which excludes Sweeps coded as Transition, Meal Time, and Nap from the schedule category. [3]Other codes for the category can be found in the COP manual.

## References


Abreu-Lima, I., Pinto, A. I., Grande, C., Cadima, J., Mata, L., Pimentel, J., et al. (2014). *Caracterização dos contextos de educação pré-escolar. Inquérito extensivo. Relatório final. [Study of the assessment of the Curriculum Guidelines and quality in preschool: final report of the case studies]*. Lisbon, Portugal: Ministry of Education.

Abreu-Lima, I., Leal, T. B., Cadima, J., & Gamelas, A. M. (2013). Predicting child outcomes from preschool quality in Portugal. *European Journal of Psychology of Education*, *28*(2), 399–420. http://dx.doi.org/10.1007/s10212-012-0120-y

Aguiar, C., Moiteiro, A. R., & Pimentel, J. (2010). Classroom quality and social acceptance of preschoolers with disabilities. *Infants and Young Children*, *23*(1), 34–41. http://dx.doi.org/10.1097/IYC.0b013e3181c9766e

Bertram, T., & Pascal, C. (2016). *Early childhood policies and system in eight countries: Findings from IEA's early childhood education study*. Amsterdam: The International Association for the Evaluation of Educational Achievement.

Bilbrey, C., Vorhaus, E., & Farran, D. (2007). *Teacher observation in preschool* Unpublished. Nashville, TN: Peabody Research Institute, Vanderbilt University.

Boer, D., Hanke, K., & He, J. (2018). On detecting systematic measurement error in cross-cultural research: A review and critical reflection on equivalence and invariance tests. *Journal of Cross-Cultural Psychology*, *49*, 713–734.

Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2015). *Too many children left behind. The U.S. Achievement gap in comparative perspective*. New York: The Russell Sage Foundation. ISBN-13: 978-0871540249.

Bratsch-Hines, M., Burchinal, M., Peisner-Fineberg, E., & Franco, X. (2019). Frequency of instructional practices in rural prekindergarten classrooms and associations with child language and literacy skills. *Early Childhood Research Quarterly*, *47*, 74–88. http://dx.doi.org/10.1016/j.ecresq.2018.10.001

Briesch, A., Chafouleas, S., & Riley-Tillman, T. C. (2010). Generalizability and dependability of behavior assessment methods to estimate academic engagement: A comparison of systematic direct observation and direct behavior rating. *School Psychology Review*, *39*, 408–421.

Broekhulzen, M., Leseman, P., Moser, T., & van Trijp, K. (2015). *Stakeholders' study. Values, beliefs and concerns of parents, and policy representatives regarding ECEC services in nine European countries* Retrieved from. CARE: Curriculum & Quality Analysis and Impact Review of European Early Childhood Education and Care, Utrecht University. http://ecec-care.org/resources/publications/

Brunsek, A., Perlman, M., Falenchuk, O., McMullen, E., Fletcher, B., & Shah, P. S. (2017). The relationship between the Early Childhood Environment Rating Scale and its revised form and child outcomes: A systematic review and meta-analysis. *PloS One*, *12*(6) http://dx.doi.org/10.1371/journal.pone.0178512

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between childcare quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, *25*, 166–176. http://dx.doi.org/10.1016/j.ecresq.2009.10.004

Cadima, J., Aguiar, C., & Barata, M. C. (2018). Process quality in Portuguese preschool classrooms serving children at-risk of poverty and social exclusion and children with disabilities. *Early Childhood Research Quarterly*, *45*, 93–105. http://dx.doi.org/10.1016/j.ecresq.2018.06.007

Campbell-Barr, V., & Bogatić, K. (2017). Global to local perspectives of early childhood education and care. *Early Child Development and Care*, *187*, 1461–1470. http://dx.doi.org/10.1080/03004430.2017.1342436

Cash, A. H., Hamre, B. K., Pianta, R. C., & Myers, S. S. (2012). Rater calibration when observational assessment occurs at large scale: Degree of calibration and characteristics of raters associated with calibration. *Early Childhood Research Quarterly*, *27*(3), 529–542. http://dx.doi.org/10.1016/j.ecresq.2011.12.006

Coelho, V., & Pinto, A. I. (2018). The relationship between children's developmental functioning and participation in social activities in Portuguese inclusive preschool settings. *Frontiers in Education: Special Education Needs*, http://dx.doi.org/10.3389/feduc.2018.00016

Cohen, B., & Korintus, M. (2017). Making connections: Reflections on over three decades of EU initiatives in Early Childhood Education and Care (ECEC). *Early Years*, *37*, 235–249. http://dx.doi.org/10.1080/09575146.2016.1181050

Decree-Law 281/2009. (2009). *Decreto-Lei n.º 281/20009 de 6 de outubro [Decree-Law No.281/2009 of October 6th]* Diário da República, n.º 193/2009. Portugal: Série I. Ministério da Saúde.

Decree-Law 54/2018. (2018). *Decreto-Lei n.º 54/2018 de 6 de julho [Decree-Law 54/2018 of July 6th]*. Portugal: Diário da República, Série I. Ministério da Educação.

Dowsett, C. J., Huston, A. C., Imes, A. E., & Gennetian, L. (2008). Structural and process features in three types of childcare for children from high and low income families. *Early Childhood Research Quarterly*, *23*(1), 69–93. http://dx.doi.org/10.1016/j.ecresq.2007.06.003

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. http://dx.doi.org/10.1037/0012-1649.43.6.1428

Early, D., Barbarin, O., Bryant, D., Burchinal, M., Chang, F., Clifford, R., et al. (2005). *Pre-kindergarten in eleven states: NCEDL's multi-state study of pre-kindergarten and study of state-wide early education programs (SWEEP)*. Chapel Hill, NC: National Center for Early Development and Learning. https://fpg.unc.edu/node/4654

Early, D., Maxwell, K., Blasberg, A., Miranda, B., Orfali, N., Li, W., et al. (2019). *Quality rated validation study report #4: Quality rated star ratings and independent measures of quality, children's growth, and work climate*. Bethesda, MD: Child Trends. https://www.childtrends.org/publications/quality-rated-validation-study-report-4

Einarsdottir, J., Purola, A.-M., Johansson, E. M., Broström, S., & Emilson, A. (2015). Democracy, caring and competence: Values perspectives in ECEC curricula in the Nordic countries. *International Journal of Early Years Education*, *23*(1), 97–114. http://dx.doi.org/10.1080/09669760.2014.970521

European Agency for Special Needs and Inclusive Education. (2017). *Inclusive early childhood education: Literature review* Retrieved from:. https://www.european-agency.org/resources/publications/inclusive-early-childhood-education-literature-review

European Commission. (2015). *Early childhood education and care systems in Europe, National information sheets 2014/15* Retrieved from. https://eacea.ec.europa.eu/national-policies/eurydice/content/early-childhood-education-and-care-systems-europe-national-information-sheets-%E2%80%93-201415_en

European Commission/EACEA/Eurydice. (2019). *Key data on early childhood education and care in Europe – 2019 edition* Retrieved from. Eurydice Report. Luxembourg: Publications Office of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CONSIL:ST_9014_2019_INIT&from=EN

European Council. (2019). *Council recommendation on high-quality early childhood education and care systems*. Brussels: Council of the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CONSIL:ST_9014_2019_INIT&from=EN

Farran, D. C., & Lipsey, M. W. (2016). Evidence for the benefits of state prekindergarten programs: Myth & misrepresentation. *Behavioral Science & Policy*, *2*(1), 9–18.

Farran, D., & Nesbitt, K. (2019). New information on evaluating the quality of early childhood education programs. In O. Saracho, & B. Spodek (Eds.), *Handbook of research on the education of Young children (2nd edition)*. New York, NY: Routledge/Taylor & Francis.

Farran, D. C., & Son-Yarbrough, W. (2001). Title I funded preschools as a developmental context for children's play and verbal behaviors. *Early Childhood Research Quarterly*, *16*, 245–262. http://dx.doi.org/10.1016/S0885-2006(01)00100-4

Farran, D. C., Meador, D. N., Christopher, C. H., Nesbitt, K. T., & Bilbrey, L. E. (2017). Data-driven improvement in prekindergarten classrooms: Report from a partnership in an urban district. *Child Development*, *88*, 1466–1479. http://dx.doi.org/10.1111/cdev.12906

Fischer, R., & Poortinga, Y. (2018). Addressing methodological challenges in culture-comparative research. *Journal of Cross-Cultural Psychology*, *49*, 691–712. http://dx.doi.org/10.1177/0022022117738086

Friedman-Krauss, A. H., Barnett, W. S., Weisenfeld, G. G., Kasmin, R., DiCrecchio, N., & Horowitz, M. (2018). *The state of preschool 2017: State preschool yearbook*. The National Institute for Early Education Research, Rutgers University. http://nieer.org/state-preschool-yearbooks/yearbook2017

Fuhs, M. W., Farran, D. C., & Nesbitt, K. T. (2013). Preschool classroom processes as predictors of children's cognitive self-regulation skills development. *School Psychology Quarterly*, *28*, 347–359. http://dx.doi.org/10.1037/spq0000031

Garvis, S., Sheridan, S., Williams, P., & Mellgren, E. (2017). Cultural considerations of ECERS-3 in Sweden: A reflection on adaptation. *Early Child Development and Care*, *188*(5), 584–593. http://dx.doi.org/10.1080/03004430.2017.1377192

Goble, P., Hanish, L. D., Martin, C. L., Eggum-Wilkens, N. D., Foster, S. A., & Fabes, R. A. (2016). Preschool contexts and teacher interactions: Relations with school readiness. *Early Education and Development*, *27*(5), 623–641. http://dx.doi.org/10.1080/10409289.2016.1111674

Gordon, R. A., Fujimoto, K., Kaestner, R., Korenman, S., & Abner, K. (2013). An assessment of the validity of the ECERS–R with implications for measures of childcare quality and relations to child development. *Developmental Psychology*, *49*, 146–160. http://dx.doi.org/10.1037/a0027899

Grammatikopoulos, V., Gregoriadis, A., & Zachopoulou, E. (2015). Evaluation of early childhood education environments and professional development: Current practices and implications. In O. Saracho (Ed.), *Contemporary perspectives on research in assessment and evaluation in early childhood education* (pp. 153–172). Charlotte, NC: Information Age Publishing.

Hamre, B. (2014). Teachers' daily interactions with children: An essential ingredient in effective early childhood programs. *Child Development Perspectives*, *8*, 223–230. https://doi.org/10.1111/cdep.12090

Harms, T., & Clifford, R. (1980). *The early childhood environment rating scale*. New York: Teachers College Press.

Harms, T., Clifford, R., & Cryer, D. (2004). *Early childhood environment rating scale-revised edition*. New York: Teachers College Press. ISBN-13: 978-0807745496.

Heine, S., Lehman, D., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective scales?: The reference-group effect. *Journal of Personality and Social Psychology*, *82*, 903–918. http://dx.doi.org/10.1037//0022-3514.82.6.903

Hoyt, W., & Kerns, M.-D. (1999). Magnitude and moderators of bias in observer ratings: A meta-analysis. *Psychological Methods*, *4*, 403–424. http://dx.doi.org/10.1037/1082-989X.4.4.403

Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology*, *16*(2), 131–152.

Institute of Education Sciences, National Center for Educational Statistics. (2019). *The condition of education: Preschool and kindergarten enrollment*. https://nces.ed.gov/programs/coe/indicator_cfa.asp

Jacobson, L. (2019). Pre-to-3: New' baby PISA' study to include US 5-year-olds. *Education Dive*, https://www.educationdive.com/news/pre-to-3-new-baby-pisa-study-to-include-us-5-year-olds/549810/

Jonsson, A., Williams, P., & Pramling Samuelsson, I. (2017). Undervisningsbegreppet och dess innebörder uttryckta av förskolans lärare. [The concept of teaching and its meaning as expressed by preschool teachers]. *Forskning om undervisning och lärande*, *5*(1), 90–109.
Kalpogianni, D. E. (2019). Why are the children not outdoors? Factors supporting and hindering outdoor play in Greek public day-care centres. *International Journal of Play*, *8*(2), 155–173. http://dx.doi.org/10.1080/21594937.2019.1643979
Keys, T., Farkas, G., Burchinal, M., Duncan, G., Vandell, D., Li, W., et al. (2013). Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child Development*, *84*, 1171–1190. http://dx.doi.org/10.1111/cdev.12048
Kline, M. A., Shamsudheen, R., & Broesch, T. (2018). Variation is the universal: Making cultural evolution work in developmental psychology. *Philosophical Transactions Biological Sciences*, *373*, 1–8. http://dx.doi.org/10.1098/rstb.2017.0059, 20170059
Li, J., & Karakowsky, L. (2001). Do we see eye-to-eye? Implications of cultural differences for cross-cultural management research and practice. *The Journal of Psychology*, *135*, 501–517.
Luttropp, A., & Granlund, M. (2010). Interaction - it depends - a comparative study of interaction in preschools between children with intellectual disability and children with typical development. *Scandinavian Journal of Disability Research*, *12*(3), 151–164. http://dx.doi.org/10.1080/15017410903175677
Mashburn, A. J. (2017). Evaluating the validity of classroom observations in the Head Start designation renewal system. *Educational Psychologist*, *52*(1), 38–49. http://dx.doi.org/10.1080/00461520.2016.1207539
Mashburn, A., Meyer, J., Allen, J., & Pianta, R. (2014). The effect of observation length and presentation order on the reliability and validity of classroom observations. *Educational and Psychological Measurement*, *74*, 400–422. http://dx.doi.org/10.1177/0013164413515882
Melhuish, E., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K., et al. (2015). *A review of research on the effects of early childhood Education and Care (ECEC) upon child development*. CARE project. http://ecec-care.org/fileadmin/careproject/Publications/reports/CARE_WP4_D4_1_review_of_effects_of_ecec.pdf
Moiduddin, E., Bush, C., Manley, M., Aikens, N., Tarullo, L., Malone, L., et al. (2017). *A Portrait of Head Start classrooms and programs in Spring 2015: FACES 2014-2015 data tables and study design* OPRE Report 2017-101. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
Moser, T., Leseman, P., Melhuish, E., Borekhuizen, M., & Slot, P. (2017). *European framework of quality and wellbeing indicators. (D6.3)*. http://ecec-care.org/fileadmin/careproject/Publications/reports/D6_3_CARE_Framework_of_Quality_and_Wellbeing_Indicators.pdf
Moss, P., Dahlberg, G., Grieshaber, S., Mantovani, S., May, H., Pence, A., et al. (2016). The organisation for economic Co-operation and development's international early learning study: Opening for debate and contestation. *Contemporary Issues in Early Childhood*, *17*(3), 343–351. http://dx.doi.org/10.1177/1463949116661126
OECD. (2015). *Education at a Glance 2015: OECD indicators*. Paris: OECD publishing. http://dx.doi.org/10.1787/eag-2015-en
OECD. (2017). *Starting Strong 2017: Key OECD indicators on early childhood education and care*. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264276116-en
OECD. (2020). *Early learning and child well-being: A study of five-year-Old's in England, Estonia, and the United States*. Paris: OECD Publishing. http://dx.doi.org/10.1787/3990407f-en
Parker, E., Workman, E., & Atchison, B. (2016). *50 state review. States pre-k funding for 2015–2016 fiscal year: National trends in state preschool funding*. Denver, CO: Education Commission of the States.
Pastori, G., & Pagani, V. (2017). Is validation always valid? Cross-cultural complexities of standard-based instruments migrating out of their context. *European Early Childhood Education Research Journal*, *25*(5), 682–697. http://dx.doi.org/10.1080/1350293X.2017.1356545
Perlman, M., Falenchuk, O., Fletcher, B., McMullen, E., Beyene, J., & Shah, P. (2016). A systematic review and meta-analysis of a measure of Staff/Child interaction quality (the classroom assessment scoring system) in early childhood education and care settings and child outcomes. *PloS One*, *11*(12), Article e0167660 http://dx.doi.org/10.1371/journal.pone.0167660
Phillips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., et al. (2017). *Puzzling it out: The current state of scientific knowledge on pre-kindergarten effects. A consensus statement*. Washington, DC: The Brookings Institution. https://www.brookings.edu/wp-content/uploads/2017/04/consensus-statement_final.pdf
Phillips, D., Gormley, W., & Lowenstein, A. (2009). Inside the prekindergarten door: Classroom climate and instructional time allocation in Tulsa's pre-k programs. *Early Childhood Research Quarterly*, *24*, 213–238. http://dx.doi.org/10.1016/j.ecresq.2009.05.002
Pianta, R. C., La Paro, K., & Hamre, B. (2008). *Classroom assessment scoring system (CLASS) manual, Pre-k*. Baltimore, MD: Paul Brookes.
Pramling Samuelsson, I. (2010). Ska barns kunskaper testas eller deras kunnande utvecklas i förskolan? [Are children's knowledge to be tested or developed in preschool?]. *Nordic Early Childhood Education Research*, *3*(3), 159–167.
Pramling Samuelsson, I., & Sheridan, S. (2009). Preschool quality and Young Children's Learning in Sweden. *International Journal of Child Care and Education Policy*, *3*(1), 1–11. http://dx.doi.org/10.1007/2288-6729-3-1-1
Pramling Samuelsson, I., Williams, P., Sheridan, S., & Hellman, A. (2016). Swedish preschool teachers' ideas of the ideal preschool group. *Journal of Early Childhood Research*, *14*(4), 444–460. http://dx.doi.org/10.1177/1476718x14559233
Rentzou, K. (2017). Using rating scales to evaluate quality early childhood education and care: Reliability issues. *European Early Childhood Education Research Journal*, *25*(5), 667–681. http://dx.doi.org/10.1080/1350293X.2017.1356599
Santos, A. I., & Alves Martins, M. (2011). Literacy practices in Portuguese kindergartens and children's conceptualisations about written language. *Revue Internationale des Sciences de l'Éducation*, *25*, 125–141. http://dx.doi.org/10.4000/dse.1043
Scarr, S., & Weinberg, R. (1986). The early childhood enterprise: Care and education of the young. *The American Psychologist*, *41*, 1140–1146.
Sheridan, S., & Williams, P. (2018). *Undervisning i förskolan: En kunskapsöversikt [Teaching in preschool: A review]*. https://www.skolverket.se/publikationsserier/kunskapsoversikter/2018/undervisning-i-forskolan-en-kunskapsoversikt
Sheridan, S., Giota, J., Han, Y. M., & Kwon, J. Y. (2009). A cross-cultural study of preschool quality in South Korea and Sweden: ECERS evaluations. *Early Childhood Research Quarterly*, *24*(2), 142–156. http://dx.doi.org/10.1016/j.ecresq.2009.03.004
Silva, I. L., Marques, L., Mata, L., & Rosa, M. (2016). *Orientações Curriculares para a Educação Pré-Escolar [Curriculum Guidelines for Preschool Education]*. Lisbon, Portugal: Ministry of Education and General Direction of Education.
Spivak, A. L., & Farran, D. C. (2016). Predicting first graders' social competence from their preschool classroom interpersonal context. *Early Education and Development*, *27*, 735–750. http://dx.doi.org/10.1080/10409289.2016.1138825
Storli, R., & Hansen Sandseter, E. B. (2019). Children's play, well-being and involvement: How children play indoors and outdoors in Norwegian early childhood education and care institutions. *International Journal of Play*, *8*(1), 65–78. http://dx.doi.org/10.1080/21594937.2019.1580338
Swedish Education Act (SFS 2010:800). https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/skollag-2010800_sfs-2010-800.
Swedish National Agency for Education. (2019a). *Barn och personal i förskolan per 15 oktober 2018 [Children and staff in preschool per 15 oktober 2018]*.
Swedish National Agency for Education. (2019b). . *Curriculum for the preschool* (18) Lpfö.
Swedish Schools Inspectorate. (2018). ). *Förskolans kvalitet och måluppfyllelse: Ett treårigt regeringsuppdrag att granska förskolan* Slutrapport [Quality and goal attainment of the preschool: A three-year government mandate to review the preschool. Final report].
Tobin, J. (2005). Quality in early childhood education: An anthropologist's perspective. *Early Education and Development*, *16*, 421–434. http://dx.doi.org/10.1207/s15566935eed1604_3
Tobin, J., Wu, D., & Davidson, D. (1989). *Preschool in three cultures*. New Haven: Yale University Press.
Tompkins, V., Zucker, T. A., Justice, L. M., & Binici, S. (2013). Inferential talk during teacher–child interactions in small-group play. *Early Childhood Research Quarterly*, *28*(2), 424–436. http://dx.doi.org/10.1016/j.ecresq.2012.11.001
Ulferts, H., Anders, Y., Leseman, P., & Melhuish, E. (2016). *D4.2:Effectsof ECEC on academic outcomes in literacy and mathematics: Meta-analysis of European longitudinal studies*. Utrecht, Netherlands: Utrecht University.
UNICEF Office of Research. (2013). *Child well-being in rich countries: A comparative overview, Innocenti Report Card 11*. Florence, Italy: UNICEF Office of Research.
United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*. https://sustainabledevelopment.un.org/post2015/transformingourworld
Vallberg Roth, A. C. (2020). What may characterize teaching in preschool? The written descriptions of Swedish preschool teachers and managers in 2016. *Scandinavian Journal of Educational Research*, *64*(1), 1–21. http://dx.doi.org/10.1080/00313831.2018.1479301
Vermeer, H., van IJzendoorn, M., Cárcamo, R., & Harrison, L. (2016). Quality of childcare using the Environment Rating Scales: A meta-analysis of international studies. *International Journal of Early Childhood*, *48*(1), 33–60. http://dx.doi.org/10.1007/s13158-015-0154-9
Westberg, J. (2019). Designing preschools for an independent and social child: Visions of preschool space in the Swedish welfare state. *Early Years*, 1–18. http://dx.doi.org/10.1080/09575146.2019.1608426
Wolery, M., Paucca, T., Brashers, M. S., & Grant, S. (2000). ). *Quality of inclusive experiences measure*. Chapel Hill, NC: University of North Carolina, FPG Child Development Center.
World Bank. (2013). *What matters most for early childhood development: A framework paper* Working Paper Series No 5, January 2013. Washington, DC: World Bank.