University of Milano Bicocca

Department of Human Sciences 'R. Massa'

Teacher-child relationship quality in ECEC settings.

A mixed-methods study of the CLASS tool application in Italy.

Author:

Valentina PAGANI

Supervisor:

Dr. Giulia PASTORI

Academic year 2015/2016

«Education is either a situation of research,
and the research produces a new pedagogy […].
Pedagogy is movement, continuous movement…
I don't believe that pedagogy can know, each day,
where it is going and where it may go;
it is a route that you discover as you travel…»
Loris Malaguzzi

For those who walked with me on this journey
and for those who, step by step, keep on traveling
to discover new paths.

Contents

Abstract

In recent years, a rising body of research has shown that, although early education matters, only high quality ECEC can make a difference ensuring a wide range of benefits for children, parents and society at large (OECD, 2013a; Pianta et al., 2009; Sylva et al., 2004; Vandell & Wolfe, 2000).

This recognition and the subsequent emerging question '*how can quality be measured?*' have drawn educational researchers' attention towards developing evaluation instruments to assess quality (Ishimine & Tayler, 2014; Fenech, 2011; Grammatikopoulos, Gregoriadis & Zachopoulou, 2015). Most of these instruments are *objective*, standard-based tools, often developed in the USA albeit widely used at international level.

The international application of the same evaluation measures, despite carrying some undeniable advantages, may also leads to pitfalls, especially if the complexities – both at cultural and methodological level – of a cross-cultural use of these instruments are not taken into account (Pastori et al., 2016; Pastori & Mantovani, 2016, Pastori & Pagani, *forthcoming*).

Despite its relevance, this issue has received only marginal attention in literature and only few studies (Douglas, 2004; Fenech, 2011; Ishimine & Taylor, 2014; Mathers et al., 2007; Sheridan, 2007) have investigated the potential risks inherent in the current *globalization* of evaluation tools.

The present thesis is aimed to address this gap, focussing specifically on the *Classroom Assessment Scoring System* (CLASS – Pianta, La Paro & Hamre, 2008), an American instrument developed to assess daily interactions between teachers and children that in recent years has experienced great international diffusion.

Specifically, this study, building on and developing further the critical-cultural reflection initiated within the European project CARE (Pastori et al., 2016; Pastori & Mantovani, 2016), adopted a *mixed-methods convergent parallel design* (Creswell & Plano Clark, 2011) to analyse the implications of the CLASS application to the Italian ECEC context.

Through a qualitative approach, teachers and pedagogical coordinators from 0-3 and 3-6 services (*nidi* and *scuole d'infanzia*) were involved in discussing the tool. Their opinions and cultural beliefs about effective teaching and models to evaluate ECEC quality were elicited and compared with the perspective proposed by the CLASS. The qualitative exploration was complemented with a quantitative analysis of the tool, in order to test even at statistical level the applicability and generalizability of the CLASS framework to the Italian ECEC context. Qualitative and quantitative data were then compared to offer a more thorough understanding of the issue at hand.

Results highlight the value of adopting a critical approach to evaluation tools, attentive to the cultural and methodological complexities when these instruments are *exported* – along with their implicit values and underpinning assumptions about what ECEC quality is and how it can be assessed – to cultural contexts different from the original ones. Moreover, they offer interesting insights to a methodological reflection on the potential offered by integrating a reflective discussion with the use of standard based instruments.

Introduction

A critique does not consist in saying that things aren't good the way they are. It consists in seeing on what type of assumptions, of familiar notions, of established, unexamined ways of thinking the accepted practices are based […]. Showing that things are not as obvious as people believe, making it so that what is taken for granted is no longer taken for granted.

Michel Foucault

In recent years, educational researchers' attention has been increasingly drawn towards monitoring the quality of early childhood education and care (ECEC). This is the result of increasing recognition of the central role played by high-quality ECEC services in ensuring a wide range of benefits for children, as well as society as a whole. Not surprisingly, this interest has resulted in the development of a number of evaluation tools to measure or assess ECEC quality (Ishimine & Tayler, 2014; Fenech, 2011; Grammatikopoulos, Gregoriadis & Zachopoulou, 2015).

The dominant approach to ECEC quality (Dahlberg, Moss & Pence, 1999, 2007; Farquhar, 1999; Fenech, 2011) has been «shaped by a body of research that has predominantly been conducted in the USA, within a positivist research paradigm, and grounded in quantitative methodologies and the agendas and perspectives of researchers» (Fenech, 2011, p. 108). Mirroring the US-base of this research, most of these evaluation measures are objective, standardised tools also developed in an American context. They are nonetheless widely used at international level, to such an extent that we assist to a sort of *globalization* of evaluation tools (Pastori & Pagani, *forthcoming*).

Undoubtedly, the international application of the same evaluation measures confers certain advantages, such as offering a common ground and a shared language to compare ECEC services and to elicit continuities across countries (Grammatikopoulos, Gregoriadis & Zachopoulou, 2015; Limlingan, 2011).

However, it may also leads to pitfalls, especially if the cultural and methodological complexities of a cross-cultural use of these instruments are not taken into account.

The first danger lies in disregarding the fact that each assessment tool, being designed in a specific context for a specific purpose, presents and promotes a specific notion of quality (Fenech, 2011; Mathers et al., 2007).

The concept of quality «has achieved such dominance that it is hardly questioned. For the most part it is taken for granted that there is some thing – objective, real, knowable – called quality» (Dahlberg, Moss & Pence, 1999, p. 4). As a consequence, the discourse on quality

> «places more emphasis on the question 'how do we identify quality?' than on the preceding questions 'what do we mean by quality and why?' and 'how and by whom has quality been defined?' This in turn prioritizes methods, especially methods of measurement […]. Because the essence of quality is its absolute and universal nature, it is particularly important to remove any element of personal speculation, interpretation or judgement, any whiff of subjectivity. These suspect behaviours must be replaced by methods of measurement that are reliable and open to scrutiny and undertaken by disinterested measurers who are clearly separated from the subject of their measurement: objectivity rules. Not only does the discourse assume a reality, a thing called quality; it assumes that this reality can be perfectly captured, given adequate and carefully controlled means» (Dahlberg, Moss & Pence, 1999, p. 94).

However, although quality is often presented as a decontextualized concept, «as a universal truth that is value and culture free and applicable equally anywhere in the field under consideration» (Dahlberg, Moss & Pence, 1999, p. 94), it is actually a «value- and cultural-based concept» (OECD, 2013b, p. 35). Hence, its definition may vary across different cultural contexts and evolve over time (Dahlberg, Moss & Pence, 1999, 2007; Moss, 1994; Tobin, 2005; Tobin, Hsueh & Karasawa, 2009; Vandenbroeck & Peeters, 2014):

> «'Quality' is never an objective reality, to be finally discovered and pinned down by experts. It is inherently subjective and relative, based on values and

> beliefs, that may not only vary among and within societies, but will undoubtedly vary over time» (Moss, 1994, p. 5).

Therefore, *exporting* international assessment tools raises the issue of their «suitability» (Douglas, 2004) and «goodness of fit» (Mathers et al., 2007) when applied to different cultural contexts. In fact, assessment tools «are generally validated by reference to the values of one particular group in one country» (Douglas, 2004, p. 184), and, accordingly, they unavoidably reflect the original cultural matrix (i.e., structural characteristics of the settings, pedagogical representations, images of children and teachers…) of their *cultural cradle* (Pastori & Pagani, *forthcoming*; Pastori et al., 2016; Douglas, 2004; Mathers et al., 2007).

Thus, an uncritical use of assessment tools across cultures may lead to the local application of instruments perceived as foreign; founded on values unshared by local professionals. Such tools may also be inadequate to fully grapple with issues of quality considered relevant at a local level. Moreover, the uncritical use of foreign tools may promote a detrimentally universalistic idea of educational standards of quality. This universalism, rather than celebrating the variety of local cultures of childhood education and recognising that «the diversity of cultural ways within a nation and around the world is a resource for creativity and the future of humanity» (Rogoff, 2003, p. 18), may contribute instead to cultural homogenisation (Pastori et al., 2016; Pastori & Pagani, *forthcoming*).

In the long term, this position may even lead to serious repercussions:

> «A cost of decontextualized quality standards in early childhood education will be that local approaches that are well adapted to their local context will be driven into extinction by ideas and programs that are less context-dependent» (Tobin, 2005, p. 427).

Furthermore, an uncritical application of evaluation tools – especially in contexts different from those where these instruments were developed – may in itself open the way to a second danger, intrinsically linked to the first.

Namely, adopting these instruments without questioning their underpinning values and conceptualizations of quality may carry the inherent risk to reduce

quality to a mere «measurement without description and conceptual understanding» (Athey, 1990, p. 8). In this simplistic perspective, the tools, ceasing to be a *means* to measure quality, may become the *main and ultimate criterion*, the *perfect predefined recipe* for identifying what quality is (Vandenbroeck & Peeters, 2014), to the point of rendering any conscious thought about the process of improving quality simply superfluous (Mathers et al., 2007).

To quote from Dahlberg, Moss and Pence (1999, p. 92):

> «We can see a growing body of experts – researchers, consultants, inspectors, evaluators and so on – whose job it is to define and measure quality. Increasingly, we rely on this expert system to make judgments for us about the services we want or need for ourselves and our children. We look to these experts to tell us that what we are getting is good 'quality'. Increasingly overloaded, we seek reassurance rather than understanding, we want the guarantee of expert assessment».

Conversely, tools and the quality criteria they propose, rather than being regarded as *incontestable*, need to be critically analysed and even called into question. They need to «be used in an intelligent and reflective manner, and as part of a wide range of methods, approaches and philosophies rather than being presented as a 'complete solution'» (Mathers et al., 2007, p. 267). They need to «allow room for reflection, creativity and a variety of uses related to time, context, goals and content» (Sheridan, 2007, p. 214), stimulating professionals' critical thinking and questioning, rather than being expected to offer all the answers.

Despite the relevance of these issues, particularly in the light of the growing diffusion of American standard-based evaluation instruments (e.g., the ECERS/ECERS-R, the ITERS/ITERS-R, and the CLASS) at international level, this topic has received only marginal attention in literature. Only few studies (Douglas, 2004; Fenech, 2011; Ishimine & Taylor, 2014; Mathers et al., 2007; Pastori & Pagani, *forthcoming*; Sheridan, 2007) have problematized and investigated the phenomenon of assessment tools *migrating* to other countries and its implications.

The international project CARE (*Curriculum Quality Analysis and Impact Review of European ECEC*) figures among these few yet notable exceptions. In fact, within the CARE framework, a qualitative analysis of the implications of the cross-cultural application of standard-based tools was initiated, focusing on the *Classroom Assessment Scoring System* (CLASS – Pianta, La Paro & Hamre, 2008), an instrument developed in the USA to evaluate process quality that in recent years has experienced great international diffusion. In particular, the qualitative exploration undertaken intended to compare local theories with the values and the cultural models embedded in the instrument (Pastori et al., 2016; Pastori & Mantovani, 2016), involving teachers of few selected ECEC centers in discussing the tool. These discussions opened and stimulated a debate on the use of standardized instruments in a reflective way, representing therefore an attempt to address not only the gap evidenced by literature at cultural level, but also the methodological issue above mentioned.

The interesting findings emerging from the CARE qualitative analysis (Pastori & Mantovani, 2016) offered the starting point for the study here presented.

Specifically, this thesis aims at extending at an Italian national level the qualitative *critical-cultural* analysis started by the CARE (to the design of which the author is a contributor – Pastori et al., 2016), involving a broader number of teachers and developing further the approach, fruitfully experimented with a limited sample within the CARE project, to explore this standardized tool with ECEC practitioners and to use it as a *stimulus* to foster their reflective thinking. Moreover, the qualitative exploration was complemented with a quantitative analysis of the tool.

In detail, adopting a *mixed-methods research design* (Creswell & Plano Clark, 2011), the implications of the CLASS application to the Italian ECEC context were analysed, involving teachers and pedagogical coordinators from 0-3 and 3-6 services (*nidi* and *scuole d'infanzia*). Through a qualitative approach, Italian practitioners' cultural values and beliefs concerning ECEC quality were explored and compared with the perspective proposed by the CLASS. Moreover, their opinions about the assessment approach that this tool entails, and the potential offered by combining a reflective discussion with the use of a standard-based instrument were examined. Quantitative data were used

to test even at statistical level the applicability and generalizability of the CLASS framework to the Italian ECEC context. Results from both qualitative and quantitative data analysis were then compared to offer a more comprehensive picture and a more nuanced understanding of the issue at hand.

As with the international study, the overarching purpose of the present thesis, therefore, is not to criticize *tout court* the cross-cultural use of standardized assessment measures, nor to deny the valuable advantages that it may confer – such an approach would be as naïve as simplistic. As Douglas (2004, p. 191) clearly points out, the problem is not with the tools or – it could be added – with their international application as such; «but rather with the fundamental paradigm presented by the search for a universal measure of quality with which to assess a diversity of values, philosophy and service provision in the child care/early childhood education sector».

Rather, the intent of this thesis is to highlight the importance and the need to adopt a *critical approach* to evaluation tools, attentive to the cultural and methodological complexities when these instruments reach – along with their implicit values and underpinning assumptions about what ECEC quality is and how it can be assessed – cultural contexts different from the original ones.

The structure of the present thesis is as follows.

Part I defines the theoretical background in which this work is settled. Chapter 1 examines how in the last decades evaluation of ECEC quality has become a crucial issue at international level, and illustrates the leading role progressively assumed by the USA in developing instruments to evaluate ECEC quality. After this general presentation, the focus narrows down to a national level, analysing the peculiar conceptualization of quality that has arisen in the Italian ECEC context and, accordingly, the development of distinctive assessment instruments and of a specific evaluation culture in the country. Chapter 2 is dedicated to the CLASS tool. Specifically, its characteristics and conceptual framework are described, along with a presentation of the process that led its primary author, the American educational psychologist Robert C. Pianta, to the development of this tool. Validity issues in cross-cultural applications of the CLASS, as illustrated in the existing literature, are also discussed.

Part II focuses on the current study. Chapter 3 outlines the research framework that provides the context for this study, and illustrates its purposes, the guiding research questions, and the methods adopted. In Chapters 4 and 5 respectively, qualitative and quantitative results are exposed, while Chapter 6 provides an overall integration and comparison of the qualitative and quantitative findings. Implications for a cross-cultural application of standard-based tools and potential advantages offered by a *critical-cultural* reflection on them are also discussed therein, as are the limitations of the study and suggestions for future research.

Finally, the concluding chapter summarizes the contributions of this thesis in addressing the issue at hand, namely, extending the discussion of the CLASS tool to a broader reflection on the international use of standard-based assessment instruments, and highlighting the potential advantages offered by adopting a *critical-cultural* approach to evaluation measures.

PART I:

The theoretical framework

Chapter 1: Quality and evaluation culture in Italian ECEC services

'Would you tell me, please, which way I ought to go from here?'
'That depends a good deal on where you want to get to,' said the Cat.
Lewis Carroll – Alice's Adventures in Wonderland

1.1 The key-role of quality

Over the last decades, the social changes that have invested most developed countries, reshaping labour market policies and family structure, have led to increasing children's attendance in ECEC (early childhood education and care) services both in terms of participation rates and time that young children spend in these contexts (UNICEF, 2008; OECD, 2013a). Across OECD countries as a whole, 79% of 4-year-olds are enrolled in ECE or ECEC programmes (OECD, 2013a). Furthermore, according to the most recent data (European Commission/EACEA/Eurydice/Eurostat, 2014), on average in the 28 current European Union Member States 93% of children between the age of 4 and the starting age of compulsory education already attend ECEC, attesting that the Europe-wide benchmark (i.e. reaching at least a level of 95% participation for all EU Member States by 2020) is close to be achieved.

A rising body of research has documented that ECEC attendance brings a wide range of benefits for children, parents and society at large, such as better child well-being; more equitable child outcomes and reduction of poverty; increased intergenerational social mobility; more female labour market participation; increased fertility rates; and, at a broader level, better social and economic development. Moreover, it plays a crucial role in improving children's cognitive and language development, socio-emotional competencies and academic success, providing a crucial foundation for future lifelong learning (Camilli et al., 2010; European

Commission/EACEA/Eurydice/Eurostat, 2014; OECD, 2012; OECD, 2013a; Sylva et al., 2004; Vandenbroeck, 2010).

However, the extent to which ECEC can exert these long-term productivity benefits for society and good outcomes for children is closely linked to the quality of the ECEC provision: increasing participation rates and expanding access to services, although paramount, is not enough without giving due attention to quality (OECD, 2013a; Pianta et al., 2009; Vandell & Wolfe, 2000). Moving beyond the simple acknowledgement that early childhood education matters, as Sylva et al. (2004) stated clearly, only high quality ECEC makes a difference.

Moreover, research has emphasized that not only the magnitude of the abovementioned positive effects is conditional on quality, but has also suggested that low quality can have a long-lasting negative impact on child development (OECD, 2012).

The increasing public investment in ECEC, along with the acknowledgment of the central role of quality – and the consequential need to guarantee high quality services to children and their families that it entails – unavoidably raise a question: *how can quality be measured?*

1.2 Measuring quality in ECEC settings

The quality of ECEC services has been an international concern since the beginning of 1980s, leading to an increased focus on the issue of its evaluation (Bondioli & Savio, 2015; Sartorio & Nigito, 2005; OECD, 2015).

The year 1980 marks a milestone in the field of educational evaluation research. This is the year in which the findings from the groundbreaking *Oxford Preschool Research Project* – which, adopting a comparative methodology, discriminated between low- and high-quality services, and was aimed at advising the then Minister of Education, Margaret Thatcher, on how to improve preschool provisions for British children – were published (Bruner, 1980; Sylva, Roy & Painter, 1980; Wood, McMahon & Cranstoun, 1980).

In the same year, the US researchers Harms and Clifford (1980) developed the *Early Childhood Environment Rating Scale* (ECERS), originally designed

for children ages 0 to 6 years, that was the forebear of a family of widely used instruments aimed at assessing quality of early childhood environments[1].

In the following decades, a lively debate over quality has sparked in Europe, especially since the publication of pivotal documents by the European Commission Childcare Network (such as *Quality in childcare services*, 1990, and *Quality targets in services for young children*, 1996) and by the European Council (*Council Recommendations on Childcare*, 92/241/EEC). These documents recognize that monitoring and evaluating ECEC quality are essential for its improvement. However, rather than establishing 'euro-standards', they attempt to foster a discussion on quality and its assessment capable to take into account the variety of values, beliefs and experiences that characterize a community complex and pluralistic as Europe:

> «A balance needs to be found between defining certain common objectives, applying them to all services, and supporting diversity between individual services» (European Commission Childcare Network, 1996, p. 11).

Meanwhile, since 1980s, the USA has increasingly assumed a leading role in developing instruments to evaluate ECEC quality (Halle, Vick Whittaker & Anderson, 2010; López Boo, Araujo & Tomé, 2016; Sartorio & Nigito, 2005; Snow & Van Hemel, 2008; Ishimine & Tayler, 2014).

The following instruments can be mentioned as some of the most known research-based observation measures primarily designed to assess quality of ECEC settings at the classroom level:

1. The *Assessment Profile for Early Childhood Programs* (APECP – Abbot-Shinn & Sibley, 1992) is an observational checklist with dichotomous items that measures the global quality of classrooms for infants, toddlers and preschoolers, addressing different dimensions (Learning environment, Scheduling, Curriculum, Interacting, Individualizing);

[1] The ECERS was the first of a highly popular family of assessment instrument elaborated by Harms and Clifford (who have then been joined by Debby Cryer) to measure child care quality in different age groups (the ITERS-R for younger children and the ECERS-R for older children) or in different contexts (ECERS-R/ITERS-R for child care centers and FCCERS-R, *Family Child Care Rating Scale*, for in-home care).

2. The *Caregiver Interaction Scale* (CIS – Arnett, 1989) provides a global rating of staff/teacher sensitivity, emotional tone and responsiveness to all children in early childhood settings;
3. The *Classroom Assessment Scoring System* provides an assessment of overall classroom quality, focussing on teacher-child interactions (see Chapter 2 for a more detailed presentation of this tool). After the original CLASS Pre-K (Pianta, La Paro & Hamre, 2008) targeted to preschoolers, other versions were developed to measure quality in infant-toddler centers (respectively, CLASS Infant – Hamre, La Paro & Pianta, 2014; and CLASS Toddler – La Paro, Hamre & Pianta, 2012);
4. The *Classroom Practices Inventory* (CPI – Hyson, Hirsh-Pasek & Rescorla, 1990) assesses the developmental appropriateness of classroom and curriculum practices, teachers' behaviours, children's activities and teacher-child interactions in preschool settings;
5. The *Early Childhood Classroom Observation Measure* (ECCOM – Stipek & Byler, 2004) focuses on the quality of preschool classrooms, assessing the nature and quality of instruction as well as the social climate and management of the classroom;
6. The *Early Childhood Environmental Rating Scale – Revised Edition* (ECERS-R – Harms, Clifford & Cryer, 1998/2005) is a revision of the original scale published in 1980, and it is the most widely used measure of early childhood environments for both evaluation and research purposes. The ECERS-R primarily focuses on structural quality of early childhood programs for children aged from 2 and half to 5 years, retaining «the original scale's broad definition of environment, including those spatial, programmatic, and interpersonal features that directly affect the children and adults in an early childhood setting» (Harms, Clifford & Cryer, 1998, p.1);
7. The *Infant/Toddler Environment Rating Scale – Revised Edition* (ITERS-R – Harms, Cryer & Clifford, 2003) is a revision of the ITERS originally published in 1990. It uses the same format as ECERS-R, but it is designed to measures global quality in childcare programs for infants and toddlers. The ITERS-R «contains items to assess provision in the environment for the protection of children's health and safety,

appropriate stimulation through language and activities, and warm, supportive interaction» (Harms, Cryer & Clifford, 2003, p. 1);

8. The *Observation Record of the Caregiving Environment* (ORCE – NICHD Early Child Care Research Network, 2000) provides an assessment of caregiver's sensitivity and responsiveness to an individual child in services for infants, toddlers, and preschoolers;
9. The *Preschool Program Quality Assessment* (PQA Preschool Version – High/Scope Educational Research Foundation, 2003) is a rating instrument designed to evaluate the overall quality of preschool classrooms and to identify staff training needs.

This brief presentation, without in any way claiming to be exhaustive, is aimed at offering an insight of the US prolific production of assessment tools. As Ishimine and Tayler (2014, p. 281) suggest, «the preponderance of US-based measures reflects greater resources and an approach to producing well-developed measures that have then been used internationally». In fact, all the aforementioned instruments were designed to be sufficiently generalizable to be applied to a variety of different contexts; and many of them, going beyond the US borders, are now widely known and used at international level (albeit not without some risks, see Chapter 2).

As it will be described in the following paragraph, some of these tools have been *exported* even to our country.

1.3 Assessing ECEC quality in Italy

Since the early 1990s, the evaluation of ECEC quality has become a crucial issue also in Italy, and various instruments have been adapted or developed *ex novo* to assess quality of ECEC settings (Bondioli, 2013; Bondioli & Savio, 2015; Ferrari, 1994, 2013; Musatti & Picchio, 2010; Sartorio & Nigito, 2005).

The specific conceptualization of quality arisen in Italian ECEC can provide a premise and a framework essential to understand the assessment instruments spread in our country.

1.3.1 The conceptualization of quality in Italian ECEC

International and European debate played a crucial role in stimulating the reflection on quality and its evaluation in Italy, with regard both to early childhood education, and to primary and secondary education. However, the discourse on quality and its resultant conceptualizations have assumed different forms, characteristics and outcomes at these two educational levels.

With regard to primary and secondary education, the issue of quality is closely interwoven with the evaluation of the outcomes.

Over the last decade, the Italian school system has been invested by important changes – as exemplarily illustrated by the creation of a National Evaluation System (2004, 2012) and the initiation of a regular survey of students' learning through standardized tests (INValSI tests).

Two factors have triggered these changes: on the one hand, the importance and diffusion gradually acquired by international studies on evaluation and the comparison of school systems[2] (Schleicher, 2011); on the other hand, at national level, the process of granting schools a degree of autonomy in terms of educational, managerial and financial functions (L. 59/1997, D.lgs 59/1998, and DPR 275/1999). Specifically, the policy of school autonomy has drawn attention to the issue of accountability, encouraging the monitoring of educational quality throughout the national territory – considered as the gateway to improve the overall performances of the system (Bracci, 2009). Thus, it has contributed to laying the foundation for going beyond self-referentiality and excessive contextual relativity of the evaluation (Allulli, 2008; Bottani & Cenerini, 2003; Cerini, 2012).

The emphasis on school productivity that characterizes this segment of the educational system has resulted in a specific conceptualization of quality – conceived as a centralized, top-down defined, and objectively measurable concept. *Outcomes* and *performance* have become key words of the debate, and the results achieved in national and international standardized tests (e.g.,

[2] Since the early 90s, the OECD has regularly published a comparative report on education systems (*Education at a glance*), and, since 2000, the international comparative surveys OECD-PISA have involved an increasing number of countries, providing regular and reliable data on the knowledge and skills of their students and the performance of their education systems.

INValSI, OEDC-PISA) have assumed increasing importance as a strategic benchmark, a steadfast touchstone for school quality.

Against this background, the dimensions of care and wellbeing, the idea of school as a place to live in, the attention to spaces and timing of school life are little considered within the current National Evaluation System.

In regards to the Italian ECEC services (*nidi d'infanzia* and *scuole dell'infanzia*)[3], the reflection on quality has followed a different path (Mantovani, 2007a, 2007b; Bondioli & Savio, 2015).

The debate, although already started in 1980s, marked a turning point in 1990s, with the publication of the above-mentioned documents by the European Commission Childcare Network (1990, 1996) and the *Council Recommendations on Childcare* (92/241/EEC). Specifically, while the former proposed an idea of quality as «a relative concept based on values and beliefs», and suggested that «defining quality should be a dynamic, continuous and

[3] In Italy ECEC is a split system, organized in two different stages according to children's age: *nidi d'infanzia* (crèche, infant-toddler centers) for children from 3 months to 3 years of age, and *scuole dell'infanzia* (preschool) for children aged 3-6 – both offering full-time provision (Mantovani, 2007a).
Although a recent Law (L. 107/2015) has proposed the creation of an integrated system 0-6, at present, *nido* is still not part of the national education system. It is instead responsibility of the Prime Minister office with links with the Ministry of Labour and Ministry of Welfare, mirroring the core of the Law 1044 (1971). In fact, with the Law that enshrined the *nidi* establishment, the State announced the intent to provide child-care services for infants and toddlers in order to support working mothers, instituting a service on demand rather than ensuring universal provision. The implementation of this Law has been inconsistent and uneven, and a vast majority of services for children aged 0 to 3 has been – and is still – developed thanks to local initiative. At national level, about 25 percent of children from 0 to 3 attend *nidi* (European Commission/EACEA/Eurydice/Eurostat, 2014), but figures range from 40 percent in some cities of Emilia Romagna to 1 percent in some areas of the South (Mantovani, 2007a). Overall, 59% of children are enrolled in public (mostly municipal) services, while 41% in services provided by private organizations (Cittadinanzattiva observatory of prices and tariffs, 2015).
The scenario is significantly different as regards *scuola dell'infanzia*. In 1968 (L.444), the Ministry of Education proclaimed the right for Italian children to pre-primary education and to a universal provision of services. Although preschool provision is part of the education system and falls under the responsibilities of the Ministry of Education, institutions providing pre-primary education are also run at local level, not only by the State, but by the municipalities and the private sector – that accounts for almost a third of all children enrolled in preschools (Istat, 2012; OECD, 2014). Albeit not compulsory, preschool enrolment rates reach almost full coverage for children between 3 and 6 (over 95% coverage for children over the age of four, OECD, 2014).
The split nature of Italian ECEC services emerges also in the initial professional development and qualification-requirements that vary from 0-3 to 3-6 settings. No professional profile has been defined at national level for practitioners employed in *nidi* (*educatori*, *educatrici*: i.e., educators), and the qualification-requirements are regulated at Regional level. Instead, for practitioners employed in preschools (*insegnanti*: i.e., teachers) a 5-years University degree in Educational Studies (*Scienze della Formazione Primaria*), that enables to pre-primary and primary school teaching, is required (DM 249/2010).

democratic process» (European Commission Childcare Network, 1996, p. 11); the latter contributed to define the first criteria for early childhood education quality:

- services should be accessible and affordable;
- services should combine safe and secure care with a broad education or pedagogical approach;
- flexibility and diversity of childcare services should be encouraged, as part of a strategy to increase choice and meet the different preferences, needs and circumstances of children and families, while maintaining coherence between different services;
- training – both initial and continuous – should be ensured to workers in childcare services;
- childcare services should work closely with parents and with local communities involving regular contact and exchanges of information.

Building on and inspired by these suggestions, an intense reflection has developed from a close collaboration between research agencies and ECEC services. As a result, the following features has been identified as the core notions underpinning the concept of quality in the Italian ECEC discourse (Bondioli, 2002, 2013; Bondioli & Ghedini, 2000; Bondioli & Savio, 2010; Gariboldi, Babini & Vannini, 2014; Mantovani, 2007a, 2007b, 2014; Musatti & Picchio, 2010; Picchio, Di Giandomenico & Musatti, 2014; Zanelli, 1998):

- *quality is a process*: quality is not intended as an endpoint, a product, the final result of a process. Quality itself is a dynamic, continuous process that unfolds over time and can never be considered definitively concluded, developing in a spiral course;
- *quality is negotiated and co-constructed*: quality is a transactional process, a negotiation among a plurality of stakeholders (policy makers, pedagogical coordinators, researchers, ECEC practitioners, families…), at various levels involved in the service life and interested in clarifying and identifying values, aims, priorities, perspectives about how the service is and how it should or could be. It is not a predetermined absolute value, the fulfilment of top-down established standards, «a knowable, objective and certain truth waiting 'out there' to be discovered and described» (Dahlberg, Moss and Pence, 1999, p. 93).

Quality is rather an empty vessel that needs to be filled with the meanings co-constructed by all the actors involved. Considering quality as negotiated means to embrace different perspectives, recognizing that a *better* or *more objective* point of view on the service does not exist and that the dialogue among different viewpoints, each one unique and peculiar, is an enriching resource;

- *quality is participation*: considering quality as a negotiation among various viewpoints necessarily entails different stakeholders' active involvement in its definition and assessment;
- *quality is contextual*: quality is not regarded as a generalizable, *one-size-fits-all* concept. It rather needs to be flexibly shaped to recognize and take into account the unique characteristics of each local reality – rooted in a specific territory and with a peculiar history, educational values and pedagogical tradition;
- *quality generates self-reflection and improvement*: quality goes beyond mere monitoring or acritical meeting the defined standards. Instead, it has a formative purpose and a transformative resonance. It fosters awareness and creates fertile ground for shared reflectivity, aimed at analysing, discussing, supporting and improving the educational practice;
- *quality is here and now*: quality deals with *beings* rather than with *becomings*. It has to focus on the ongoing educational process, on the present experience that children live within the service, rather than paying attention only to their future outcomes.

This perspective on quality has had a significant impact on the Italian early childhood research and pedagogy landscape.

The first consequence is the lack of longitudinal studies to assess Italian children's long-term outcomes and the role played in their subsequent development by attending high-quality ECEC services (Brilli, Del Boca, & Pronzato, 2016; Varin, 2007). This is hardly surprising considering the emphasis on the *hic et nunc* dimension of quality:

> «Little attention is devoted to investigating the relationship between quality of educational services for early childhood and subsequent success in learning on the part of children, the so-called longitudinal effects. This is a theme in

> which Italian preschool and early education services have never shown much interest, because the cultural, political, and educational reasons to invest in early education and to define and evaluate quality are rooted in the correspondence and compatibility of the services with the community, and with the ideas and representations of children, rather than projected in a more "product-" or performance-oriented perspective that gains strength only when the compulsory school years begin. A "good" service is a service open to all children and good for them in the present, rather than a service that produces good students in the future» (Fortunati, 2007, p. 1124).

Therefore, as Picchio, Di Giandomenico and Musatti (2014, p. 135) state overtly: «we propose that evaluation will be based on the analysis of what happens to children within the ECEC service and not on its hypothesized effects on their future development».

In second place, it has deeply influenced the idea of evaluation, which, besides, is inextricably linked to the concept of quality.

On the one hand, the negotiated and participatory nature of quality has shaped the *approach to evaluation*, which, drawing on a constructivist framework, is realized *with* practitioners rather than *on* them (Becchi, 2000; Bondioli & Ferrari, 2004; Bondioli & Ghedini, 2000; Bondioli & Savio, 2010, 2015; Di Giandomenico, Musatti & Picchio, 2008). In fact, in order to acquire real meaning, evaluation data needs to be co-constructed by observers/evaluators and practitioners, who are actively involved in processes of sense-making in observing and evaluating themselves, their services and classrooms.

On the other hand, the *tools* selected and developed to assess quality in Italian ECEC services have necessarily mirrored its underlying conceptualization.

1.3.2 Tools to assess ECEC quality in Italy

As already mentioned, since 1990s, several experiences have been initiated to evaluate the educational quality of Italian ECEC services and many assessment instruments have been adapted or developed *ex novo*.

This paragraph will trace the most significant milestones in this field, presenting some of the most relevant tools elaborated in the Italian context and specifically designed to assess ECEC settings. Table 1.1 provides information on the characteristics of the measures that will be presented, as well as their purposes and key constructs.

A first, significant milestone is represented by the groundbreaking experiences carried out by researchers from the University of Pavia in strict cooperation with the Region Emilia Romagna and the Region Umbria, which desired to both assess and improve the quality of ECEC public provision. In these experiences, two American scales, the ITERS (Harms, Cryer & Clifford, 1990) and the ECERS (Harms & Clifford, 1980), were translated in Italian, resulting respectively in the publication of the SVANI (*Scala di Valutazione dell'Asilo Nido*, 1992) and the SOVASI (*Scala per l'Osservazione e la valutazione della Scuola dell'Infanzia*, 1994). Although the adaptations to the Italian context involved some modifications[4] of the original tools, their structure remained substantially unaltered.

Moving in the same direction, another research team from the University of Pavia proposed an Italian version of the ACEI scale (Darder & Mestres, 1994), the ASEI (*Autovalutazione dei servizi educativi per l'infanzia*, 2000). This tool, unlike the SVANI and the SOVASI, is conceived as an instrument for internal self-evaluation and allows teachers and educators to assess the overall service quality as they perceive it.

However, in the scenario presented, the question is: how can the adaptation and use of pre-existing assessment tools – which unavoidably convey a specific, *foreign* 'philosophy' of education and quality – be reconciled with the apparently conflicting principles underpinning the Italian reflection on quality?

The answer comes from the specific approach that has always guided the application of these instruments: *striped off* their absolute value, they have been

[4] The most relevant adjustment made by Ferrari and Livraghi to take into account the peculiar organization of Italian *nidi* regards the introduction in the SVANI of two additional items to evaluate the practices of child's *inserimento* (the first entry into the *nido* centre), which plays an important role in child's transition between family and *nido*.
With regard to the SOVASI, the most significant difference compared to the ECERS is the target age group. While the ECERS was originally conceived to assess quality of early childhood environments for children ages 0 to 6 years; the SOVASI was designed exclusively for preschool settings (children aged from 3 to 6 years). Therefore, four items (items 2, 7, 10 and 14) specifically addressed to younger children were excluded and other two items (items 13 and 31) were slightly modified.

regarded as a *filter* to look critically at the service from a different perspective, a *stimulus* to elicit a thorough discussion on educational quality among practitioners (Bondioli & Ferrari, 2004).

Ferrari explains clearly the core of this approach (2002, p. 28) in regard to the use of the SVANI:

> «The group to which I belong has never used the SVANI to establish ranking lists of services, or to press for a passive emulation of the excellent level defined by the instrument; but rather to initiate a reflection on quality evaluation among coordinators and educators, to train those working in the *nido* to use such tools in order to build a profile to be discussed together. However, above all [our] working group wished to contextualize the evaluation process triggered, to stimulate an internal discussion on the *nido* starting from a tool that could offer a decentralized perspective to reflect on their own reality, to get back to the here and now coming from far away» (translation by the author).

Nonetheless, this approach has not prevented noticing the limits embodied in the use of tools developed in different cultural and educational contexts. The adapted instruments were «found to neglect several elements that were considered crucial in the Italian ECEC culture […]. Thus, new tools and procedures of evaluation were developed in different sites, following the local educational culture» (Musatti & Picchio, 2010, p. 149).

The first, emblematic experience that effectively illustrates this new orientation regards the development of the ISQUEN (*Indicatori e Scala di valutazione della Qualità Educativa del Nido* – Becchi, Bondioli & Ferrari, 1999). Recalling the principles of Guba and Lincoln's *Fourth Generation Evaluation*[5] (1989), the ISQUEN is the result of a process of critical discussion on the SVANI, which involved researchers, pedagogical coordinators and educators from the Region Umbria.

The process started with the recognition that «some indicators featured in the American instrument did not seem to be completely in line with a specific

[5] Deeply rooted in the constructivist framework, the *Fourth Generation Evaluation* (FGE) relies on negotiated co-creation of social reality. According to FGE, the boundaries and parameters of the *evaluand* (i.e., the entity being evaluated) should be generated by the participants themselves via a process of expression of their claims, concerns and issues, and subsequent negotiation that contributes to stakeholders' empowerment.

image of the *nido* shared by practitioners of the Region» (Ferrari, 2004a, p. 157, translation by the author). For instance, the SVANI does not properly investigate educators' professionalism and expertise (such as practices of observation, documentation, designing and planning educational and care activities…); and it does account only marginally families' participation in the service life – both key-aspects of quality according to the Italian pedagogical tradition (Ferrari, 2004a; Sartorio & Nigito, 2005).

Due to its shortcomings, the SVANI did not appear to be completely adequate to assess the quality of the Umbrian *nidi*. Therefore, researchers and educators proceeded to a critical analysis and 'deconstruction' of the SVANI that led to the development of a new instrument, suitable for *capturing* the peculiar characteristics of this local reality.

In the following years, further tools[6] were elaborated *ex novo* adopting a similar participatory approach, in which constant dialogue and negotiation with practitioners and other stakeholders played a key role. In fact, their active involvement in the co-construction of those instruments ensured to fully respect and capture the peculiar characteristics of the Italian (and often even local) ECEC services.

[6] As illustrative examples, the following tools could be mentioned (see Table 1.1 for a concise description of their characteristics and the process that led to their development): the AVSI (Bondioli & Ferrari, 2008), the DAVOPSI (Bondioli & Nigito, 2008), the SCIN (Zanelli, Sagginati & Fabbri, 2004), the SPRING (Marcuccio & Zanelli, 2013), the PRADISI (D'Ugo, 2013), the Ri.Qua (Gariboldi, Babini & Vannini, 2014), and the instruments elaborated in the services of the Municipality of Milan (Franchi & Caggio, 1999) and in Region Tuscany (Region Tuscany & Istituto degli Innocenti, 2006).

Table 1.1

Selection of instruments elaborated in the Italian context to assess quality in ECEC (Alphabetical order)

INSTRUMENT	AUTHOR(S), YEAR	DEVELOPMENT PROCESS OF THE INSTRUMENT	PURPOSE	UNIT OF ANALYSIS	TARGET AGE GROUP	KEY CONSTRUCTS OF MEASURE	TYPE OF EVALUATION
Autovalutazione dei servizi educativi per l'infanzia: ASEI	Darder & Mestres (1994-2000)	The tool is the Italian adaptation of the ACEI realized by Gusmini	To assess the overall service quality as perceived by the staff and to support practitioners' reflection	ECEC service	Children aged from 0 to 6 years (*Nido*, *Scuola dell'infanzia*)	Educational project, Organization and management of the service	Self-evaluation, Formative evaluation
Autovalutazione della scuola dell'infanzia: AVSI	Bondioli & Ferrari (2008)	Researchers elaborated the instrument. Then, it was tested on the field and underwent the critical analysis of a panel of practitioners and experts in EC education	To assess quality of preschool, regarded as a complex formative environment	ECEC service	Children aged from 3 to 6 years (*Scuola dell'infanzia*)	Educational experience, Professional activities, Adults and their relationships, Guarantees	Self-evaluation, Formative evaluation
Dispositivo di Analisi e Valutazione dell'Organizzazio-ne Pedagogica della scuola dell'infanzia: DAVOPSI	Bondioli & Nigito (2008)	The tool is the result of research and in-service training experiences that involved practitioners in Region Liguria. Then, it was tested on the field and underwent the critical analysis of various stakeholders	To assess relevant aspects of preschool organization and to foster teachers' reflection	ECEC service	Children aged from 3 to 6 years (*Scuola dell'infanzia*)	Timing, Space (Indoor and outdoor), Grouping	Self-evaluation, Formative evaluation
Indicatori e Scala di valutazione della Qualità Educativa del Nido: ISQUEN	Becchi, Bondioli & Ferrari (1999)	Researchers and Umbrian educators realized a critical analysis and a 'deconstruction' of the SVANI that led to the development of a new instrument, adequate to capture the peculiar characteristics of this regional reality	To assess the global quality of the *Nido* and to support practitioners' reflection on the service structure, its functioning and its educational project	ECEC service, Classroom	Children aged from 0 to 3 years (*Nido*)	Subjects, Contexts and practices, Practitioners' expertise, Guarantees	Self-evaluation, Formative evaluation

Table 1.1. Continued

INSTRUMENT	AUTHOR(S), YEAR	DEVELOPMENT PROCESS OF THE INSTRUMENT	PURPOSE	UNIT OF ANALYSIS	TARGET AGE GROUP	KEY CONSTRUCTS OF MEASURE	TYPE OF EVALUATION
Manuale di rilevazione della qualità	Region Tuscany & Istituto degli Innocenti (2006)	The tool is the result of an action-research that involved practitioners, pedagogical coordinators, parents and other stakeholders in Region Tuscany	To assess the overall service quality	ECEC service	Children aged from 0 to 3 years (*Nido*)	Structural features, Human resources, Organizational features, Educational style, Relationships with the territory and other stakeholders, Maintenance of facilities and hygiene	Formative evaluation, External evaluation
Prassi Didattiche dell'Insegnante di Scuola dell'Infanzia: PRADISI	D'Ugo (2013)	The tool is the result of a research that involved pedagogical coordinators and teachers in the Municipality of Bologna, starting from the recognition of the limitations of the scale SOVASI. Then, it was tested on the field and underwent the critical analysis of a panel of experts in EC education	To assess teacher's educational practices aimed at fostering children's development	Individual teacher	Children aged from 3 to 6 years (*Scuola dell'infanzia*)	Daily routines, Promotion of children's skills, Teacher's educational choices	Self-evaluation, Formative evaluation, External evaluation

Table 1.1. Continued

INSTRUMENT	AUTHOR(S), YEAR	DEVELOPMENT PROCESS OF THE INSTRUMENT	PURPOSE	UNIT OF ANALYSIS	TARGET AGE GROUP	KEY CONSTRUCTS OF MEASURE	TYPE OF EVALUATION
Questionario elaborato negli asili nido milanesi	Franchi & Caggio (1999)	The tool is the result of an experience of reflection and discussion on the SVANI that involved pedagogical coordinators and educators of the Municipality of Milan. It was developed in order to integrate the SVANI with additional information and better address the peculiar characteristics of this local reality	To assess the global quality of the *Nido*	Classroom	Children aged from 0 to 3 years (*Nido*)	Space and materials, Personal care routines, Safety and maintenance, Learning activities, Daily work organization for adults' and children's wellbeing, Relationships with preschool	Self-evaluation, Formative evaluation, External evaluation
Riflessione Qualità: RI.QUA	Gariboldi, Babini & Vannini (2014)	The tool is the result of an action-research that involved pedagogical coordinators of FISM preschool in the Province of Bologna	To assess the quality of the service and to foster educators' reflection on their educational practices	ECEC service	Children aged from 3 to 6 years (*Scuola dell'infanzia*)	Identity, School experience, Organization of the context, Reflecting on the experience, Constraints and resources	Self-evaluation, Formative evaluation
Strumento per la Costruzione/Condivisione dell'Identità pedagogica dei Nidi: SCIN	Zanelli, Sagginati & Fabbri (2004)	The tool is the result of an action-research that involved pedagogical coordinators and educators in the Province of Forlì-Cesena	To assess the quality of the service and to foster educators' reflection on their educational practices	ECEC service	Children aged from 0 to 3 years (*Nido*)	Educational context, Teamwork, Relationships with parents	Self-evaluation, Formative evaluation

Table 1.1. Continued

INSTRUMENT	AUTHOR(S), YEAR	DEVELOPMENT PROCESS OF THE INSTRUMENT	PURPOSE	UNIT OF ANALYSIS	TARGET AGE GROUP	KEY CONSTRUCTS OF MEASURE	TYPE OF EVALUATION
Scala per l'Osservazione e la valutazione della Scuola dell'Infanzia: SOVASI	Harms & Clifford (1980-1994)	The tool is the Italian adaptation of the ECERS realized by Ferrari & Gariboldi	To assess the global quality of the *Scuola dell'infanzia*	Classroom	Children aged from 3 to 6 years (*Scuola dell'infanzia*)	Personal care routines, Furnishings and display for children, Linguistic and cognitive experiences, Motor activities, Creative and expressive activities, Social development, Adult needs	External evaluation
Strumento per lo Sviluppo di Processi Riflessivi e Indagini valutative nei Nidi da parte dei Gruppi di lavoro educativi: SPRING	Marcuccio & Zanelli (2013)	Pedagogical coordinators and educators in the Province of Forlì-Cesena were involved in the revision process of the tool SCIN	To assess the quality of the service and to foster educators' reflection on their educational practices	ECEC service	Children aged from 0 to 3 years (*Nido*)	Organization of the educational context, Teamwork, Contextualization and flexibility, Relationships with families and the territory, Evaluation processes	Self-evaluation, Formative evaluation, External evaluation
Scala di valutazione dell'asilo nido: SVANI	Harms, Cryer & Clifford (1990-1992)	The tool is the Italian adaptation of the ITERS realized by Ferrari & Livraghi	To assess the global quality of the *Nido*	Classroom	Children aged from 0 to 3 years (*Nido*)	Furnishings and display for children, Personal care routines, Listening and talking, Learning activities, Interaction, Program Structure, Adult needs	External evaluation

Table 1.1. Continued

INSTRUMENT	AUTHOR(S), YEAR	DEVELOPMENT PROCESS OF THE INSTRUMENT	PURPOSE	UNIT OF ANALYSIS	TARGET AGE GROUP	KEY CONSTRUCTS OF MEASURE	TYPE OF EVALUATION
Strumenti per valutare la qualità della scuola materna elaborati dal Comune di Milano	Franchi & Caggio (1999)	Preschool teachers and pedagogical coordinators were involved in a process of discussion and analysis that led to the construction of two assessment instruments (Tools A and B), adequate to address the peculiar characteristics of the services in the Municipality of Milan	**Tool A** To assess global preschool quality and the teachers' professionalism	Classroom	Children aged from 3 to 6 years (*Scuola dell'infanzia*)	Service organization, Safety and wellbeing, Space and materials, Educational and teaching practices, Diversity and differences, Educational continuity, Teacher's professionalism, Relationships/ interactions	Self-evaluation, Formative evaluation, External evaluation
			Tool B To assess the global quality of the Scuola dell'infanzia	ECEC service	Children aged from 3 to 6 years (*Scuola dell'infanzia*)	Children, Educational project, Elements and instruments of professionalism, Organizational constraints and resources	Self-evaluation, Formative evaluation, External evaluation

It is important to stress that all these evaluation experiences share some common features. First, they all strongly echo the distinctively Italian conceptualization of quality abovementioned.

In second place, in all these experiences, the tool development was carried out in the framework of *formative evaluation* (Bondioli & Ferrari, 2004; Bondioli & Savio, 2015). In this perspective, rather than being a 'report card' or a 'final judgement', the assessment provided by the instrument was intended as a starting point for enhancing reflective thinking and awareness and, thus, for improving service quality. In order to fully achieve this goal, evaluation became a dialogical and *transactional* process (in the sense used by Dewey and Bentley, 1949), in which the crucial phase was represented by the *restituzione* (i.e., data returning) and discussion of the data collected with all the actors of the service (Becchi, 2000; Ferrari, 2004b; Gusmini, 2004).

Finally, in many cases, the *formative evaluation* framework paved the way for and was accompanied by processes of *metaevaluation*, in which the evaluative experience and the assessment tools adopted themselves became subject of evaluation. In fact, using a tool within a formative approach (even a locally co-constructed tool) does not imply an uncritical alignment to the model of quality and education that it coveys (Becchi, 2000; Bondioli, 2013; Bondioli & Ferrari, 2004; Bondioli & Savio, 2015; Ferrari, 2013; Vannini, 2014). Therefore, the instrument itself should be examined, discussed and tested to verify the extent to which it includes and represents practitioners' educational ideas, values and goals.

Hence, even the identification of divergences between the pedagogical model proposed by the instrument and the one shared within the service can be an enriching opportunity. In fact, it should not imply the elimination or change of the dissonant practice on the behalf of a supposed superiority of the tool, nor an overall rejection of the instrument. Rather it should be the beginning of a further stage of investigation aimed at questioning the reasons for these divergences. It should represent a reflective exercise, an occasion not only to enhance practitioners' awareness of the strengths and limits of the tool, but also to think about and to make explicit their own educational values, ideas of good practices, beliefs and quality conceptions (Bondioli, 2013; Vannini, 2014).

1.4 Taking stock of the Italian ECEC evaluation culture today

As discussed in the previous paragraphs, the evaluation culture in Italian ECEC settings has been profoundly shaped by the peculiar vision of quality emerged in our country, according to which:

> «Quality should be discussed, shared among all actors […]. It is the balance to work towards and the prospect to achieve, what we aspire to, [what] we primarily want to ensure and implement, rather than unequivocally "measure it"» (Mantovani, 2014, p. 25, translation by the author).

This notion has inspired many experiences that have led to the construction of various assessment tools, adopting a participatory approach, often attentive to local characteristics.

Moreover, it has contributed to strip assessment instruments off their 'prescriptive aura' and to consider them as interlocutors to dialogue with, compare with and even to question.

However, in order to offer a more thorough image of the present scenario, some additional clarifications should be provided.

Firstly, although the evaluation of ECEC quality has become a crucial issue in Italy since the 1990s, a strong evaluation culture is not fully established yet. In fact, as a recent survey (Savio, 2015) seems to suggest, the diffusion of evaluation experiences is not pronouncedly widespread across the Nation and still presents considerable blind spots, even in those Regions (such as Emilia Romagna or Tuscany) that have been particularly active in this field[7].

Moreover, a creeping, subtle suspicion towards evaluation seems to persist (Calzolari, D'Ugo & Vannini, 2010; D'Ugo & Vannini, 2014). Although practitioners recognize that evaluation processes are a crucial condition for professional development and for improving service quality (Savio, 2015), «the

[7] The survey (Savio, 2015) involved 209 respondents – the majority of which was represented by *nido* educators (98), pedagogical coordinators (64) and preschool teachers (17), contacted in occasion of a national congress (*Convegno nazionale del Gruppo nazionale Nidi-Infanzia*, Montecatini, March 2012). Despite its limited and selected sample, some data provided by this survey can offer an interesting insight into the current state of the Italian ECEC evaluation culture. For instance, it is noteworthy that 46.88% of respondents declared that they had never participated to evaluation experiences, and that the majority of them (74.76%) perceived the evaluation culture as spread too thinly among ECEC services.

topic continues to be highly contradictory and is often dismissed by equating evaluation with words such as stigma, judgment, selection» (Calzolari, D'Ugo & Vannini, 2010, pp. 12-13, translation by the author). This feeling is particularly emphasized when the evaluation process relies on assessment tools not negotiated, «developed by others, often theoreticians, who do not know what it means 'being in school' with children in a socio-cultural and political context constantly changing and increasingly challenging» (Calzolari, D'Ugo & Vannini, 2010, pp. 15, translation by the author). In this perspective, a participatory approach to evaluation can represent a way to face the ghosts that still surround evaluation and to approach assessment tools in a more accessible, tamer form that allow to see them as a resource for dialogue, rather than as an 'intransigent judge' (Ferrari, 2013).

Against this background, not surprisingly, most of the standard-based assessment instruments validated and widely diffused at international level are rarely used – and sometimes even barely known – in our country.

This is the case with the *Classroom Assessment Scoring System* (CLASS – La Paro, Pianta & Stuhlman, 2004; Pianta et al., 2008), that will be presented in the following Chapter.

Chapter 2: The *Classroom Assessment Scoring System* (CLASS): theoretical framework and validity issues in cross-cultural applications

No significant learning occurs
without a significant relationship.
James Comer

The CLASS is a standardized observational system based on developmental and educational theory, which assesses quality of educational contexts, focussing specifically on daily interactions between teachers and children in classroom environments (La Paro, Pianta & Stuhlman, 2004; Pianta, La Paro & Hamre, 2008). This tool, developed by the American educational psychologist Robert C. Pianta and his colleagues, has experienced in recent years great diffusion and appreciation at international level.

However, the CLASS is only the final step in Pianta's reflection and research work, which started in 1990s and led to the definition of a theoretical framework to portray teacher-child relationships and to the development of specific tools to investigate them.

Briefly retracing the path taken by this author can contribute to better understand the characteristics, strengths and limits of the CLASS, to which this chapter is dedicated.

2.1 Before the CLASS

2.1.1 A conceptual model of teacher-child relationships

Pianta's interest in teacher-child relationships arose from his personal experience as a teacher, during which he realized that relationships – more than instructional or learning processes – are the true linchpin around which the school experience is organized (Pianta, 1999).

This initial insight was then confirmed by a sizable literature that provided evidence that strong and supportive relationships play a key role in children's both academic and social-emotional development, and represent a promotive and protective resources to all children and especially to those with identified *risk factors* (e.g., children from disadvantaged social, economic, and cultural backgrounds, see Birch & Ladd, 1998; Hamre & Pianta, 2001, 2006; Pianta, 1999; Gregory & Weinstein, 2004). However, it also constituted the starting point for Pianta's research, which culminated in 1999 with the proposal of an interdisciplinary model (Pianta, 1999) – then updated and extended in 2003 (Pianta, Hamre & Stuhlman, 2003) – to conceptualize relationships between children and teachers.

Although in previous decades this topic had been addressed by diverse areas of psychology, the study of teacher-child relationships was not an area of inquiry unto itself until the 1990s. The lack of focus was mainly due to the poor dialogue and integration across diverse theoretical frameworks. They all separately examined definite components of teacher-child relationships; however, none of them alone could adequately encompass the intrinsic complexity of this construct and offer a comprehensive understanding of the dynamic, multilevel interactions that take place in schools (Pianta, 1999; Pianta, Hamre & Stuhlman, 2003).

Overcoming the limits of this 'insularity', the model proposed by Pianta had the merit of interweaving different theoretical traditions, drawing on principles and concepts of developmental systems theory (DST – Ford & Lerner, 1992; Lerner, 1998; Sameroff, 1995), attachment theory (Bowlby, 1969), and developmental psychopathology paradigm (Cicchetti & Cohen, 1995).

The model focuses at the level of teacher-child relationships as the key unit of analysis:

> «A relationship between a teacher and child is not equivalent to only their interactions with one another, or to their characteristics as individuals. A relationship between a teacher and a child is not wholly determined by that child's temperament, intelligence, or communication skills. Nor can their relationship be reduced to the pattern of reinforcement between them. Relationships have their own identities apart from the features of interactions or individuals» (Pianta, Hamre & Stuhlman, 2003, p. 206).

Teacher–child relationships are conceptualized as open dyadic systems that are not only affected by actual behaviors and qualities of the partners, but also by each individual's mental representation of the relationship (Pianta, 1999; Pianta, Hamre & Stuhlman, 2003). In fact, as depicted in Figure 2.1, they are complex entities constituted by several components:

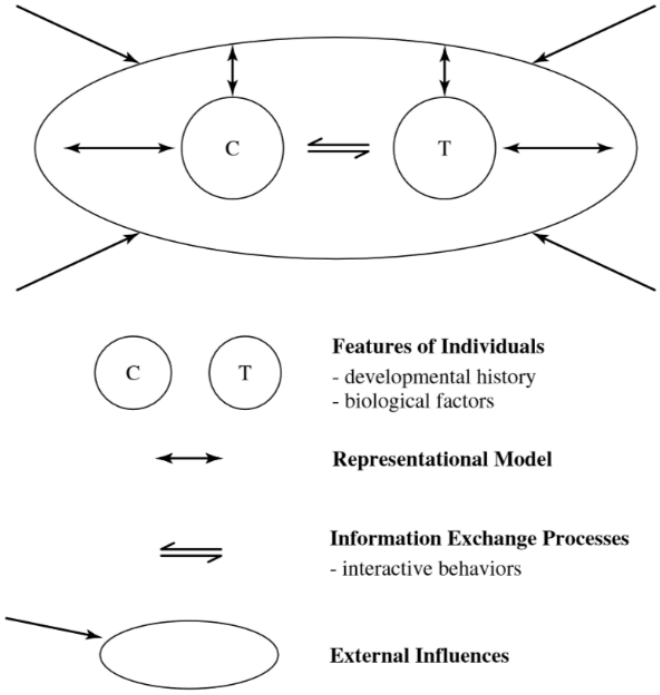


Figure 2.1

A conceptual model of teacher-child relationship

Source: Pianta, Hamre & Stuhlman (2003)

1. child's and teacher's *individual features* (biologically predisposed characteristics, personality, self-perceptions and beliefs, developmental history…);
2. *representational models* (each individual's representation of the relationship and of the roles of each relational partner). These models are themselves conceived as open systems, since the information stored in them, while fairly stable, is open to change based on new experiences;
3. *processes of feedback* by which information is exchanged between teacher and child (behavioral interactions, language, and communication).

Due to their systemic nature, relationships are more than simply the sum of these components. They are rather a product of the dynamic, reciprocal interactions over time of feedback processes, representations, and characteristics of the two individuals involved.

Moreover, the model assumes that relationships interact with systems at similar levels (e.g., families and peer groups) through a transactional process, and are exposed to external influences of the systems in which they are embedded (schools, classrooms, communities, cultures). For instance:

> «Cultures can prescribe timetables for expectations about students' performance or the organization of schools that can shape how students and teacher relate to one another. […] State regulations mandate standards for student performance that affect what teachers must teach, and at times how they must teach it» (Pianta, Hamre & Stuhlman, 2003, p. 213).

Furthermore, although, both teacher and child bring an assortment of goals, feelings, needs, and behavioral styles that can affect their experiences with one another in the classroom, teacher-child relationships also embody a certain degree of asymmetry. This asymmetry reflects differences in roles and maturity of the relational partners and its balance is subject to considerable variation across age, grade, or schools. Therefore, the responsibility for the quality of the relationship lies mainly with the teacher.

2.1.2 Assessing the quality of the relationship from the teacher's perspective

A further step in Pianta's reflection was to outline practical implications of this conceptual model.

Specifically, Pianta suggested that, in order to enhance the protective relational resources potentially available within educational contexts, it is necessary, in the first instance, to assess quality of teacher-child relationships. In fact, assessment is the fundamental prerequisite for then improving those relationships that are not – adapting Donald Winnicott's (1896-1971) terms – *supportive enough* (Pianta, 1999).

Although recognizing that each of the components described in his conceptual model could be a valid entry point for intervention, Pianta initially focused his attention on teachers' perspective and specifically on their representational models, proposing two tools, the STRS and the TRI, to assess them.

The *Student-Teacher Relationship Scale* (STRS – Pianta, 1994, 2001) is a self-report instrument designed to assess teacher's perception of his/her relationship with a particular child. The STRS measures student-teacher relationship patterns in terms of conflict, closeness, and dependency, as well as the overall quality of the relationship.

The *Teacher Relationship Interview* (TRI – Pianta, 1999) is a semi-structured interview that elicits teachers' internal representations of the relationship with a particular pupil. During the interview, teachers describe the relationship with the child, providing examples of specific types of interactions and describing their own and the child's affective responses to these interactions.

The STRS is currently the only standardised and validated instrument available for assessing teachers' representational models and is especially useful as a screening measure to identify relational difficulties or strengths. However, the TRI can provide a more variegated, in-depth picture that can represent a useful starting point for relationship-focused consultation with teachers (Koomen et al., 2006).

2.2 The CLASS: a tool to assess effective teacher-child interactions

2.2.1 The theoretical framework

Further developing his reflection on teacher-child relationships, but shifting the focus from teachers' representational models to behavioral interactions, Pianta proposed another instrument, the *Classroom Assessment Scoring System* (Pianta, La Paro & Hamre, 2008).

Although other tools to assess quality of educational settings (e.g., see Paragraph 1.2) generally attach particular or exclusive importance to structural features of classrooms (such as curriculum, adult-child ratio, group size, teacher education, physical and organizational aspects of the classroom…), the CLASS focuses solely on *process quality*. Specifically, this observation instrument analyses teacher-child interactions – in their socio-emotional and instructional components – and what teachers do with the materials they have (La Paro, Pianta & Stuhlman, 2004; Pianta, La Paro & Hamre, 2008).

As Pianta and colleagues point out, daily interactions between teachers and children are among the most feasible aspects of teachers' jobs that can be reliably observed and assessed (Hamre et al., 2013).

Most importantly, although both structural features and dynamic aspects of classrooms are important for children's social and developmental outcomes,

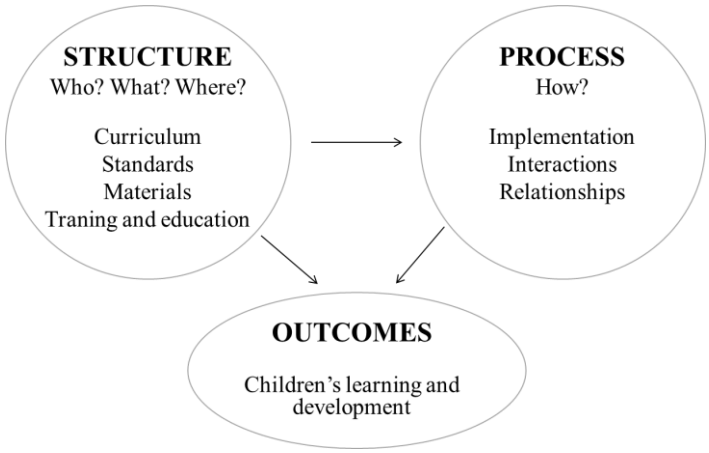


Figure 2.2

Structure, process and outcomes diagram

Source: La Paro, Hamre & Pianta (2012)

research has shown not only that process quality provides powerful predictors of children outcomes (Hamre et al., 2013; Howes et al. 2008; Mashburn et al. 2008; Montie, Xiang & Schweinhart, 2006), but also that «the structural quality is mediated by, of flows through, process quality» (La Paro, Hamre & Pianta, 2012, p. 2 – see Figure 2.2). Therefore, structural features of the classroom become valuable when teachers fully exploit them for engaging children and providing learning opportunities; hence, the term *teacher effectiveness* or *effective teaching* (La Paro, Hamre & Pianta, 2012).

Furthermore, addressing both emotional and instructional features of the classroom, the CLASS provides – with a few notable exceptions (Eccles & Roeser, 1999; Pressley et al., 2003) – a more comprehensive and systematic framework than most other models of classroom practice (Hamre et al., 2007).

The CLASS measure presents a *multilevel, latent structure*, in which a wide range of effective classroom interactions is organized in broad categories, called *domains*. Within each domain are *dimensions*, which capture more specific details about teacher-child interactions hypothesized to be relevant in promoting children's learning and social development. Each of these dimensions, in turn, is described by explicit *indicators*, which are then operationalized in specific, observable descriptions of classroom proximal interactions, defined *behavioural markers* (Hamre et al., 2007; La Paro, Hamre & Pianta, 2012).

Although the various versions of the instrument currently available (to assess effective teaching in infant, toddler, preschool, elementary, and secondary education classrooms and settings) share the same multi-level structure, the ways these domains and dimensions are manifested and organized are peculiar to particular developmental levels (Hamre et al., 2007; La Paro, Hamre & Pianta, 2012). For instance, the Pre-K version (Pianta, La Paro & Hamre, 2008) organizes effective teacher-child interactions into three broad domains (Emotional Support, Classroom Organization, and Instructional Support) embracing overall ten dimensions; whereas the domains featuring in the Toddler version (La Paro, Hamre & Pianta, 2012) are only two – Emotional and Behavioral Support, and Engaged Support for Learning – organized in eight dimensions. Tables 2.1 and 2.2 on the next pages provide an overview of the CLASS domains and dimensions from the Pre-K and Toddler versions.

Table 2.1

Domains and dimensions of the CLASS Pre-K (Pianta, La Paro & Hamre, 2008)

Domain	Dimension	Description
Emotional Support	Positive Climate	Reflects the overall emotional tone of the classroom and the connection between teacher and students.
	Negative Climate[8]	Reflects the overall level of expressed negativity in the classroom.
	Teacher Sensitivity	Encompasses the teacher's awareness and responsiveness of students' academic and emotional needs.
	Regard for Student Perspectives	Reflects the degree to which the teacher's interactions with students and classroom activities place an emphasis on students' interests, motivations and autonomy.
Classroom Organization	Behavior Management	Captures the teacher's ability to anticipate problem behavior and use effective methods to prevent and redirect misbehavior.
	Productivity	Reflects how well the teacher manages instructional time and routines, and offer students opportunities to be involved in learning activities.
	Instructional Learning Formats	Considers how well the teacher maximizes students' interest, engagement and ability to learn from activities.
Instructional Support	Concept Development	Focuses on the teacher's use of instructional discussions and activities to promote students' high-order thinking skills and cognition.
	Quality of Feedback	Reflects the degree to which the teacher provides effective feedback to expand students' learning, understanding and persistence.
	Language Modeling	Assesses the quality and amount of the teacher's use of language-stimulation and language-facilitation techniques.

[8] In both CLASS Pre-K and Toddler versions, Negative Climate is the only dimension that uses a reverse scoring scale in which low scores indicate little if any negative climate present in the classrooms and, therefore, a higher classroom quality.

Table 2.2

Domains and dimensions of the CLASS Toddler (La Paro, Hamre & Pianta, 2012)

Domain	Dimension	Description
Emotional and Behavioral Support	Positive Climate	Reflects the overall emotional tone of the classroom and the connection between teacher and children.
	Negative Climate	Reflects the overall level of expressed negativity in the classroom.
	Teacher Sensitivity	Encompasses the teacher's awareness and responsiveness of children's individual needs and emotional functioning.
	Regard for Child Perspectives	Reflects the degree to which the teacher's interactions with children and classroom activities place an emphasis on children's interests, motivations and autonomy.
	Behavior Guidance	Captures the teacher's ability to promote behavioral self-regulation in children and to use effective methods to support positive behavior and minimize problem behavior.
Engaged Support for Learning	Facilitation of Learning and Development	Considers how well the teacher facilitates activities to support children's learning and developmental opportunities.
	Quality of Feedback	Reflects the degree to which the teacher provides effective feedback to expand students' learning, understanding and participation.
	Language Modeling	Assesses the quality and amount of the teacher's use of language-stimulation and language-facilitation techniques.

As highlighted by the authors, in this balance between its *latent structure*, hypothesized as *grade-invariant*, and the *heterotypic continuity*, that accounts for variation across grades and ages in the dimensions of teacher-child relationship and their specific behavioral displays, lies another innovative characteristic of the tool.

In other words, the diverse CLASS versions provide context-specific and developmentally sensitive parameters for each age level. Nonetheless, the tool offers a common metric and vocabulary across grades, addressing the need of continuity and coherence in education. In fact, an underlying assumption of the CLASS is that effective teacher-child interactions share commonalities across

age levels, mirroring an *invariant latent structure* (Hamre et al., 2007; La Paro, Hamre & Pianta, 2012).

This idea is fully expressed and exemplified considering that the organization of teacher-child interactions into three major domains (Emotional Support, Classroom Organization, and Instructional Support) that characterizes the CLASS Pre-K is also shared by all the versions for the subsequent educational levels. In fact, according to the authors, it corresponds to a unique underlying conceptual framework (*Teaching through Interactions framework* – Hamre & Pianta, 2007; Hamre et al., 2013), which posits that these three areas of interaction are important for students from preschool age through high school since the effect of teachers and classrooms on students' learning is located in the interactions that take place between teachers and students.

2.2.2 The observation procedure

The CLASS tool allows observing and coding nearly all of the activities, both structured and unstructured, that take place in classroom settings – with the only exception of recess and outdoor play. Snack and mealtime can be coded as well, since, especially in ECEC services, a lot of teaching can occur even during these times (Pianta, La Paro & Hamre, 2008).

According to the manual (La Paro, Hamre & Pianta, 2012; Pianta, La Paro & Hamre, 2008), the live observation procedure includes a minimum of four (up to a maximum of six) 30-minute cycles.

During each cycle, the observer is required to watch, without interruption, everything that happens at the classroom level, with particular attention to the teachers' behaviors and responses, for a period of 15-20 minutes. Notes should be taken for each dimension during every observation cycle: these notes will provide the basis for coding and will help the observer to make judgement about a code.

The observation is followed by a 10-minute period for recording codes. The observer should derive numerical rating for each CLASS dimension per observation cycle, based on the range, frequency, intention and tone of interpersonal and individual behaviour during the observation time. Codes are

assigned abiding by the manual that provides detailed information on the dimensions, the specific indicators, and observable behaviour markers as well as anchor examples.

Each dimension is rated on a 7-point Likert scale ranging from 1 to 7. The score represents the extent to which that dimension is characteristic of the classroom:

- scores of 1-2 (low range) mean the classroom is low on the aspect described by the dimension examined;
- scores of 3-5 (mid-range) are given when classrooms show a mix of effective interactions with periods when interactions are not effective or are absent in relation to the dimension considered;
- scores of 6-7 (high range) mean that, with regard of that specific dimension, effective teacher-child interactions are consistently observed throughout the observation period.

Furthermore, the manual (La Paro, Hamre & Pianta, 2012; Pianta, La Paro & Hamre, 2008) suggests some principles that should guide the observer during observation and scoring:

- Remaining objective: «the observer must guard against injecting external explanations for what he or she sees taking place within the classroom» (Pianta, La Paro & Hamre, 2008, 12). When assigning codes, observers should not be influenced by any information other than what they have directly observed and should not take the perspective of the teacher;
- Independence of cycle: each cycle must be considered independently of the others;
- Independence of dimensions: due to a certain extent of overlap among different dimensions, a single event in the classroom may contribute to the scoring on more than one dimension. However, each dimension must be rated independently;
- Weighting single incidents: the score should reflect the experience of the average child in the classroom across the whole observation cycle. Therefore, a single incident should not be given too much weight in the overall score.

The CLASS has been validated also in coding video-clips of classroom (Mashburn et al., 2007). Videotaping may be completed by teachers or by outsiders, and the same procedure described above for live observation applies to coding video-clips.

2.2.3 The observation training and certification

The CLASS should be scored by trained and certified observers. In fact, in order to use this tool, it is necessary to attend a two-day observation training and then pass a reliability test. The test relies on an online format, providing each applicant five videos to be watched and coded. At least 80% reliability with the master codes is required to be certified as a reliable observer[9].

An annual recertification is demanded to maintain the licence and ensure adherence to the criteria established by the CLASS over time.

2.2.4 Psychometric properties

Several studies have provided evidence of the good psychometric properties of the instrument applied to US classrooms.

Particularly, in relation to *reliability*, the observers' certification process underpins a clear and comprehensive understanding of the CLASS purposes and procedures, and ensures high interrater reliability.

Moreover, research data suggest that CLASS scores are highly stable across time and that observations for the recommended four cycles provide an adequate sampling of stable features of teacher-child interactions across the diverse domains (La Paro, Hamre & Pianta, 2012; Pianta, La Paro & Hamre, 2008).

With regard to *validity*, the authors have demonstrated that the CLASS is associated empirically with other measures of similar constructs (criterion validity). For instance, CLASS Toddler and Pre-K scores show convergent and divergent validity in correlation respectively with the ITERS-R and the ECERS-

[9] Each applicant has to observe and score videos previously coded consensually by at least three CLASS master coders. Ratings that are within one point of each other (along the 1-7 rating scale) are considered to reflect an acceptable degree of accuracy in rating.

R. As expected, the relations are stronger with ITERS-R/ECERS-R scales that assess process quality and lower with those scales that measure structural quality features (Pianta, La Paro & Hamre, 2008).

Furthermore, results from several studies have showed that classroom quality, as assessed by the CLASS, is linked to various children's academic, social, and behavioral outcomes during their early years, as well as at the end of preschool and first grade (predictive validity – Hamre et al., 2013; Howes et al., 2008; La Paro, Hamre & Pianta, 2012; Mashburn et al., 2008; Pianta, La Paro and Hamre, 2008).

Finally, recent studies have provided evidence that the organization of classroom interactions into three broad domains of effective teaching postulated by the Teaching through Interactions model fits observational data collected from a sample of over 6,400 US classrooms from preschool to twelfth-grade (Hafen et al., 2015; Hamre et al., 2013).

2.2.5 Uses of the CLASS tool

The CLASS has been used as part of many research projects interested in documenting quality of classroom environment and exploring its association with children's current or future development and learning, covering a broad array of US classroom contexts (Hamre et al., 2013; La Paro, Hamre & Pianta, 2012; Pianta, La Paro & Hamre, 2008).

It has also been used for program planning and evaluation, as well as in professional development programs (e.g., MyTeachingPartner system). In particular, with regard to the latter aspect, the authors highlight that the CLASS tool can provide teachers with an objective and direct feedback of their actual classroom practices pinpointing their areas of strength and challenges. Furthermore, it can contribute to preservice teachers' professional development, offering a framework for understanding the components of their teaching that really matter for children (La Paro, Hamre & Pianta, 2012; Pianta, La Paro & Hamre, 2008).

Moreover, the CLASS has been applied for accountability purposes and is currently included in the regular monitoring of the federal Head Start and Early

Head Start programs (Aikens et al., 2012; La Paro, Hamre & Pianta, 2012). In addition, several states in the US have integrated the CLASS into their quality rating and improvement systems.

Besides, the interest in this tool is not limited to the United States. In fact, it has been extensively used at international level to evaluate efficacy of teacher-child interactions and to analyse the relationship between its scores and several children's outcomes in Australia (Tayler et al. 2013), Belgium (Buyse et al., 2008; Declercq & Laevers, 2015), Canada (Bouchard et al., 2014), Chile (Leyva et al., 2015; Treviño, Toledo & Gempp, 2013), China (Hu et al., 2016), Colombia (Maldonado-Carreño & Votruba-Drzal, 2014), Ecuador (Araujo et al., 2014), Finland (Pakarinen et al., 2010; Salminen et al., 2012), Israel (Ziv & Aviezer, 2014), France (Dessus, Cosnefroy & Joët, 2014), Germany (von Suchodoletz et al., 2014), the Netherlands (Slot, 2014), Portugal (Cadima et al., 2010).

2.3 Validity issues in cross-cultural applications of the CLASS

In the light of these characteristics (a solid, research-based theoretical framework; a good psychometric foundation; documented links with children outcomes; an observation procedure not excessively time-consuming…), the growing popularity of the CLASS both in the US and at international level is not surprising.

However, despite its diffusion, recent studies (Ishimine and Tayler, 2014; Pastori & Pagani, *forthcoming*; Sandilos et al., 2014) have discussed some issues that question the validity of this tool when applied in cultural contexts different form its original one.

According with the purposes of the present thesis and with the specific interest in ECEC quality, in the next paragraphs, these issues will be examined paying particular attention to literature focused on the CLASS Toddler and Pre-K.

2.3.1 The reliability testing

A first crevice highlighted by scholars in the otherwise solid framework provided by the CLASS regards its albeit innovative procedure to be certified as a reliable observer.

The reliability test guarantees that each certified observer has a solid understanding of the CLASS tool, and enables the CLASS to maintain a high standard of inter-rater reliability.

However, as long as the videos provided for the test depict US classroom and are therefore representative of US culture, the soundness of this procedure is questionable when the tool is applied at international level. As Ishmine and Tayler (2014) have clearly pointed out: *is it still a valid process to gain and maintain the certification as a reliable observer using US-based classroom videos and then apply the skills in international contexts?*

2.3.2 The structural validity of the CLASS in international contexts

Another issue regards the international validation of the CLASS framework (Sandilos et al., 2014).

CLASS validity is highly established in the US and several researchers have tried to adapt this tool to different cultural ECEC contexts, investigating its psychometric properties and its structural validity (Bouchard et al., 2014; Cadima et al., 2010; Declercq & Laevers, 2015; Dessus, Cosnefroy & Joët, 2014; Leyva et al., 2015; Pakarinen et al., 2010; Slot, 2014; Treviño, Toledo & Gempp, 2013; von Suchodoletz et al., 2014; Ziv & Aviezer, 2014).

Some of these studies have suggested that the factorial model underpinning the CLASS framework does not always sufficiently describe classroom quality when applied to international contexts.

For instance, Pakarinen and colleagues (2010) proposed a validation of the CLASS Pre-K using data from 49 Finnish kindergarten classrooms. Results of the confirmatory factor analysis (CFA) indicated that the three-domain model (i.e., the Teaching through Interaction framework, Hamre et al., 2013) did not fit the data well. A closer inspection revealed that the Negative Climate

dimension showed poor discriminatory validity: since most classrooms scored similarly, at a low level, this dimension was not relevant to differentiating classroom quality in Finland. Thus, they tested whether a single-factor solution could improve the fit of the model to the Finnish data, but found no evidence to support this alternative model. Then, the authors decided to remove the Negative Climate item from the original three-factor model and allow the residual of the Quality of Feedback item to correlate with the Concept Development item. The resulting final model fitted the data well. However, the validity of removing a dimension from a structured instrument that should include it was not examined and its consequences not questioned.

To cite another example, Leyva and others (2015) tested the three-factor model in 91 public prekindergarten classrooms in Chile. To increase up to an acceptable level the model fit, they introduced a modification in the Teaching through Interactions model, allowing the following observed variables to correlate: Behavioral Management and Positive Climate, Productivity and Teacher Sensitivity, Negative Climate and Productivity, and Instructional Learning Formats and Regard for Student Perspectives. Thereby, the authors concluded that their data supported the validity of the Teaching through Interactions framework in Chile and «the generalizability of the inferences regarding the value of teacher-child interactions in children's learning and development, based on a conceptual framework and an observational measure developed in North America, to urban Chile» (Leyva et al., 2015, p. 14). However, those residual correlations were not included in the original model (Hamre et al., 2013). Thus,

> «although the correlation of residuals was considered to be a relatively minor modification, it still presents a concern for the CLASS model, as the correlations reveal associations among the dimensions that are not being explained by the three domain factors. These associations could be resulting from the presence of key characteristics measured by the indicators that are not explained by the current factors» (Sandilos et al., 2014, p. 910).

Moreover, as Mathers and colleagues (2007, p. 268) argue:

> «While helpful in terms of dovetailing with local contexts, this [adapting the tools, dropping or amending elements which do not fit with the local context] raises some issues in terms of comparability. If the scales are altered – that is, users adhere to certain elements of the scales and discard others – then they are no longer the reliable and valid instruments proven by research. In addition, if different users retain and discard different parts of the scale/s, they are no longer valuable as comparative tools».

Similar psychometrical inconsistencies with the original model have been reported also in other countries (e.g., Bouchard et al., 2014; Declercq and Laevers, 2015; Dessus, Cosnefroy & Joët, 2014; von Suchodoletz et al., 2014). However, although the CLASS has been used internationally quite extensively, these validity issues have not been properly investigated yet.

These findings suggest that more research is needed to clarify potential differences in the structure and function of classroom interactions at the international level, and raise questions about applicability of the CLASS outside the US: *beyond statistics, what are the meaning of these inconsistencies and their implications for daily interactions in the classrooms? Might they mirror cultural features of effective teacher-child relationships not captured by the tool?*

2.3.3 The cultural sensitivity issue

These considerations draw attention to an even more crucial issue, so far only marginally addressed (Pastori et al., 2016; Pastori & Mantovani, 2016; Pastori & Pagani, *forthcoming*).

The CLASS framework takes into account that teacher-child interactions can be affected by cultural variability and recognizes that behavioural markers can be culturally bound (Hamre, Goffin & Kaft-Sayre, 2009; Pianta & Hamre, 2009a; Vitiello, 2013).

For instance, both in the Toddler and Pre-K versions, the Positive Climate dimension includes, inter alia, the indicator Respect, that examines the respect that teaches and children demonstrate for each other. One of the behavioral markers for this indicator is eye contact. The CLASS recognizes that, although

this behavior conveys respect in US contexts, it may assume different meanings in other cultures that, for example, can discourage eye contact between adults and children. Therefore, in these situations, the CLASS manual invokes observer sensitivity and recommends not considering eye contact among display of respect, relying on other behavioural markers to assess Positive Climate (Hamre, Goffin & Kaft-Sayre, 2009; Pianta & Hamre, 2009a).

Still, as argued in the CARE project (Pastori et al., 2016; Pastori & Mantovani, 2016), the acknowledgment of cultural differences, albeit not totally absent, seems to be limited to the behavioural level, without affecting the overarching structure and the dimensions it postulates. In fact, the underlying assumption that «children benefit from high quality teacher-child interactions, as defined by CLASS, regardless of race or ethnicity» (Hamre, Goffin & Kaft-Sayre, 2009, p. 62) suggests that the dimensions of classroom quality assessed by the CLASS are relevant across cultures (Hamre, Goffin & Kaft-Sayre, 2009; Pianta et al., 2009; Vitiello, 2013; Hamre et al, 2013):

> «The Teaching through Interactions framework offers only one window into teacher effectiveness. However, we argue that interactions among teachers and students are among the most important aspects of teachers' jobs. There is also initial evidence that this model for understanding classroom interactions is consistent across other cultures» (Hamre et al., 2013, p. 482).

However, as studies drawing on the field of socio-cultural and anthropological research advise, culture shapes the way that adults and children interact (Vygotsky, 1978; Alexander, 2000; Rogoff, 2003; Tobin, 2005; Tobin, Hsueh & Karasawa, 2009). Thus, conclusions about the supposed universality of any conceptualizations of quality (regarding structural features, as well as pedagogical and interactional aspects) should be drawn with caution, since ECEC quality, early childhood pedagogies and teachers-child relational patterns «are not universal or culture free but instead are reflection of values and concerns of particular people in a particular time and place» (Tobin, 2005, p. 426). In fact, as Rogoff (2003) points out, they might have both a universal and a cultural-related nature.

It should be stressed that this caveat, that seems overshadowed in the CLASS framework, was instead taken into account within the conceptual model

of teacher-child relationships proposed by Pianta himself in 1999 (see Paragraph 2.2.1), which recognized the significant role of culture in affecting interactions between teachers and pupils.

Hence, transposing these reflections to the current discussion on the CLASS, the emerging questions are: *is it enough to account cultural variability and complexity only at the behavioural-marker level? Can the conceptualization of effective teacher-child interactions and the subsequent dimensions proposed by the CLASS transcend cultural differences?*

These are the questions that contributed to guide the design and realization of the research here presented.

PART II:

The present study

Chapter 3: Methodology

Every research arises from a curiosity,
a problem, a question.
Susanna Mantovani

3.1 The CARE project

The starting point for the present study was offered by the European project *Curriculum Quality Analysis and Impact Review of European ECEC* (CARE).

The CARE (January 2014-December 2016) is a collaborative project that includes 11 partners and countries, covering all regions of Europe. It is funded by the European Union to address issues related to the quality, inclusiveness, and individual, social, and economic benefits of early childhood education and care in Europe[10].

More specifically, the CARE central objective is to develop an evidence-based and culture-sensitive European framework of developmental goals, quality assessment, curriculum approaches, and policy measures for improving the quality and effectiveness of early childhood education and care.

One of the main research actions of the CARE project was aimed at identifying common and culturally different key-elements of process quality. In this respect, a quantitative multiple case-study of 0-3 and 3-6 ECEC centers in 8 European countries[11] was carried out, and video-data from four 'good practice' ECEC centres (two infant-toddler centers and two preschools) in each country were collected (Reggio Emilia and Milan in Italy). The CLASS tool

[10] 7th Framework Programme SSH-2013, European coordinator Paul Leseman (Utrecht University). Belgium, Denmark, Italy, Finland, Germany, Greece, the Netherlands, Norway, Poland, Portugal, and the UK are taking part to the CARE project. For a more detailed presentation of the project, see the CARE website: http://ecec-care.org.
[11] Denmark, England, Finland, Germany, Italy, the Netherlands, Poland and Portugal.

(Toddler and Pre-K versions) was applied to the European video-clips to carry out a quantitative analysis of the encodings[12] (Slot et al., 2016).

In order to enhance the tool ecological validity and to make this quantitative study more comprehensive, the Italian team[13] proposed to complement it with a parallel qualitative and cultural ethnographic study (Gillen et al., 2007; Rogoff, 2003; Tobin, Hsueh & Karasawa, 2009). This qualitative study involved European practitioners and experts in ECEC in discussing the tool, assumed as a powerful highlighter to elicit different cultural perspectives about ECEC quality and to compare them with the pedagogical traits and values implicitly present in the tool itself (Pastori et al., 2016; Pastori & Mantovani, 2016).

3.2 Purpose and research questions

The present study, acknowledging the interest and relevance of the topic explored by the qualitative study conducted within the European project, aims at developing further the *critical-cultural discussion* on the CLASS initiated by the CARE.

Its main overarching objective is to shed light and address in depth the gap – evidenced by the existing literature (see Introduction and Paragraph 2.3) – in examining the implications, both at cultural and methodological level, of applying this standard-based tool outside of its *cultural cradle*, with specific regard to the Italian ECEC context.

Therefore, firstly, the study extended at national level the qualitative exploration of the tool, involving a broader number of teachers from 0-3 and 3-6 services (*nidi* and *scuole d'infanzia*), and integrated the qualitative study with a quantitative analysis of the tool.

[12] M. K. Lekkernan (Jyvaskyla University) is the international leader of the WP2 (Curriculum, Pedagogy, and Classrooom Quality: promoting effectiveness of ECEC, Task 2.3); P. Slot (Utrecht University), J. Cadima (Instituto Universitário de Lisboa) and J. Salminen (Jyvaskyla University) are responsible for the quantitative analysis. J. Salminen also conducted a qualitative discourse analysis on educational dialogues.

[13] The Italian research team, supervised by Professor Susanna Mantovani, consists of researchers from University of Milan-Bicocca – Department of Human Sciences (G. Pastori, C. Bove, P. Braga, F. Zaninelli, S. Cescato, V. Pagani, G. Banzi e T. Morgandi) and from Reggio Children (C. Giudici, C. Rinaldi, P. Cagliari, M. Castagnetti, S. Bonilauri, L. Colla, M. Ruozzi, M. Nicolosi – till November 2014). The *critical-cultural* study on the CLASS was led by G. Pastori, V. Pagani and S. Mantovani.

Moreover, recognizing the peculiar characteristics that the evaluation culture has assumed in Italian ECEC services (see Chapter 1) and its apparent distance from the *objective* approach implemented by the CLASS, the exploration of the tool at methodological level was developed further, examining the potential opportunities offered by combining a reflective discussion with the use of a standard-based instrument.

Specifically, to investigate these issues, the following research questions were posited:

1. *Are there continuities between the conceptualization of effective teaching rooted in the Italian pedagogical tradition and the one embedded in the tool? and are there any differences, discrepancies or missing elements – transcending the behavioural-marker level – in the framework provided by the CLASS that are nonetheless crucial to fully capture the quality of teacher-child interactions in the Italian context?*
2. *Do the possible incongruences emerged at qualitative level mirror statistical inconsistencies in the CLASS framework when applied to the Italian data?*
3. *Is it possible to integrate in a fruitful way two apparently opposing perspectives on quality evaluation such as the standard-based assessment provided by the CLASS and the participatory-reflective approach typical of the Italian early childhood pedagogical tradition?*

3.3 Method

3.3.1 Research design

To answer these research questions, this study was conducted using a *mixed-methods research design* (Creswell & Plano Clark, 2011). Specifically, a mixed-methods *convergent parallel design* (see Figure 3.1) was adopted. The purpose of this design is «to obtain different but complementary data on the same topic» (Morse, 1991, p. 122) to gain a more thorough understanding of the research problem. In this design, the researcher uses concurrent timing to collect two independent strands of qualitative and quantitative data during the same phase

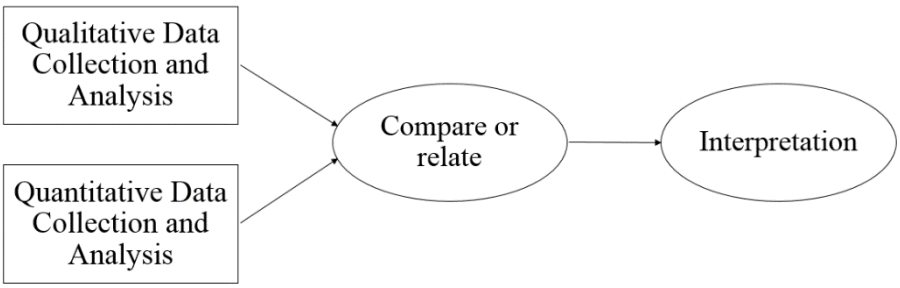


Figure 3.1

The convergent parallel design

Source: Creswell & Plano Clark (2011)

of the research process, prioritizing the methods equally and keeping the strands independent during analysis. Then, the two sets of results are merged into an overall interpretation looking for convergence, divergence, contradictions, or relationships of the two sources of data (Creswell & Plano Clark, 2011).

This choice was led by the recognition – at the core of the mixed-methods approach – that, since «qualitative research and quantitative research provide different pictures, or perspectives, and each has its limitations […] the combination of quantitative and qualitative data provide a more complete understanding of the research problem than either approach by itself» (Creswell & Plano Clark, 2011, p. 8).

In the present case, a mixed-methods design seemed the best choice not only to address all the research questions posited, but also to enhance the solidity of the study.

In fact, a qualitative approach alone, despite being the most appropriate for eliciting Italian practitioners' viewpoints and opinions about teacher-child interactions and the CLASS tool, did not allow to compare the current results with the ones emerged from the other studies that have explored and tested the CLASS framework outside the USA adopting exclusively a quantitative approach.

On the other hand, relying only on quantitative data – as other researchers have done before – entailed significant limitations (see Paragraph 2.3.2). As

literature suggests, a mere statistical investigation offers a too narrow picture of the issue examined if not supported by the detailed, in-depth understanding of the problem provided by qualitative data. In fact, despite acknowledging the possible psychometric inconsistencies of the CLASS when applied to other cultural contexts, it cannot explain the meaning of these inconsistencies, nor identify their implications for daily interactions in the classrooms.

Therefore, in this study, a qualitative approach was adopted to: a) explore Italian practitioners' cultural values and beliefs concerning effective teaching and ECEC quality and to compare them with the perspective proposed by the tool; b) involve participants in discussing the evaluation model proposed by the tool also at methodological level.

Quantitative data (i.e., preschool and infant-toddler centers classroom observations coded using, respectively, the CLASS Pre-K and Toddler) were used to describe classroom quality as postulated by the CLASS in Italian preschools and infant-toddler centers, and to examine the applicability and generalizability of the Teaching through Interaction framework to the Italian ECEC context. Results from both qualitative and quantitative data analysis were then compared to offer a more comprehensive picture and to bring greater insight into the problem.

3.3.2 Participants

The present study gathered data from 23 preschool and 7 infant-toddler centers classrooms[14] (see Table 3.1 for an overview of participants' characteristics).

Firstly, the procedures and the primary goals of the study were explained in detail to pedagogical coordinators, teachers and parents. Teachers were selected to participate on a voluntary basis and were asked for their written consent. Similarly, parents were asked for consent for their children's participation. Moreover, informed consent was requested from the 8 pedagogical coordinators who expressed their interest in participating to the qualitative part of the study.

[14] The four 'good practice' classrooms (two preschool and two infant-toddler center classrooms) selected for the CARE case-study aforementioned (see Paragraph 3.1) were included in the sample.

Preschools. Teacher–child interactions were observed in 23 preschools classrooms drawn from 9 public preschools (5 state-run preschools – 14 classrooms, and 4 municipal preschools – 9 classrooms) in the Italian Provinces of Arezzo, Como, Milano, Monza-Brianza and Reggio Emilia.

Forty-six preschool teachers (44 female, 2 male), two of whom were special education teachers, participated to the study. Teachers were on average 47.11 years old (min = 25, max = 63, SD = 10.34) with experience working in ECEC settings ranging from 9 months to 41 years (M = 22.74 years, SD = 12.07). All teachers had at least a secondary education degree[15], and eight of them held university degrees[16].

Preschool class size (i.e., total number of children enrolled in the class) ranged from 19 to 29 children (M = 25.09, SD = 2.68). All classrooms served children between the ages of three and six, provided full time service, and were Italian speaking.

Class size and the number of teachers in the classroom (both classroom teachers and special education teachers) were used to calculate the child-teacher ratio. The ratio ranged from 7.33 to 14.50 (M = 11.87, SD = 2.31). However, this ratio was generally lower at the beginning and at the end of the school day since teacher shifts overlapped only for the central hours – i.e., generally from 10.30 a.m. till 1.30 p.m. (M = 2.87 hours, min = 2, max = 4.5, SD = 0.86).

On average, 15.59 (SD = 6.46) children were present during the observation cycles (min = 5, max = 27), depending on the type of activity (routine, small group activities, large group activities, meal/snack…) taking place.

Infant toddler centers. The study involved 29 teachers from 7 infant-toddler centers (6 municipal infant-toddler centers and 1 service run by a Consortium, involving collaboration between public institutions, no profit organizations and private companies) in the Italian Provinces of Milano and Reggio Emilia. All participants were females and on average 35.40 years old (min = 25, max = 51, SD = 7.91). Their work experience in day-care ranged from 10 months to 29

[15] Although a 5-years University degree in Educational Studies is currently required to teach in Italian preschools (DM 249/2010), the majority of the preschool practitioners involved in the study began their teaching career well before the promulgation of the pertinent Law.

[16] Nine teachers did not fill out completely the questionnaire reporting their education or work experience.

years (M = 11.72 years, SD = 7.31). All teachers had at least a secondary education degree, and six of them held university degrees.

Teacher-child interactions were observed in 7 classrooms drawn from 5 infant-toddler centers[17].

Classroom size ranged from 8 to 30 children (M = 21.43, SD = 6.83). All classrooms served children from 15 months to 3 years of age, provided full time service, and were Italian speaking.

The child-teacher ratio ranged from 4.00 to 6.33 (M = 5.48, SD = 0.84), and the co-presence of at least two teachers was guaranteed throughout the entire school day.

On average, 6.10 (SD = 2.97) children were present during the observation cycles (min = 2, max = 18), depending on the type of activity (routine, small group activities, large group activities, meal/snack…) taking place.

[17] All participating infant-toddler center teachers were involved in the qualitative part of the study, whereas observational data were gathered only from 7 classrooms to date. Due to delays in collecting parental informed consents, classroom observations still have to be conducted in 5 classrooms.

Table 3.1

Participants' and service characteristics

	PRESCHOOL				INFANT-TODDLER CENTER			
	N				N			
Services involved	9				7			
Teachers participating in the qualitative study	46				29			
Teachers participating in the quantitative study	46				15			
Classrooms video-observed	23				7			
	M	**DS**	**Min**	**Max**	**M**	**DS**	**Min**	**Max**
Teacher's age (years)	47.11	10.34	25	63	35.40	7.91	25	51
Teachers' working experience (years)	22.74	12.07	0.75	41	11.72	7.31	0.83	29
Class size	25.09	2.68	19	29	21.43	6.83	8	30
Child-teacher ratio	11.87	2.31	7.33	14.50	5.48	0.84	4.00	6.33

3.3.3 Data collection: procedure and measures

Qualitative data. The qualitative part of the study aimed at exploring Italian practitioners' cultural values and beliefs concerning effective teaching and ECEC quality, comparing them with the CLASS perspective, and eliciting their opinions about this instrument and the standard-based assessment approach that it entails.

In each ECEC service, teachers and pedagogical coordinators were involved in a cycle of three reflective seminars organized in the following steps:

1. a very first step consisted in providing an introduction to the CLASS (Toddler or Pre-K version). Specifically, the theoretical framework was presented, CLASS domains and dimensions were explained in detail showing also US exemplary video-clips, and observational and scoring procedures were described;
2. in the second seminar, participants observed a video-clip from another Italian infant-toddler center/preschool, and encoded it using the CLASS. Then, the codes assigned by the certified observer (CLASS perspective) and by Italian practitioners (pedagogical-cultural perspective) were compared. This step was specifically designed to allow teachers and pedagogical coordinators to become more familiar with the tool and gain a better understanding of what aspects of teaching practices the CLASS lens encompasses, practicing with a material more 'neutral' and less 'emotionally charged' than their own video-clips;
3. finally, participants observed the video-clips from the infant-toddler center/preschool they were employed in, and the feedback provided by the CLASS was shared and discussed.

In each reflective seminar, teachers and pedagogical coordinators were involved in focus groups to elicit their opinion about the CLASS tool and discuss its framework. The set of questions that guided these discussions is presented in Table 3.2.

On the one hand, the CLASS was assumed as a powerful trigger to explore and make explicit participants' pedagogical values and teaching choices through

the comparison with the cultural perspective embedded in the tool itself (Pastori et al., 2016; Pastori & Mantovani, 2016).

On the other hand, discussing the CLASS represented an opportunity to reflect on quality evaluation and, becoming more familiar with procedures and requirements of this standard-base tool, to appraise more knowingly complementary advantages and limits of this assessment approach compared to the participatory-reflective one.

Table 3.2

Guiding questions (from Pastori et al., 2016)

MAIN THEMES	SUB-TOPICS	QUESTIONS
Cultural values and pedagogical practices	Continuities	Are there any dimensions/indicators in the instrument that seem familiar? If any, which ones?
	Disagreements	Are there any dimensions/indicators in the instrument that you would eliminate? If any, which ones?
	Missing elements	Are there any dimensions/indicators in the instrument that you would add (missing dimensions/indicators in the tool you consider key-ones of the teacher-child relationship)? If any, which ones?
	Differences	Are there any dimensions/indicators in the instrument that you perceive as more exposed to a different cultural interpretation? If any, which ones?
Quality evaluation approaches	General opinion	What do you think about the standard-base assessment approach proposed by the CLASS tool?
	Strengths	Do you see any interesting elements in this approach? If any, which ones?
	Limits	Do you see any limits/weaknesses in this approach? If any, which ones?

Quantitative data. The quality of teacher-child interactions in preschools and infant-toddler centers was assessed using, respectively, the CLASS Pre-K (Pianta et al., 2008) and Toddler (La Paro, Hamre & Pianta, 2012).

Preschool observations were carried out from March to June 2015, whereas infant-toddler centers observations started in March 2015 and are still ongoing.

Classrooms were videotaped on a randomly chosen school day. Observations begun in the morning when the instructional activity started (about 9.30 a.m.) and lasted approximately 3 hours (up to lunchtime). Observations were scheduled in order to ensure that classroom videotaping occurred on days that were typical of the usual environment for the classroom (i.e., not on a day when half the class or the classroom teachers were absent or sick; not when festivals, special projects or festivities that could disrupt the regular daily routine were planned).

Moreover, a visit of the observer was planned before starting the videotaping to let children explore the video-equipment, get accustomed to the presence of the observer, and address the novelty in the classroom. Preschoolers were told why they were being videotaped prior the beginning of the video-observation and were allowed to share any concern about the process.

Six 15-20 minute video segments (cycles) were selected to code each classroom videotapes. Selection criteria were developed in accordance with the observation procedure guidelines included in the CLASS manual (La Paro, Hamre & Pianta, 2012; Pianta et al., 2008), and with previous studies that have used the CLASS for video-observations outside the United States (Araujo et al., 2014; Hu et al., 2016; Leyva et al., 2015). These criteria comprised:

1. each segment lasted a minimum of 15-20 minutes without interruption[18];
2. segments could include both structured and unstructured times;
3. segments included a sampling of the school morning (beginning, midday, and end of the morning);
4. at all times there were at least a teacher and five preschoolers or two toddlers in the video.

[18] As the CLASS manual suggests (La Paro, Hamre & Pianta, 2012; Pianta et al., 2008), shorter observational segments can be accepted and coded nonetheless provided that they have a minimum duration of 10 minutes. The sample exceptionally included 4 shorter segments, lasting from 11 to 14 minutes, that met the criteria described (e.g., teachers ended an activity and children went to recess before the completion of the 15-20 minute cycle; a child, whose parents did not give their consent to take part to the study, stepped into the classroom in the middle of the video-observation).

These criteria ensured that selected segments were representative of the average experience of children in the observed classrooms.

The classrooms were observed by the author, who went through the rigorous CLASS reliability process and exceeded the recommended level of reliability (i.e., 80% of codes within and between dimensions within one scale point of agreement) for both the Toddler and Pre-K versions prior to data collection. The trained observer watched each of the six classroom video segments selected, and then rated each dimension on the 7-point scale for about 10 min, as recommended by the CLASS manual.

Teachers were also asked to complete questionnaires on teacher and classroom demographics. Questionnaires were given to the teachers at the end of the observation day and returned by mail or on the occasion of a subsequent seminar scheduled to discuss the tool.

3.3.4 Data analysis

Qualitative data. The focus groups with teachers and pedagogical coordinators were audiotaped and transcribed verbatim. Then, the qualitative data were analysed conducting a thematic analysis (Boyatzis, 1998; Braun & Clarke, 2006).

This approach aimed at identifying common, salient patterns within qualitative data, and, «through its theoretical freedom, […] provides a flexible and useful research tool, which can potentially provide a rich and detailed, yet complex, account of data» (Braun & Clarke, 2006, p. 78).

Specifically, a semantic approach was adopted, focusing on the explicit meaning of the data, rather than on the latent or interpretative level (Boyatzis, 1998; Braun & Clarke, 2006). The analysis was applied to the content of the entire data set, and was conducted from a data-driven inductive perspective (Boyatzis, 1998), in which patterns and themes linked to research questions are drawn from the data rather than being imposed on them prior to data collection and analysis (i.e., deductive or theory-driven coding).

In accordance with the guidelines proposed by Braun and Clarke (2006), the thematic analysis was conducted following these steps: (a) familiarization with the data, (b) initial coding, (c) searching for recurring themes, (d) reviewing themes, (e) defining and naming themes, and (f) reporting.

Quantitative data. The quantitative part of the study aimed to describe quality of teacher-child interactions as defined by the CLASS in Italian ECEC settings, and to examine to what extent the Teaching through Interaction framework was adequate to understand classroom processes in the Italian ECEC context.

Prior to the analysis, data screening was conducted at univariate level for both the Pre-K and Toddler dataset, using the computer software package Statistical Package for Social Scientists (SPSS 23.0) and following the procedures outlined by Kline (2011) and Tabachnick and Fidell (2001).

Then, the descriptive statistics of the individual CLASS Pre-K and Toddler dimensions were examined and compared with those collected in previous studies conducted at international level (respectively, samples from the USA, Chile, China, Finland, and Germany in the case of the CLASS Pre-K, and samples from the USA and the Netherlands with regard to the CLASS Toddler).

Next, structural validity investigations were carried out. Due to the limited number of observation cycles realized in infant-toddler centers, this analysis was limited only to the data collected in preschools[19]. First, the magnitude of

[19] The reliability of factor analysis is dependent on sample size, and a wide range of recommendations regarding sample size in factor analysis has been proposed.
For instance, regarding the minimum necessary sample size (N), Gorsuch (1983) and Kline (1979) recommended that N should be at least 100. Guilford (1954) claimed the minimum desirable N to be 200. Comrey and Lee (1992) classed 300 as a good sample size, 100 as poor and 1000 as excellent. Even considering the minimum N per measured variable (p), the suggested ratio varies considerably according to different scholars. For instance, Gorsuch (1983) proposed a minimum ratio of 5. Everitt (1975) recommended that the ratio should be at least 10. Kass and Tinsley (1979) recommended to have between 5 and 10 cases per variable.
As a result, «the wide range in these recommendations causes them to be of rather limited value to empirical researchers» (MacCallum et al., 1999, p. 97). Moreover, further research has suggested that such guidelines are not sufficiently sensitive to a variety of important characteristics of the data (MacCallum et al., 1999; Velicer & Fava, 1998). For instance, MacCallum and colleagues (1999) have shown that level of communalities of the measured variables plays a critical role. Specifically, when each common factor is overdetermined (i.e., at least three or four measured variables represent each common factor) and the communalities are high (i.e., ≥ 0.70 on average), accurate estimates of population parameters can be obtained with samples as small as 100. With communalities in the 0.5 range, samples between 100 and 200

the correlations between the individual dimensions of the CLASS was examined. Then, exploratory factor analysis (EFA) was used to identify the factor structure underlying the current set of data without imposing any restrictions based on *a priori* theoretical assumptions regarding the relations between observed variables (Fabrigar et al., 1999). A principal components extraction method with direct oblimin rotation was performed using SPSS 23.0. Expecting that the resultant factors would be correlated (Hamre et al., 2007; Hamre et al., 2013), oblimin rotation was preferred over varimax rotation.

Finally, confirmatory factor analysis (CFA) was carried out for the Teaching through Interaction three-factor model suggested for the CLASS Pre-K (Hamre et al., 2007; Hamre et al., 2013), that assumes three positively correlated latent constructs (i.e., the three CLASS domains: Emotional Support, Classroom Organization, and Instructional Support; see Figure 3.2). The aim was to examine whether this model would also fit the Italian data. In accordance with Hamre et al. (2013), alternative models were also tested (i.e., a single domain model of Effective Teaching, and a two-domain model of Social and Instructional Support). AMOS 23.0 software was used to estimate each model.

The goodness of fit of the estimated models was evaluated using multiple fit indices as recommended by Tanaka (1993) and Schermelleh-Engel, Moosbrugger & Müller (2003), in order to measure how well each considered model represented the data drawn from the current sample (Kline, 2011).

Due to the sensitivity of the chi-square test (χ^2) to sample size (Kline, 2011; Schermelleh-Engel, Moosbrugger & Müller, 2003), this statistic was deemphasized when evaluating the fit of each model, and alternative goodness-of-fit measures were also examined. Thus, the absolute fit indices primarily considered in this study consisted of:

a. the root mean square error of approximation (RMSEA);
b. the standardized root-mean-square residual (SRMR).

---

can be good enough provided that there are well-determined factors (not a large number of factors with only a few indicators each).
According to the criteria proposed by MacCallum (MacCallum et al., 1999), the current Pre-K sample of 138 cases (i.e., 6 observational cycles rated with the CLASS Pre-K per each preschool classroom) – presenting a relatively small number of factors, each one represented by at least three measured variables, and moderate communalities (i.e., M = 0.57, SD = 0.12) – can be considered adequate to perform factor analysis.

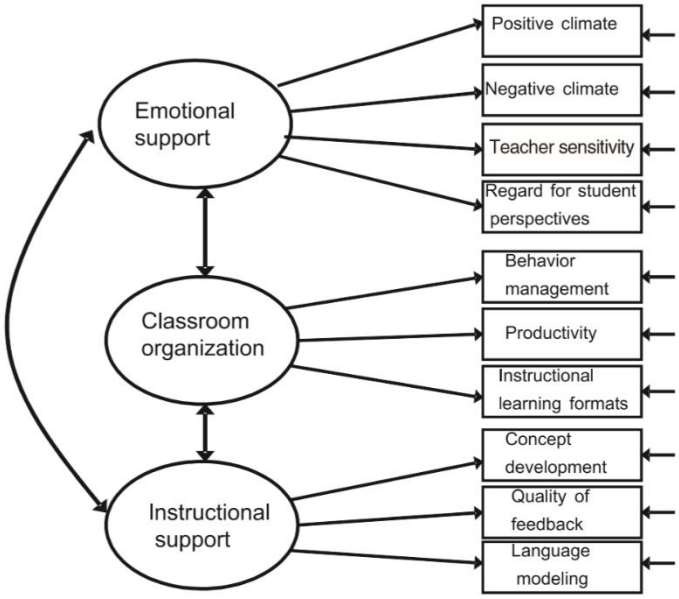


Figure 3.2

The theoretical three-factor model of classroom quality (Hamre et al., 2013)

The following relative goodness-of-fit indices were also used to evaluate model fit:

c. the comparative fit index (CFI);

d. the Tucker–Lewis index (TLI).

Criteria for fit were as follows: RMSEA values ≤ .05 can be considered as a good fit, values between .05 and .08 as an adequate fit, and values between .08 and .10 as a mediocre fit, whereas values > .10 are not acceptable (Browne & Cudeck, 1993; Schermelleh-Engel, Moosbrugger & Müller, 2003).

Values for the SRMR range from zero to 1.0 with well fitting models obtaining values less than .05, whereas values smaller than .10 can be interpreted as acceptable (Schermelleh-Engel, Moosbrugger & Müller, 2003).

For the TLI and CFI, values of .90 or higher are generally taken as indicative of acceptable and values greater than .95 as indicative of good fit (Hu & Bentler, 1999). However, according to Schermelleh-Engel, Moosbrugger & Müller (2003), TLI and CFI values of .97 seem to be more reasonable as an indication of a good model fit than the often-reported cutoff value of .95, whereas values greater than .95 can be interpreted as an acceptable fit.

In the following chapters, the results of the analysis of qualitative (Chapter 4) and quantitative data (Chapter 5) will be presented. An overall integration and comparison of these findings will be provided in Chapter 6.

Chapter 4: Qualitative data

Sometimes to get a better look at what is near to us
we have to move and gain a new perspective.
Giuseppe Mantovani

The qualitative part of the study was aimed, on the one hand, at eliciting CLASS cultural assumptions through a critical discussion of the tool with Italian practitioners; on the other hand, at examining the possibility to integrate a standard-based assessment with the participatory-reflective approach typical of the Italian ECEC pedagogical tradition.

These two aspects will be illustrated respectively in Paragraphs 4.1 and 4.2.

4.1 Discussing the CLASS cultural values and assumptions with Italian practitioners

In order to shed light and address the issues emerging from the literature review in examining the cultural implications of applying this standard-based tool outside of its *cultural cradle* (Pastori & Pagani, *forthcoming*), Italian ECEC teachers and pedagogical coordinators were involved in a discussion not only *on* the instrument, but also *with* it[20].

[20] The qualitative analysis of the implications of the cross-cultural application of the CLASS tool was initiated within the CARE project, based on the data collected from the four Italian 'good practice' ECEC centers selected. Results of this qualitative analysis were presented at international conferences (EECERA – Pastori, Pagani & Mantovani, 2015; EARLI –Pastori, Pagani & Mantovani, 2016) and were included in the working paper '*Study 5: A Cultural Analysis of ECEC Quality in 7 Countries*' published on the CARE website (Pastori & Mantovani, 2016). They will shortly be published with the integration of Portuguese and Dutch data.

In fact, the lens provided by the CLASS allowed them to reflect on – and sometimes become more aware of – their conceptualization of quality in ECEC, recognizing *continuities* between their local-cultural perspective and the one offered by the tool, and, by contrast, points of *disagreements* with the CLASS framework. At the same time, this dialogue enabled teachers to identify key-features of effective relationships not captured by the tool (*missing elements*) or that the instrument invests with different meanings and interpretations from the cultural, scientific, or pedagogical point of view (*differences*).

In the following Paragraphs, the main recurring themes emerging from this dialogue will be presented. For reasons of clarity, the discussion is organized in four sections: continuities, differences, missing elements, and disagreements. However, although these four levels will be examined separately, as the reader will notice, many of the themes illustrated are closely intertwined, often reflecting founding values and principles rooted in the Italian pedagogy.

4.1.1 Continuities

4.1.1.1 The centrality of teacher-child relationship

One of the most innovative aspects that differentiates the CLASS from other assessment tools is the emphasis placed on the key role of teacher-child interactions in defining classroom quality.

The stress on relationship was welcomed by the Italian practitioners interviewed, who considered teacher-child interactions as a crucial factor to define ECEC quality, and therefore judged the observational lens provided by the CLASS as interesting and valuable:

The results presented in this section – integrating and extending the data initially collected for the CARE project by involving a larger sample of Italian ECEC services – are part of a forthcoming paper ('*Is validation always valid? Cross cultural complexities of standard-based instruments migrating out of their context*'– Pastori & Pagani, *forthcoming*) that will be published on the European Early Childhood Education Research Journal.

I think it's a valid thing, that's what we as a school have been saying: the relationship first, and then, in a complementary way, everything else. […] We, as a school, have always tried to put the relationship first, because otherwise everything falls. (Preschool teacher 16)

The relationship is the pivot around which everything else moves. (Preschool teacher 1)

Indeed, in Italy, in order to promote the development of significant and stable relationships, ECEC teachers generally accompany the same group of children throughout the whole period of their attendance (Musatti, 2007). Besides, *pedagogia delle relazioni* (i.e., pedagogy of relationships) – a pedagogy where interpersonal relationships are considered fundamental means for supporting children's social, emotional and cognitive development – is one of the key words in describing the pedagogy of early childhood education in Italy (Mantovani, 2007b; Cochran, 2011):

«We consider relationships to be the fundamental, organizing strategy of our educational system» (Malaguzzi, 1993, p. 10).

4.1.1.2 The dual concept of relationship

Moreover, Italian teachers appreciated the dual and comprehensive concept of relationship that this tool considers, addressing both emotional and instructional features of the classroom:

We fully agree with the idea of not separating learning and care moments, routines, activities, and also [of considering] *the school day as a place of learning and relationship.* (Pedagogical coordinator 2)

This point of agreement with the framework provided by tool was also confirmed by the particular appreciation received by some CLASS dimensions: Positive Climate, Regard for Student Perspectives/Regard for Child

Perspectives, Concept Development/Facilitation of Learning and Development, and Quality of Feedback. On the one hand, these dimensions suggest the importance to offer children a warm, enjoyable and respectful environment (Positive Climate) in which their interests and ideas can be taken into account (Regard for Student Perspectives/Regard for Child Perspectives). On the other hand, they underline that high-quality teaching is focused on the process of learning and on stimulating children's reasoning and thinking rather than on rote instruction (Concept Development/Facilitation of Learning and Development, and Quality of Feedback).

Therefore, they mirror some core-idea at the heart of the Italian pedagogy:

a) the attention for children wellbeing (*pedagogia del benessere*, i.e., pedagogy of wellbeing, Mantovani, 2007b);
b) the concept of an active, constructive, and competent child (Malaguzzi, 1993, 1994, 1998; Mantovani, 2007b; Montessori, 1949, 2004);
c) the importance, also stated in the National Guidelines for preschool (called *Indicazioni*, Ministry of Education, 2012), to consider ECEC services «not only as care environments, as they are also learning contexts: these are two dimensions that are and must be interwoven» (Sarsini & Di Bari, 2015, p.82, translation by the author).

However, despite the appreciation of the overall framework, the teachers involved in the study pointed out nonetheless some dissonances in the conceptualization of ECEC quality and effective teaching provided by the tool.

4.1.2 Differences: the concept of learning

At the content level (methodological differences will be addressed in Paragraph 4.2), a relevant difference regards the concept of children's learning. Both the CLASS Toddler and Pre-K versions feature a specific domain (Engaged Support for Learning and Instructional Support, respectively) that takes into account children's learning and how this process happens in the context of relationships. However, the CLASS seems to convey a

conceptualization of learning that focuses solely on its cognitive and linguistic aspects.

Conversely, Italian practitioners shared a broader vision of what learning is that, mirroring the Italian pedagogical tradition (Mantovani, 2007a, 2007b), embraced also children's socio-emotional development and the role of teachers in fostering it:

> *What do we want to promote? Only cognitive and linguistic development? At least it should be considered social development along with cognitive development… Learning cannot just be cognitive and linguistic!* (Pedagogical coordinator 3)

> *There is also the socio-emotional dimension of learning.* [The CLASS framework] *encompasses only cognitive and linguistic development.* (Pedagogical coordinator 6)

> *Relinquishing this level* [the promotion of socio-emotional learning] *is a missed opportunity. I don't mean a loss, a lack, but a missed opportunity.* (Infant-toddler center teacher 9)

> *Moreover,* [the CLASS] *does not consider learning autonomy.* (Pedagogical coordinator 2)

> *In my opinion, perhaps even the reference to autonomy is missing, because, especially in preschool, we strive a lot for allowing children to learn autonomy.* (Preschool teacher 39)

Therefore, according to Italian practitioners, it is crucial for a high-quality ECEC service to provide children opportunities to learn to cooperate, to be part of a group or a community, to be responsible for others, to regulate their emotions and to understand and recognize those of others, to acquire basic daily-life skills.

However, all these aspects are not included in the CLASS definition of learning. Even the emotional support provided by teachers, although well

developed by the authors and considered a crucial feature to analyse classroom interactions, is never conceived as a learning theme.

Furthermore, Italian teachers considered intercultural and inclusive education as another important aspect of children's learning experience in early childhood educational settings, and pointed out the CLASS failure to take account this theme in its definition of classroom quality:

> *The reference to intercultural skills is completely missing.* (Infant-toddler center teacher 10)

As Vitiello (2013, p. 7-8) points out, «although CLASS scores are correlated with teachers' acceptance of diversity, the CLASS measure does not specifically assess cultural competence, cultural sensitivity, or teaching strategies specific to dual language learners».

However, according to Italian teachers, fostering intercultural competencies should be considered an essential aspect of what children should learn during their ECEC experience. Particularly, teachers should promote an inclusive environment and provide children opportunities to learn to deal with and respect any form of diversity and difference amongst individuals and groups.

This position is supported by recent studies that suggest that, as the society becomes increasingly multicultural, so grows the need to promote children's development of intercultural competencies and that early childhood education plays a pivotal role in establishing this foundation (Barrett, Huber & Reynolds, 2014; Gay, 2002; Perry & Southwell, 2011).

This perspective is also rooted in the Italian long tradition of inclusiveness: Italian ECEC services are conceived since their foundation – well before Italy became a multicultural country – as inclusive and universal, addressed to all children, regardless to their conditions and origins (Canevaro, 2007).

4.1.3 Missing elements

4.1.3.1 The lack of attention paid to the classroom physical environment

Italian teachers mentioned various key-features of the teacher-child relationships not captured by the CLASS.

A first issue deals with the lack of attention paid by the CLASS to the classroom physical environment. This position might seem in conflict with the appreciation expressed for the prominence of relationship. Nonetheless, a more careful look reveals that the contradiction is only apparent.

As described in Chapter 2, the CLASS framework focuses exclusively on process quality and does not assess structural features of classroom, which include the organization and aesthetics of the space (Hamre et al., 2013; Pianta, La Paro & Hamre, 2008).

However, according to the Italian pedagogical tradition (Cochran, 2011; Fortunati & Tognetti, 2003; Gandini, 1995; Malaguzzi, 1993, 1998; Mantovani, 2007a, 2007b; Musatti, 2007; Pontecorvo, 1991), the physical environment serves as scaffolding for the educational processes. The organization and aesthetics of the space plays a key role in sustaining children's learning, self-confidence, independence and socialization, and creates the condition for constructing significant, supportive relationships:

> «The facilities, the choice of materials, and the attractive way in which they are made available constitute an invitation to explore that comes from the environment. Everything is designed with the expectation of creating communications, exchanges between one person and another, and connections between people and things […]. The space is designed and arranged to facilitate meetings and doing [things] together» (Gandini, 1995, p. 236, 237, translation by the author).

> «Thinking the space so that it can be appropriate for the unfolding of children's experience denotes the attention to listening to their needs, and this choice anticipates and even supports the care for relationship and interaction

between the adult and the children within the educational context» (Fortunati & Tognetti, 2003, p. 138, translation by the author).

«The individual lives into the space not only because there are materials, possibilities, dimensions at his disposal […] but also because he is able to use this space and to move within it according to a project. Within it he builds his autonomy, he learns what he can and cannot do; he knows that he can do [things] by himself or along with other children; he knows that he can move and learns to move according to his own autonomy» (Pontecorvo, 1991, p. 74, translation by the author).

In this regard, well-known are the core-concepts of «prepared environment» and «regia educativa» (Montessori, 1912, 1949) and «environment as third teacher» (Malaguzzi, 1998; Gandini, 1995).

From this perspective, the relationships teachers want to promote are *embedded in* the space and materials. Therefore, the concern expressed by the participants involved in the current study is not surprising:

Relationship is also how you structure the space. I mean, if you have in mind a certain kind of teacher-child relationship and where you're headed, when you design the space and decide what to place in it, where to place it and how to place it, in my opinion [in these choices] *the relationship is already taking place.* (Pedagogical coordinator 4)

I have a doubt about the basic choice, which is observing the relationship notwithstanding the context, the creation of the context. In my opinion those things go hand in hand and I don't know if it's correct to separate them. For instance, in our school, the particular attention to the arrangement of the environment is an integral part of the educational project and it certainly affects even the relationship – not only adult-child relationship, but also relationships among peers. Therefore, this basic choice... I'm not saying I'm against it, but it leaves me a bit puzzled. (Preschool teacher 13)

It's clear that it [the CLASS framework] *puts us in a very unnatural and uncomfortable condition, because a tool that explicitly assesses adult-child*

interaction without considering space and materials is in contradiction with our pedagogy. (Pedagogical coordinator 1)

4.1.3.2 A not enough active role assigned to children as resources

A second issue concerns the reciprocal roles of teachers and children.

According to the CLASS framework, effective teachers take into account children's ideas, opinions and interests and support their independence (Regard for Student Perspectives and Regard for Child Perspectives dimensions in the Pre-K and Toddler versions respectively) – a view that is shared also by the Italian pedagogical tradition.

However, according to Italian teachers, the conceptualization provided by the tool does not cover enough a wider idea of children's active role in the school life and in the peer social life (Malaguzzi, 1994; New, 1995). The excessive emphasis placed on the adult results in underestimating the children, who are not fully seen as resources and often seem passive receivers of teacher's interventions. Consequently, even teacher's function in supporting their competences to share and co-construct projects, activities, routines, knowledges, products, and social rules is not recognized and valued:

Children are a resource for each other. There [in the CLASS framework] *the only resource is the teacher.* (Pedagogical coordinator 2)

Actually, it's a cultural difference. I mean, [in the CLASS framework] *there is an adult that manages, that is in charge of organization, expectations, objectives.* (Infant-toddler center teacher 15)

This perspective has implications at various levels. Firstly, it leads to a reductive concept of learning as a top-down process from teachers to children (this aspect will be discussed in more detail in the next Paragraph):

There's a one-sided idea of learning: from teacher to child. (Pedagogical coordinator 2)

In the second place, it results in conceiving behaviour management as an exclusive competence of the adult. Conversely, Italian teachers promote a different approach, a pedagogical guidance of behavior that progressively encourages children to co-construct shared rules, provides chances to discuss about them, their uses and need, and allows children to contribute to enforce them, without over-loading children of responsibility, but encouraging the achievement of a good competence in sustaining each other in respecting rules:

> *Rules are co-established* [with children], *they are not given, they are discussed together.* (Preschool teacher 15)

> *For instance, we ask children: 'we have four toys, there're four of you. What happens? How would you like to use them? What do you think we could do to share and use them together?'. And then dynamics of relationship are built on their answers,* [dynamics] *that they themselves have constructed and that will try to implement.* (Infant-toddler center teacher 15)

Furthermore, Italian teachers judged the CLASS consideration of children's perspectives as limited because it does not include a further level: the possibility for children to contribute with their interests, curiosities, and proposals to actively and significantly build and shape the classroom curriculum:

> [The projects we realized] *are like links in a chain, they're interconnected because they arose from each other.* (Preschool teacher 43)
> [They arose from] *children's curiosities. We didn't establish them beforehand, they were not ends in themselves…* (Preschool teacher 44)

4.1.3.3 The marginal role assigned to peer relationships

Another element that Italian practitioners criticized is closely linked to the previous one, and regards the far too marginal role that the CLASS assigns to peer relationships:

I didn't see the social dimension among children, how the adult supports cooperation among children. (Preschool teacher 14)

[In the CLASS framework] *the adult is the protagonist.* […] *Everything revolves around the adult.* (Pedagogical coordinator 2)

Another thing that puzzles me is that [the CLASS] *considers teacher-child relationship, but the group of children is never taken into account in parallel with teacher-child relationship.* (Pedagogical coordinator 1)
'Child', singular. (Infant-toddler center teacher 14)

Although the tool assesses quality at the level of the classroom, the focus is mainly placed on one-to-one interactions between the teacher and the child:

«The CLASS system centers observer judgments regarding the relative value of teachers' behaviors or interactions toward students on the basis of how individual students react and how teachers respond to individuals» (Pianta & Hamre, 2009a, p. 547).

In this framework, peer relationships are considered nearly exclusively from the socio-emotional point of view (e.g., observing positive connection among peers or children appearing comfortable with one another is considered a manifestation of Positive Climate):

Relationship among peers [is missing]. *I mean, teacher-child relationship is extensively analysed, that is not the case with child-child relationship. […]* [The CLASS] *considers only the respect that children have towards one another, in general, but there are not indicators to encompass how children play by themselves, how the group is organized…* (Preschool teacher 23)

The core of this relevant cultural difference is effectively captured by New (1995, p. 277, translation by the author):

> «Certainly, American preschool teachers recognize the need to give high priority to social relationships among children and to each child's social development. Accordingly, teachers in this country endeavour to assist children becoming part of small groups, encourage their efforts to communicate complex feelings (*tell him how you feel when he takes your truck away*), and give importance to children's pro-social behaviors such as being friendly, cooperative and helpful. However, most of these strategies are designed to cultivate child's social skills in order to promote the individual rather than the group».

Conversely, not featuring a low teacher-children ratio, Italy has developed significant reflections on sociability among peers, learning through peer interactions and on the theme of the double focus of teachers – on the individual and on the group (Malaguzzi, 1993, 1994; Musatti, 2007; Teruggi, 2007). Peer relationships are a cornerstone of Italian pedagogy to such an extent that a pedagogical coordinator even affirmed:

> *It* [the marginal role assigned to peer relationships] *is the cardinal sin. This is what separates us mostly* [from the CLASS framework]. (Pedagogical coordinator 2)

Specifically, according to Italian teachers and pedagogical coordinators, the limited attention paid to peer relationships by the tool results in disregarding two aspects crucial to define quality of teacher-child/children interactions.

First, the CLASS does not highlight the cooperative, social nature of learning, not emphasizing that «[what children learn] emerges in the process of self and social construction [since] children do not passively endure their experience but become active agents in their socialization, co-constructed by their peers» (Rinaldi, 1993, p. 105). Therefore, it does not consider relationships among peers as a key-factor in enhancing children's learning and socio-cognitive development:

> [The group of peers] *can stimulate learning.* (Preschool teacher 19)

The idea of a co-constructed learning is overshadowed. There's always a strong presence of the adult. Reciprocal help or support among peers… is not taken into account. (Preschool teacher 14)

It is probably in the exchange and in what children already know that the most interesting learning occurs. (...) Some theories can be shared and generate new suggestions in other children, perhaps producing new ideas, something that can be built together. (Infant-toddler center teacher 9)

Secondly, the CLASS does not give due emphasis to teachers' intentional action in fostering social interactions, communications and reciprocal support among peers, sustaining a sense of interdependence and of belonging to a community of children:

At times those children that are more capable seem to 'sniff each other', to find each other and band together, leaving more immature peers on the sidelines. Thus, in these situations the duty of the teacher is to try to involve the others, let them interact with each other. (Preschool teacher 39)

I say to the children: 'let's try to ask [to a peer]*'. It's not because I don't want to… but because I think it's more appropriate that everyone, even among peers, doesn't feel lost when there's not the teacher…* [that they] *cooperate to find a solution. Actually, this aspect – the relationship among peers – is not taken into account* [by the CLASS]*. […] Maybe because it's not a direct assessment on the teacher, although… it's a consequence of teacher's constant action, isn't it?* (Preschool teacher 16)

We often divide [the class]*: 5-year olds with some 4-year olds, and 3-year olds with a group of 4-year olds. The group of 4-year olds represents a stimulus for the youngest. In this way, everyone can find their own place, can have a voice.* (Preschool teacher 10)

Due to this sensible and refined – yet often subtle – *regia* of the adult, different viewpoints, disagreements and even conflicts among peers can be intended as part of a productive communication and precious learning

opportunities offered by the group that teacher should seize rather than avoid as negative and undesirable:

> *Maybe what is missing is […] gathering children's expressions and thoughts not to accumulate them, but to call them into question, relaunch them, to create a 'circulartity'* [circolarità] *of ideas […]. We try to do this […] in order to broaden* [children's] *viewpoints and opinions. 'Circularity'* [circolarità] *is one of our goals.* (Preschool teacher 15)

> [In the CLASS framework, teachers] *don't work on 'circularity'* [circolarità] *among children. […] We work a lot on developing arguing skills, explaining your thinking and being able to communicate it to others […] in order to let others understand you and to be able yourself to accept others' viewpoints. In order to realize this 'circularity'* [circuitazione], *a key-factor is the adult's role, the adult is absolutely central.* (Pedagogical coordinator 1)

> *It's up to the adult to generate this process of socio-cognitive 'conflictualization'* [conflittualizzazione socio-cognitiva]. *[…] We work intentionally with children to generate micro-processes of conflict, both at the relational and cognitive level. They actually contribute to enhance processes* [of development]*... precisely because we intentionally work on that at the educational level.* (Pedagogical coordinator 1)

According to New (1995, p. 277, translation by the author), this approach «reflects the high value attributed to children's communicative intentions, and the belief that the exchange within the group promotes concept development; that authorized 'discord' challenges American standards of proper social behavior».

4.1.4 Disagreements

4.1.4.1 The educational value of 'Productivity'

A CLASS dimension that raises several perplexities among Italian teachers was Productivity. This dimension encompasses how well teachers manage instructional time, routines and transitions and provide students with opportunities to be involved in instructional activities (Pianta, La Paro & Hamre, 2008). It does not consider the quality of instruction, but only how efficiently the teacher deals with disruptions and managerial tasks and keeps the children *busy*, offering them something to do throughout the observation.

In particular, one of the indicator included in this dimension, Maximizing Learning Time, suggests that in a highly productive classroom teachers provide clearly defined learning activities for students throughout the day and prevent children waiting due to an excessively slow pacing. As a result, «in this classroom, it would be difficult to imagine more instructional time being squeezed out of the day» (Pianta, La Paro & Hamre, 2008, p. 52):

> *The concept behind* [this dimension] *is an idea of time. In my opinion, it's an idea of life, of humanity. Thus,* [an idea of] *time that must be necessarily productive, assessable.* (Pedagogical coordinator 2)

This notion of time seems ingrained in the American educational culture (Fyfe, 1994; New, 1995): «slowing down seems to run counter to our [American] traditional value and understanding of productivity. Parents and teachers often equate productivity with how many activities are completed in a day, how much of the curriculum is covered in a given period of time» (Fyfe, 1994, p. 24).

Conversely, Italian teachers did not identify themselves in this idea of productivity and questioned the educational value of having children be engaged in constantly doing something:

> *Productivity. Maybe I'd think twice about that. I mean, it's true that you have to maximize the time and everything... it's also true that you have to be very*

relaxed. […] Even doing nothing for the children is still doing something. (Preschool teacher 16)

It is important to highlight that this criticism did not emerge only among preschool teachers, but also among teachers from infant-toddler centers. This remark is particularly interesting because, unlike the Pre-K version, in the CLASS Toddler the term *productivity* is never explicitly mentioned, suggesting that this dimension may be an implicit feature of teacher-child interactions underpinning the CLASS framework:

The allusion to productivity struck me because it seems that if you're busy and over-stimulated, everything's fine; if you just stop, you get a low score. It struck me that it's not given a value in terms of development to a stop-phase, which perhaps is only apparent… we often see children who perhaps in front of our proposal seem not involved. Before redirecting them, we try to figure out if they are observing or doing something that will let them actively participate later. This underestimation of downtimes – that are never dead times – struck me. (Infant-toddler center teachet 12)

In Italy, early childhood years are considered as «a very precious time in one's life, a time that should be tasted, explored, and experienced without haste. 'Where is the hurry?' is a question posed in early childhood pedagogy in Italy today» (Mantovani, 2007b, p. 1118). Therefore, Italian teachers encourage children to experience a relaxed time, with opportunities to think and fantasize, in the belief that imposing «adult time on children's time […] negates children being able to work with their own resources» (Malaguzzi, 1994, p. 55). Even *wasting* a bit of time is allowed as long as it contributes to provide a holistic rich experience, free of pressure:

«The conviction is that the experience in settings especially designed for children, rich and free of pressure, is by no means a waste of time. It is rather an important training ground for consolidation of the sense of self, of social competence, of exploration and of research attitudes» (Mantovani, 2007a, p. 1114).

Italian teachers' care not «to fall in the trap of productivity» (Staccioli, 2008, p. 381, translation by the author) mirrors and is supported by peculiar characteristics of the Italian ECEC services. In Italy, children spend on average six to nine hours per day, five days a week in the service. These conditions allow teachers to rely on an extended amount of time (especially compared to other countries where ECEC programs are generally part-time – e.g., the Netherlands or the UK) to interact and work with children, and consent projects and activities to unfold over a period of days, weeks, sometimes even months.

Therefore, ECEC services are conceived as a *daily life context* shared by children and teachers, and the priority, rather than focusing on the *program of the service*, is to ensure children a high-quality *life in the service*, exposing them to slower – albeit not less meaningful – rhythms than their lives outside of school (Mantovani, 2007b; Cosmai & Caggio, 2008; Staccioli, 2008).

The idea of providing children rich, meaningful activities to be experienced without haste leads to a second perplexity regarding the dimension Productivity:

> *I feel a strong resistance towards this idea: [...] not considering the significance* [of the activity], *not considering that the teacher should ask himself what kind of experience he's offering to children.* (Pedagogical coordinator 2)

> [The focus is] *on doing something, on doing things over time, rather than on constructing meanings reasoning on things. […] It's just doing, it's a product, rather than a meaning.* (Pedagogical coordinator 3)

According to Italian practitioners, offering something to do all the time, regardless of its educational quality, is pointless if the activities provided are not significant for the children. This idea is clearly explained by Staccioli (2008, p. 384, translation by the author):

> «Learning comes from *real*, shared, engaging activities, activities that have a meaning and that meet profound needs. In this way, learning become stronger, times are enlivened. 'Dead times' in preschool are not *routines*, rest or doing nothing. Dead times are those of hasty and directive learning, of activities

planned by default according only to adults' goals. […] Children need *not to hurry*, at least at school».

Furthermore, a third criticism in the concept of productivity regards the meaning attached to waiting. In contrast to the perspective suggested by the CLASS that emphasizes the importance of having children busy and interprets a slow pacing in negative terms, Italian teachers recognized that even waiting can assume an educational value for children, offering them opportunities to learn to tolerate delay and turn-taking, and above all to respect their peers:

In my opinion, at times it's right for children to experience situations in which they do not necessarily have an active role, but have to wait, because… because in life you cannot always be busy. […] It's right to propose even these moments, albeit seemingly boring, provided that they don't take a too long time. I mean, in small doses, children also have to get used to tolerate waiting times. (Preschool teacher 16)

There is also an educational value in learning to wait. [For instance,] *calling attendance: it's clear that it's a repetitive activity, it's the same every day for the whole year, there are many* [children]*... but at least it helps children to learn self-regulation, waiting* [their turn] *in silence, even a sense of respect for others…* (Preschool teacher 19)

[It helps] *to develop listening skills.* (Preschool teacher 17)

4.1.4.2 When non-intervention is effective as well

A second point of disagreement is cross-dimensional. It concerns the definition of what strategies can be considered effective in interacting with children, and specifically when *non-intervention* can be considered a good practice.

A first aspect of this issue regards indicators related to management of unacceptable behaviours or conflicts among children.

According to the CLASS, in high-quality classrooms, the adult should manage these behaviours in an effective and timely manner. The CLASS does

not contemplate the possibility that the teacher lets children find their own solution to a conflict, refraining from intervening even for a long time (or at least until there is a real escalation), or that other children mediate or enforce the rules (see Paragraph 4.1.3.3). It does not consider these modes of behaviour management to be effective strategies, and would not take into account these dimensions, while it values the quick intervention of the teacher or even the prevention of the conflict.

This interpretation seems to mirror the American educational culture, but does not necessarily reflect the pedagogies typical of other countries. For instance, the logic of *non-intervention* facing misbehaviours or even a conflict among peers is widespread in educational services for Japanese children (Tobin, Wu & Davidson, 1989; Tobin, Hsueh & Karasawa, 2009), and also partly in Italy:

> *Dealing with conflict management, […] promptness isn't always a quality indicator*. (Pedagogical coordinator 6)

> *We keep coming back to the idea of teachers, children and their relationship. Conflict: how do I manage it? If I have an idea of damping, of teacher's modelling, I'll act in a way. If I think that children are a resource for each other, then I'll manage conflict in different way. […] Considering conflict management in terms of adult's timely intervention – there's a conflict, I try to prevent it, I act promptly to resolve it – for a culture like the Italian one, you know, is very reductive. Because actually it's all about the climate you've created before, how you construct spaces for debate among children,* [spaces] *for dialogue, negotiation, respect, valorisation of differences, how you divide a group – that maybe is a bit complex and conflictual – in smaller groups. You can deploy many dimensions to prevent, to anticipate a conflict. Then, if you have to resolve it, the cultural dimension comes into play: I would never consider promptness as a quality indicator […]. That doesn't mean that I'm not aware, that I don't care at all. It means that I allow time to children's autonomy, to let them find their own solution…* (Pedagogical coordinator 2)
> [It means] *also believing that children are competent and can do that.* (Preschool teacher 15)

At times you notice that something has happened, but they [the children] *have to fix the problem themselves. Thus, you keep a watchful eye on them, observing how making peace goes. If everything goes smoothly, that's fine. Otherwise, you try to intervene.* (Preschool teacher 28)

A similar consideration applies also to the Teachers Sensitivity dimension. The CLASS suggests that highly sensitive teachers should be aware and respond quickly when students send behavioural signals indicating a need for help or attention, providing support in a timely manner (Pianta, La Paro & Hamre, 2008). However, Italian teachers challenged this conceptualization of sensitivity. They agreed that a good teacher has to be aware of children's needs and concerns; nevertheless, they thought that this awareness should not necessarily result in the adult's timely intervention:

This approach [non-intervention] *isn't wrong, because you don't have to intervene immediately… what's important is that you have observed the child.* […] *Many times when we see that a child is in trouble, we do not intervene waiting for him to overcome his difficulties or to ask for help from a peer, precisely because it's right to allow* [children] *handle challenges on their own and not always call 'teacher, teacher, teacher'. Therefore, a teacher could notice* [a child in trouble] *and decide intentionally not to intervene.* […] *Dealing with some children you have to adopt this strategy* [non-intervention] *because they constantly request your help, therefore at that point you have to ignore them because they have to learn to do things by themselves. I mean, the adult has to progressively step aside and leave more room to the child.* (Preschool teacher 16)

Perhaps it should be observed more carefully how… not whether the adult intervenes promptly as soon as a problem occurs, but how the adult… I mean, you can see if the adult is observing the child who, for instance, is trying to pull up the shade and can't do that but [the adult] *is nonetheless observing him, or if* [the adult] *is paying attention to something else, or if* [the adult] *despite noticing the child ignores him or rolls her eyes.* (Preschool teacher 18)

Non-intervention strategies can be appropriate in some circumstances even in regard with language facilitation. The CLASS seems to suggest that effective teachers have to play a highly active role in stimulating children's linguistic development (e.g., often repeating or extending their comments, describing with words adult's and children's actions…). However, Italian practitioners recognized also the value of silence, as a mean to provide time and space both for the child and for the teacher to think and reflect without being overwhelmed by too many words – welcoming Montessori's caveat quoting Dante Alighieri: «Let thy words be counted» (1912, p. 108):

> *Sometimes we remain silent for… maybe not for several minutes, but… in order to take our time to figure out if it's the time for talking or the time to be silent.* [The CLASS] *doesn't take into account that, it doesn't take into account that… because maybe in that fraction of a second you want to rethink about the possible thoughts the child is having, his expectations…* (Infant-toddler center teacher 9)

Italian teachers' viewpoints would introduce a new dimension, which could be defined 'Effective strategies of non-intervention', to recognize the value and encompass the possibility for the teacher to take time to observe the child and the peers' reaction before deciding if and how intervening:

> *At times if you intervene immediately, you can't understand what is the child's real problem.* (Preschool teacher 16)

> *During some activities, you feel less the urge to intervene […] and then it depends on what are your goals: during some activities, you decide in advance that your aim is to observe* [children]. (Infant-toddler center teacher 5)

> *The teacher becomes more an observer, a mediator where his presence is requested. The teacher, at least in preschool, does not generally intervene if children are able to interact spontaneously with each other, to make comments and even be a stimulus for each other. Thus, it's unlikely that the teacher interferes in these interactions, except when a problem arises and children request teacher's intervention.* (Preschool teacher 39)

Besides, echoes of Loris Malaguzzi's and Maria Montessori's pedagogy are recognizable in this strategy:

> «We need to know how to recognize a new presence, how to wait for the child. This is something that is learned, it is not automatic. We often have to do it against our own rush to work in our own way. We will discover that our presence, which has to be visible and warm, makes it possible for us to try to get inside the child and what that child is doing. And this may seem to be passive, but it is really a very strong activity on our part. […] If it [the interaction between children] is left to ferment without adult interference and without that excessive assistance that we sometimes give, then it's more advantageous to the child. We don't want to protect something that doesn't need to be protected» (Malaguzzi, 1994, p. 54, 55).

> «To stimulate life, leaving it free, however, to unfold itself, that is the first duty of the educator. For such a delicate mission great art is required to suggest the right moment and to limit intervention, last one should disturb or lead astray rather than help the soul which is coming to life and which will live by virtue of its own efforts» (Montessori, 2004, p. 141).

The CLASS does not consider the approach proposed by Italian practitioners as an example of effective teaching, and would not take into account (or even would give low scores) to these alternative strategies – 'staying back', 'being silent', providing to children times and spaces to observe, think, develop their own ideas, activities, both by themselves and in collaboration with peers.

4.2 Discussing the CLASS evaluation approach with Italian practitioners

Besides the exploration of the cultural assumptions underpinning the CLASS framework, the discussion on the tool addressed also the methodological level, examining its approach to evaluation and assessment.

A first aspect that Italian practitioners pointed out concerned some dissonances they perceived at this level.

4.2.1 Methodological dissonances

4.2.1.1 Unfamiliarity of a standard-based approach

As discussed in Chapter 1, in Italian ECEC services a strong evaluation culture is not fully established yet, and a subtle suspicion towards evaluation still persists especially when it relies on assessment instruments not negotiated and attentive to local characteristics (Calzolari, D'Ugo & Vannini, 2010; Savio, 2015).

Therefore, not surprisingly, a certain apprehension was the first reaction of some teachers to the CLASS, a *foreign* standard-based tool little – or not at all – known:

> *This kind of assessment is a little scary.* (Preschool teacher 11)

Specifically, some teachers reported their limited appreciation of its quantitative, highly structured evaluation apparatus:

> *I don't like rating scales.* (Preschool teacher 22)

> *I don't really like assigning numerical ratings. I mean, this kind of evaluation… I don't like it. I don't' like this need to pigeonhole things with numbers or letters…* (Preschool teacher 16)

The limited diffusion of standard-based tools that characterizes Italian ECEC evaluation culture is responsible – at least partly – for this attitude, contributing to sharpen the perception of unfamiliarity and distance towards the assessment approach proposed by the CLASS:

> [It's] *a bit aloof, a bit rigid.* (Preschool teacher 19)

This numerical approach to evaluation… it is indeed true that after all somehow you have to evaluate, but maybe it can be done in a way less… (Preschool teacher 16)

Less rigid… […] we are less technical and more humanistic. […] Besides, it's a structured tool, thus a certain degree of quantification is needed. (Preschool teacher 19)

4.2.2.1 The observational procedure

However, the main dissonance perceived at methodological level deals with the observational procedure proposed by the CLASS and its implications.

As presented in Paragraph 2.2.2, the CLASS requests to conduct 20-minute cycles of observation followed by a 10-minute period for recording codes. A minimum of four up to a maximum of six observational cycles should be obtained throughout a school day in order to «provide an adequate sampling of classroom quality» (Pianta, La Paro & Hamre, 2008, p. 96). Moreover, in order to guarantee objectivity, the observer should avoid «injecting external explanations for what the observer sees taking place» (Pianta, La Paro & Hamre, 2008, p. 12), focusing exclusively on the visible behaviour.

These requirements significantly diverge from the principles underlying the participatory-reflective approach typical of Italian early childhood pedagogical tradition.

A first issue relates to the limited time allowed for coding:

To analyse the video-clip we should have watched it more than once, and then maybe watch again segments [of particular interest]. (Preschool teacher 16)

On the one hand, Italian teachers are not familiar with standardized instruments and with observational approaches that invite to assign ratings rapidly. On the other hand, the time limitation for coding, which from the CLASS perspective should contribute to preserve reliability (since observers' overthinking could undermine their objectivity), is not necessarily considered a value. Conceiving quality as a recursive process that develops in a spiral course

(Bondioli, 2002), according to Italian teachers repeated observations of the same data could foster the emergence of new insights and reflections, drawing the attention to significant elements unnoticed upon the initial observation.

Alongside this, what Italian practitioners particularly criticized was an idea of evaluation that does not take into account – actually, demands ignoring – the local context peculiarities, the shared pedagogical principles of the service underpinning teachers' choices, the meanings and intentions behind teachers' behaviors and actions, and the pre-existing relational stories with children:

> [The CLASS] *loses sight of the context. Whenever someone else watched our videos, I felt compelled to contextualize, to make explicit* [what might be implicit]*, to do justice to everything* [the video-clip] *did not show, because there are reflections behind* [what you can see in the video-clip] *that go beyond, that have been generated over time. […] Everything we do is closely linked to an experience that unfolds over time. Thus, assessing that twenty-minute segment becomes complex, difficult, there's always something missing. […] I feel a resistance* [towards the tool]*, it makes me put blinders on. I'm always temped to say: 'maybe the teacher acted in that way because…'. […] What is missing is the history, the history resulting from an ongoing path of growth.* (Preschool teacher 15)

This need of contextualization is deeply rooted in the Italian definition of ECEC quality, intended as a locally shaped, rather than standardized, concept:

> *Any evaluation process is a contextual process* [ingrained] *in your culture of school and education, in that educational project, in that particular nido, in that particular preschool, in that teacher, in that educator who works with that group of children. The evaluation process in our perspective [...] is, by definition, a contextual process.* (Pedagogical coordinator 2)

This idea that relationships unfold and develop dynamically over time led Italian teachers to question the CLASS claim of capturing relational quality through the observation of few hours of a single school day, and to feel the need to extend the observation over a prolonged period of time:

Something is missing. Something significant is missing. You get only a snippet in those twenty minutes. (Preschool teacher 15)

Certainly, it's difficult to capture [relational quality] *in a flash […]. It's difficult to assess what lies behind a such brief moment. Perhaps* [the observation] *should be repeated over time, over situations. Then the assessment could be more objective.* (Preschool teacher 19)

In my opinion, the observation time should be extended. One thing is to observe over two or three days, another thing is to observe for a month, because this way you can see the continuity of action and everything else. Otherwise, taking just a brief moment… (Preschool teacher 41)
[It] *is just a snapshot, a moment…* (Preschool teacher 42)
It's biased… (Preschool teacher 23)
You see, it's a bit distorted, instead if you encompass the continuity… (Preschool teacher 42)

[Teacher-child] *relationship is dynamic. How does teacher-child relationship evolve? I mean, is it possible to observe it in the same way today and tomorrow? Does* [the CLASS] *take into account these changing aspects of children and even of the relationship itself?* (Infant-toddler center teacher 15)

The message that Italian teachers' voices seem to convey is that assessing teacher-child relationship as a decontextualized object – deprived of its history, meanings and dynamic nature – is liable to undermine evaluation validity, not taking into account the complexity of the *evaluandum*:

It [the CLASS] *doesn't encompass the complexity of the situation. Throwing overboard all the systemic reflection, the complexity of the situation seems to emerge from the sum of different indicators, of singular specific episodes. But, actually, it's more than this, it's about the dynamic…* (Pedagogical coordinator 2)

It should be stressed that the recognition of this contextual and temporal dimension, although disregarded by the CLASS, was present in Robert Pianta's earlier works:

> «Relationships have a history, a memory; they are patterns of interactions, expectations, beliefs and affects organized at a level more abstract than observable behaviors. That is why when one wants to observe properties of a relationship, they must be observed over time, over situations, and from multiple windows» (Pianta, 1997, p. 14).

4.2.2 Interesting and valuable elements

As the previous Paragraphs have highlighted, Italian teachers did not ignore the limits and unavoidable dissonances that the transposition of the CLASS to the Italian context may entail:

> *I had a doubt about this tool and it still remains: the tool is definitely objective, but in my opinion there should be other instruments to read to context, the history. It definitely offers a different perspective than other tools that have a diverse focus. In my idea of school, maybe my Italian idea, I think we should have a broader perspective.* (Preschool teacher 3)

> *The tool should be developed further.* (Preschool teacher 41)

> *We are not that similar to the Americans. […] Thus, we have to be careful before accepting it* [the CLASS framework]. (Preschool teacher 16)

Nonetheless, they recognized that even a tool founded on cultural and methodological assumptions often distant from those underpinning Italian pedagogical tradition such as this can offer opportunities to foster teachers' professional development:

> *It is clear that* [the gap with the CLASS framework] *makes us experience dissonances, a sense of distance, difficulties. However, this matter of fact shall*

not preclude us from considering the formative opportunities that this tool can offer. (Pedagogical coordinator 2)

4.2.2.1 A systematic frame to observe teacher-child relationship

Firstly, the CLASS offers a systematic frame, fruit of a solid theoretical and empirical research, to assess quality in ECEC, providing categories and specific behaviours to observe.

Despite being quite unfamiliar to Italian practitioners – or perhaps especially for this reason, this systematic and structured approach can represent a map, a compass useful to train and guide the gaze in observing educational practices and teacher-child relationship, whereas the holistic perspective typical of the Italian ECEC culture – attentive to embrace systemic complexity – at times may result quite dispersive, offering less defined benchmarks to compare with:

They [the Americans] *have a tradition of structured instruments. In our humanistic approach, we… […] perhaps we lack a systematization.* (Preschool teacher 19)

Actually all those areas are familiar. I mean, I think that we all keep in mind these goals in our daily educational practice. However, actually it's the first time that I see them so categorized, so well-sorted, even with a reference to a behavioural range – low, mid, or high – in which you can identify yourself. […] It's the first time that I see them organized in such a specific and systematic way. (Infant-toddler center teacher 7)

[The CLASS] *forces us to a process – which we are not used to, but that sometimes can be useful in professional development and self-evaluation experiences – of deconstructing educational relationships – which are a complex, systemic phenomenon – in their components. It's clear that this process is unnatural to us, but that's precisely why it can be interesting. […] It allows you to widen your gaze and can highlight criticalities regarding the group experience, the teacher-child relationship. […] Sometimes using a tool like this to conduct observations can be interesting, along with other*

instruments, of course. In the right dose, I think it's an interesting tool. Of course, like all the tools that come from that culture and are developed in a research context, it has a certain degree of obsessiveness and distance compared to our contextual approach… but I think that in the right dose it's nonetheless interesting. (Pedagogical coordinator 1)

An analytic tool can be useful especially in a culture like ours that tends to focus on complexity. (Pedagogical coordinator 2)

It's interesting because it meticulously plumbs all the various domains […]. It has immediately arisen my interest because it meticulously squeezes out things that maybe… it makes you aware of valid, interesting elements. (Preschool teacher 12)

Not surprisingly, the CLASS systematic frame was welcomed as a useful support to observation especially in those services where teachers were less or not at all used to observe themselves and their practices:

It's a stimulus to reflect on what you do in a positive way. We don't usually think about it, we act and that's all. (Preschool teacher 41)

Inevitably, it makes you reflect on your practices. […] You can notice… (Preschool teacher 16)
…many things that actually you miss. (Preschool teacher 22)
We should videotape and then observe [our practices] *more often, because it makes you notice children's aspects that otherwise you miss…* (Preschool teacher 16)

It makes you face some behaviours or strategies that you didn't know you implemented and makes you say: 'geez, but if in that situation I had done…' or 'geez, I do the same!'. […] It makes you think about your behaviour, your practices, how you intervene, your educational style, your shortcomings, your limits. (Preschool teacher 12)

4.2.2.2 A stimulus to reflectivity

In the second place, the tool can represent a stimulus to enhance teachers' reflective processes of self-observation and self-evaluation.

On the one hand, it can introduce a new point of view to look at teachers' everyday educational practices, and even to question consolidated strategies to interact with children:

> *It's a great stimulus.* (Preschool teacher 41)
> *It makes you reflect on many things.* (Preschool teacher 42)

> *It makes you reflect on yourself, on what you can actually do and what you can't do, on how you do it.* (Preschool teacher 4)

> [It's an opportunity] *to question yourself together with your colleagues. […] This categorization is a bit scary, but perhaps it could make you feel the desire to examine yourself.* (Preschool teacher 3)

> [It's useful] *to keep from falling into habits, from going through the motions.* (Preschool teacher 6)

> *In my opinion, the opportunity to discuss, to watch* [your videos] *is really formative, because it makes you reflect, it makes you grow professionally, it makes you improve, definitely. This numerical evaluation… […] I prefer the discussion that makes you grow professionally.* (Preschool teacher 16)
> *The opportunity to discuss.* (Preschool teacher 17)

> *A thing that I consider really valid of this experience is that it is an opportunity to reflect on our educational practices. […] We are all quite aged, we are all in our fifties or older. And that says a lot, because in words we can do everything, we are the best. In practice, we can fall into ineffective behaviours, without intent, but because, since we have been teachers for a long time and we have had many experiences, we sometimes fall into presuming to know and to recognize those aspects regarding* [teacher-child] *relationship.* (Preschool teacher 1)

For instance, one of the avenues of awareness and reflection opened up by the feedback provided by the tool concerns the CLASS Pre-K dimension Concept Development featuring in the CLASS Pre-K and the dimension Quality – present in both the CLASS Pre-K and Toddler versions. These dimensions, as pointed out in Paragraph 4.1.1.2, were among the most appreciated within the CLASS framework, but they reached only modest levels in many classrooms from both infant-toddler centers and preschools (see Chapter 5). Italian teachers agreed with this picture taken through the CLASS lens and recognized these dimensions as an area of improvement. Specifically, they valued the idea, suggested by the tool, that interventions aimed at expanding children's learning, understanding, and persistence could be provided at any time of school day and during any activity since they represent just another manifestation of teacher-child relationship, rather than being 'confined' to a limited number of activities specifically designed to stimulate children's reasoning:

> *I always insist on learning, on how children learn, on how adults support* [learning] *[…] because often in infant-toddler centers partly due to children's age, partly due to difficulties with the space, partly due to the history of the nido… they* [the teachers] *are very competent, very attentive to relationships with parents and with children in terms of listening and so on… but at the methodological level […] we still have to work on how supporting children's cognitive development within and through* [teacher-child] *relationship.* (Pedagogical coordinator 4)

> *Watching the video-clip, I noticed that we didn't stimulate a lot children's reasoning.* [We asked children:] *'Did you like it? Didn't you like it?', that's all.* (Preschool teacher 41)
>
> *It's a stimulus for us to focus more on this aspect. (…) It's a stimulus, because in every activity we can foster this aspect. At times, we are a bit too expeditious, hasty in finishing an activity. Instead, those are important aspects, opportunities – both for the children and for us – waiting to be seized.* (Preschool teacher 42)

On the other hand, even unfamiliar, contentious issues may represent an opportunity to make explicit and become more aware of their own pedagogical values, interpretations and conceptualizations of quality:

> *It is structured in a way in which we cannot recognize ourselves completely. The way it's structured takes itself a cultural stand. […] That doesn't change the fact that the tool can be useful… despite… no, especially since the tool allows you to make these cultural comparisons.* (Infant-toddler center teacher 15)

> [The comparison with the tool], *even by contrast, is a formative process very interesting and really enriching. Actually, all the dissonances you feel can help you to orient yourself better. Then, we can take those indicators and use them according to our idea of relationship. […] Taking some parts* [of the tool] – *even in an unsystematic, unorthodox way – can contribute to foster our processes of observation and evaluation.* (Pedagogical coordinator 1)

Thus, in this perspective of the dialogue *on* and especially *with* the tool, the CLASS, rather than becoming a standard to attain to, a predefined recipe that establishes what effective teachers do and how, becomes an occasion to discuss on and question teacher's educational practices, a *trigger* to foster teachers' self-reflection and, accordingly, their professional growth.

Therefore, Italian practitioners' voices seem to indicate that even a standard-based approach to evaluation may be not only compatible, but also fruitfully integrated with the participatory-reflective perspective typical of the Italian early childhood pedagogical tradition.

Chapter 5: Quantitative results

I am no poet. I do not love words for the sake of words.
I love words for what they can accomplish.
Similarly, I am no arithmetician.
Numbers that speak only of numbers are of little interest to me.
Patrick Rothfuss – The Wise Man's Fear

In order to answer to the research questions postulated, the qualitative critical cultural exploration of the CLASS tool presented in Chapter 4 was complemented with a quantitative analysis to examine, even at statistical level, the applicability and generalizability of the Teaching through Interaction framework to the Italian ECEC context.

This chapter will present the descriptive results of the application of the CLASS Pre-K and Toddler to the Italian ECEC services, and the factor analysis conducted on the preschool sample.

5.1 The application of the CLASS Pre-K in the Italian preschools

5.1.1 Descriptive statistics and comparisons with other countries

The first step of the quantitative analysis consisted in examining the descriptive statistics of the CLASS Pre-K domains and dimensions and comparing them with results reported at international level.

The means, standard deviations and score distribution ranges for the individual CLASS Pre-K dimensions in Italian sample are shown in Table 5.1 and Figure 5.1.

Table 5.1

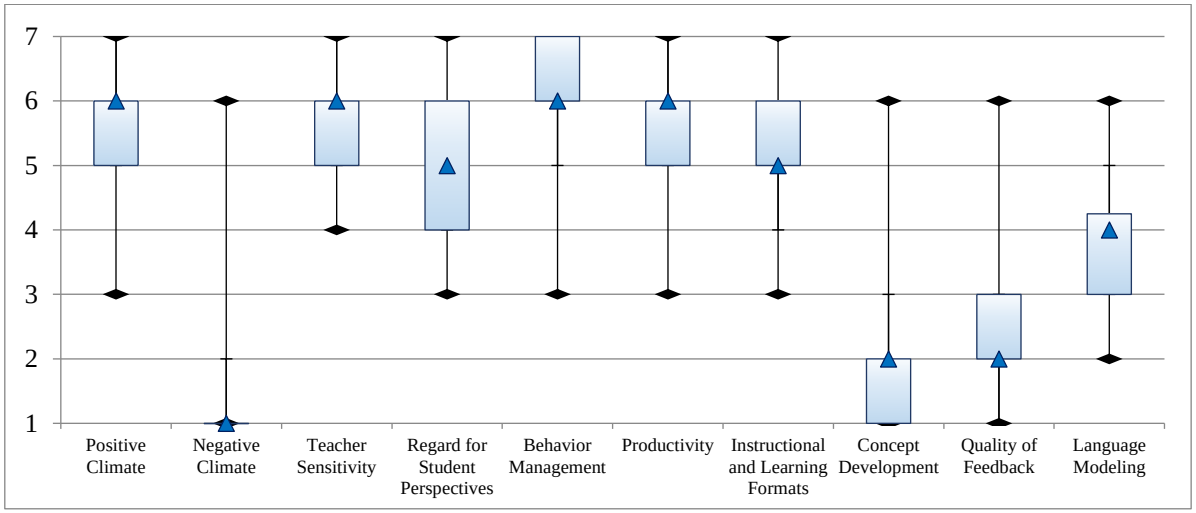
Means, standard deviations and ranges for teacher-child interactions in preschool classrooms

Domain and Dimension	**M**	**SD**	**Min**	**Max**
Emotional Support	**5.77**	**0.60**		
Positive Climate	5.63	0.84	3	7
Negative Climate	1.26	0.64	1	6
(Reversed Negative Climate)	(6.74)			
Teacher Sensitivity	5.63	0.72	4	7
Regard for Student Perspectives	5.09	1.07	3	7
Classroom Organization	**5.63**	**0.71**		
Behavior Management	5.95	0.90	3	7
Productivity	5.72	0.85	3	7
Instructional Learning Formats	5.23	0.88	3	7
Instructional Support	**2.71**	**0.65**		
Concept Development	1.94	0.80	1	6
Quality of Feedback	2.41	0.65	1	6
Language Modeling	3.80	0.96	2	6

Note. Each scale ranges from 1 to 7 points.

Figure 5.1

Box plot and score distribution ranges for teacher-child interactions in preschool classrooms



Note. The top and bottom of each rectangular box denote the 75th and 25th percentiles, respectively, with the triangle inside the boxes showing the median. Vertical bars extending from each box represent the extreme values.

At the domain level, the overall level of Emotional Support (M= 5.77, SD = 0.60) and Classroom Organization (M = 5.63, SD = 0.71) was moderately high, with five dimensions in the middle to high range (scoring between 5 and 6) and one dimension (Negative Climate reversed score) in the high range.

In contrast, the overall level of Instructional Support was rather low (M = 2.71, SD = 0.65), with two dimensions (Concept Development and Quality of Feedback) in the low range and one dimension (Language Modeling) in the middle range.

The standard deviations for the dimensions ranged from 0.64 to 1.07, and for most dimensions were approximately one scale point. The two dimensions registering the lowest variation were Negative Climate (SD = 0.64) and Quality of Feedback (SD = 0.65). However, in the specific case of Negative Climate, this value alone could provide a partially biased picture. In fact, as Figure 5.1 effectively illustrates, despite the score distribution ranging between 1 and 6, overall the ratings for Negative Climate were very low (only one observational cycle received a rating higher than 3 and the majority of the video segments – 80.3% – were scored as 1). This result suggests that this dimension poorly

differentiated classroom quality among Italian preschool classrooms, as already pointed out by Pakarinen and colleagues (2010) with regard to the Finnish sample.

The general picture of classroom quality depicted by the CLASS seems to suggest that: a) Italian preschool teachers had generally positive, warm, and supportive interactions with children; b) moderate language stimulation and facilitation were provided; c) classroom were rather effectively organized in terms of behavior and instructional time management. Conversely, Concept Development and Quality of Feedback means in the low range seem to suggest that, for the most part, the observed activities focused more on basic skills than promoting children's reasoning and thinking.

This pattern – higher levels of Emotional Support and Classroom Organization, with lower scores of Instructional Support dimensions – is similar to those reported in previous studies conducted at international level (see Table 5.2 and Figure 5.2).

However, as Figure 5.2 shows, it should be stressed that, compared with the USA, Finnish, German, Chilean and Chinese samples, Italian preschool classrooms registered higher scores in all the six dimensions of Emotional Support and Classroom Organization. With regard to Instructional Support, Quality of Feedback and Concept Development ratings were rather low and resulted higher only compared to those from the Chilean and Chinese samples; whereas Language Modeling scores were lower only than the Finnish ones.

Table 5.2

Comparison between the Italian preschool sample and samples from five other countries. Descriptive statistics

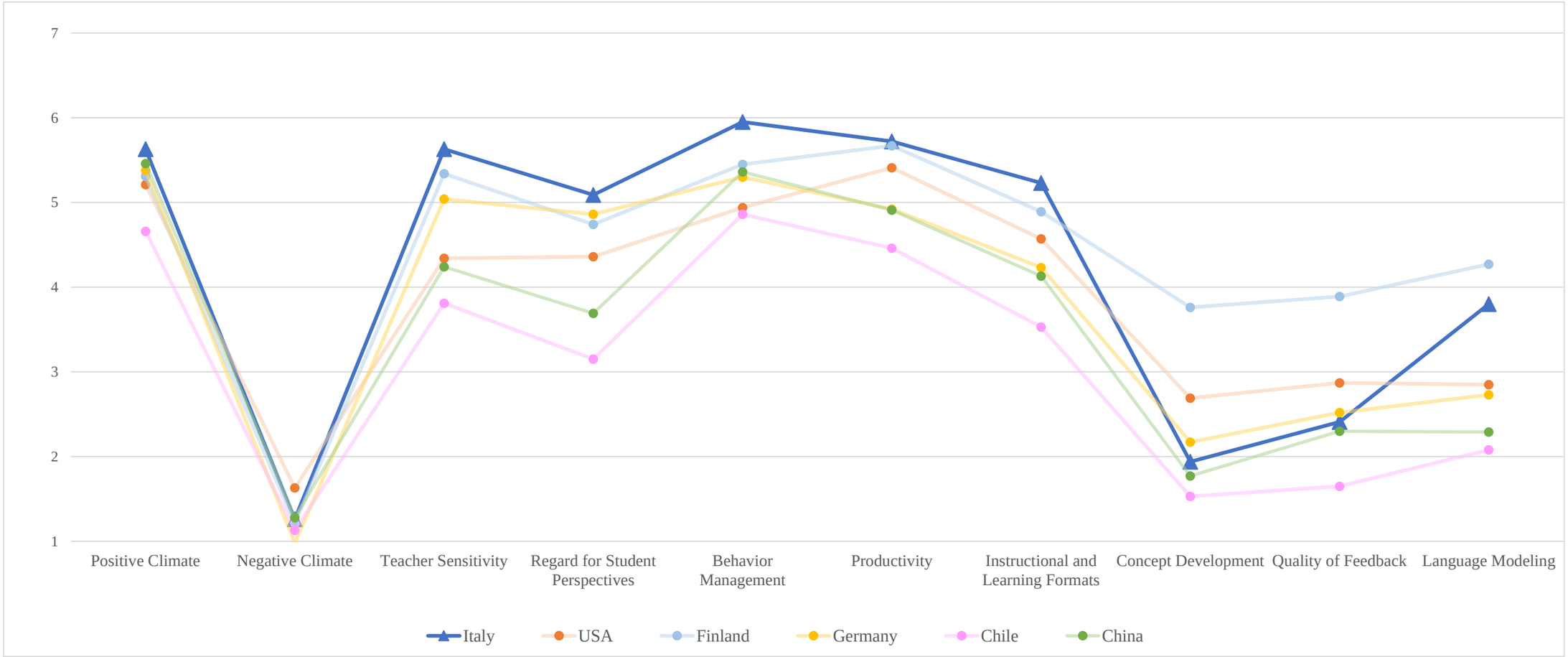
Dimension/Domain	Italy (N = 23)	USA (N = 164)	Finland (N = 49)	Germany (N = 63)	Chile (N = 91)	China (N = 180)
Dimension						
Positive Climate	5.63 (0.84)	5.21 (0.90)	5.31 (0.83)	5.38 (0.95)	4.66 (0.71)	5.46 (1.02)
Negative Climate	1.26 (0.64)	1.63 (0.69)	1.21 (0.38)	1.98 (0.27)[a]	1.13 (0.21)	1.28 (0.43)
Teacher Sensitivity	5.63 (0.72)	4.34 (0.94)	5.34 (0.74)	5.04 (0.89)	3.81 (0.91)	4.24 (1.00)
Regard For Student Perspectives	5.09 (1.07)	4.36 (0.97)	4.74 (0.82)	4.86 (0.90)	3.15 (0.86)	3.69 (0.96)
Behavior Management	5.95 (0.90)	4.94 (0.88)	5.45 (0.85)	5.30 (0.99)	4.86 (0.78)	5.36 (0.94)
Productivity	5.72 (0.85)	5.41 (0.82)	5.67 (0.45)	4.92 (1.10)	4.46 (0.77)	4.91 (0.99)
Instructional Learning Formats	5.23 (0.88)	4.57 (0.78)	4.89 (0.67)	4.23 (0.97)	3.53 (0.74)	4.13 (0.86)
Concept Development	1.94 (0.80)	2.69 (0.68)	3.76 (0.85)	2.17 (0.78)	1.53 (0.54)	1.77 (0.59)
Quality of Feedback	2.41 (0.65)	2.87 (0.85)	3.89 (1.04)	2.52 (0.81)	1.65 (0.60)	2.30 (0.74)
Language Modeling	3.80 (0.96)	2.85 (0.73)	4.27 (0.86)	2.73 (0.76)	2.08 (0.74)	2.29 (0.64)
Domain						
Emotional Support	5.77 (0.60)			5.54 (0.65)	4.65 (0.54)	5.03 (0.69)
Classroom Organization	5.63 (0.71)			4.82 (0.87)	4.29 (0.63)	4.80 (0.81)
Instructional Support	2.71 (0.65)			2.47 (0.68)	1.75 (0.55)	2.12 (0.61)

Note. Data are M (SD). USA sample: Pianta et al. (2008), Finnish sample: Pakarinen et al. (2010), German sample: von Suchodoletz et al. (2014), Chilean sample: Leyva et al. (2015), Chinese sample: Hu et al. (2016).

[a]In the original article, the mean score (6.02) for this dimension (with a negative connotation) was reversed. In this table, for consistency with the presentations of other study results (i.e., a higher score indicates less positive classroom interactions), the original mean score of 6.02 is not reversed to become 1.98.

Figure 5.2

Comparison between the Italian preschool sample and samples from five other countries



Note. Data are M. USA sample: Pianta et al. (2008), Finnish sample: Pakarinen et al. (2010), German sample: von Suchodoletz et al. (2014), Chilean sample: Leyva et al. (2015), Chinese sample: Hu et al. (2016).

5.1.2 Exploratory factor analysis

After examining the descriptive statistics, structural validity was investigated. The CLASS Pre-K observations were used for structural analysis.

Prior to the analysis, data screening was conducted in order to examine the inter-correlation between variables verifying the absence of singularity or extreme multicollinearity, and to test the assumption of univariate normality.

First, the magnitude of the correlations among dimensions and among domains was examined (see Table 5.3). The results suggested that the presence of multicollinearity or singularity was not a significant concern as significant correlation among individual CLASS dimensions ranged from .16 (between Behavior Management and Concept Development) to .62 (between Productivity and Instructional Learning Formats). Overall, the results suggested modest to moderate intra-domain correlations and weak inter-domain correlations. The only exception was represented by correlations among Emotional Support dimensions and Classroom Organization dimensions that presented some moderate convergence values. This result paralleled the high correlation (.66) registered between these two domains, compared to the weak correlations between Emotional Support and Instructional Support (.30) and between Classroom Organization and Instructional Support (.32).

Secondly, skewness and kurtosis of individual CLASS dimensions were examined. Standardized indices are considered highly skewed or kurtotic at > 2.0 and > 7.0, respectively (Fabrigar et al., 1999). Based on these criteria, only one dimension, Negative Climate, demonstrated severe skewness (3.87) and severe kurtosis (21.86), resulting non-normally distributed. Due to these extreme values, Negative Climate was excluded from factor analysis.

Then, exploratory factor analysis (EFA) was carried out in order to identify the factor structure underlying the dataset without imposing any restrictions, whereas in confirmatory factor analysis (CFA) – which aims to confirm theory – restrictions are made to loading matrix (i.e., items are allowed to be loaded on their ‘own’ factor only).

A principal component analysis (PCA) was conducted on the CLASS dimensions (excluding Negative Climate) with oblique rotation (direct oblimin).

The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis, KMO = .77, which was above the acceptable limit of .50 (Field, 2009). Bartlett's test of sphericity χ^2 (36) = 403.75, $p < .001$, indicated that correlations between items were sufficiently large for factor analysis.

An initial analysis was run to obtain eigenvalues for each component in the data. Two components were retained in the final analysis having eigenvalues over Kaiser's criterion of 1 (Kaiser, 1960). In combination, they explained 57.40% of the variance. Table 5.4 shows the factor loadings after rotation. The value of .40 was used as cutoff for acceptable factor loadings (Field, 2009).

All the CLASS dimensions from Emotional Support and Classroom Organization domains loaded together on the first factor. All the CLASS dimensions from Instructional Support loaded together on the second factor. Both two factors had high scale reliabilities as measured by Cronbach's alpha: Factor 1 = .81, Factor 2 = .72.

Thus, the EFA results do not seem to provide preliminary support for the three-factor model posited in the Teaching through Interactions framework. Conversely, the two-factor solution suggested by the EFA is closely similar the two-factor model tested by Hamre and colleagues (2013) – as an alternative to the three-factor solution, with Emotional Support and Classroom Organization combined into a single factor, called Social Support, while the Instructional Support domain remained intact as a second factor. The only difference consists in the exclusion of the Negative Climate dimension in the EFA model, due to its extreme skewed and kurtotic distribution.

Table 5.3

Correlations among the CLASS Pre-K domains and dimensions

	Classroom Organization	**Instructional Support**	1	2	3	4	5	6	7	8	9
Emotional Support	**.66****	**.30****									
1. Positive Climate			–								
2. Negative Climate			-.45**	–							
3. Teacher Sensitivity			.44**	-.39**	–						
4. Regard for Student Perspectives			.39**	-.35**	.29**	–					
Classroom Organization	**–**	**.32****									
5. Behavior Management			.50**	-.41**	.31**	.42**	–				
6. Productivity			.39**	-.27**	.39**	.31**	.31**	–			
7. Instructional Learning Formats			.58**	-.40**	.45**	.37**	.52**	.62**	–		
Instructional Support	**–**	**–**									
8. Concept Development			.23**	-.08	.17*	.21**	.16*	.29**	.37**	–	
9. Quality of Feedback			.06	-.07	.14*	.10	-.09	.09	.09	.48**	–
10. Language Modeling			.37**	-.20**	.14	.22**	.23**	.34**	.28**	.50**	.43**

Note. *$p < .05$ **$p < .01$

Table 5.4

Factor loadings after rotation for the two-factor solution

DOMAIN AND DIMENSION	FACTOR 1	FACTOR 2
Emotional Support		
Positive Climate	**.77**	.03
Teacher Sensitivity	**.64**	.00
Regard for Student Perspectives	**.61**	.03
Classroom Organization		
Behavior Management	**.79**	-.18
Productivity	**.65**	.15
Instructional Learning Formats	**.81**	.08
Instructional Support		
Concept Development	.13	**.78**
Quality of Feedback	-.19	**.87**
Language Modeling	.20	**.71**

Note. Factor loadings above the cutoff value of .40 are evidenced in bold.

5.1.3 Confirmatory factor analysis

Next, a confirmatory factor analyses (CFA) was conducted to test with the current sample the theoretical model proposed by the Teaching through Interactions framework (Pianta et al., 2008; Hamre et al., 2013) assuming three positively correlated factors (Emotional Support, Classroom Organization, Instructional Support). Furthermore, three alternative solutions were tested.

The first alternative model tested was a model with 10 dimensions loading on one global domain (Effective Teaching). Then, a model with 10 dimensions loading on two domains (Social Support and Instructional Support) was tested. Both these models are based on the work of Hamre and colleagues (2013).

Finally, the two-factor solution excluding Negative Climate dimension as indicated in the EFA was tested.

Before conducting the analysis, five univariate outliers were identified and deleted listwise (Field, 2009).

Fit indices for the four tested models are presented in Table 5.5. According to Schermelleh-Engel, Moosbrugger & Müller (2003), TLI and CFI values of .97 indicate a good model fit, whereas values greater than .95 can be interpreted as an acceptable fit. RMSEA and SRMR values less than .05 indicate good fit and values less than .10 indicate acceptable fit (Browne & Cudeck, 1993; Schermelleh-Engel, Moosbrugger & Müller, 2003). Based on these cutoff-points, although the original three-factor model (illustrated in Figure 5.3) showed better fit than the one- and two-domain solutions examined by Hamre and colleagues (2013; illustrated respectively in Figures 5.4 and 5.5), it did not fit the Italian data well nonetheless. Moreover, the high correlation between Emotional Support and Classroom Organization (.85) indicated that there was still a notable overlap among domains.

Table 5.5

Fit indixes for the structural models

STRUCTURAL MODEL	**FIT INDICES**						
	χ^2 (*p*)	df	χ^2/df ratio	*TLI*	*CFI*	*RMSEA*	*SRMR*
Original CLASS Pre-K model	59.781 (.002)	32	1.868	.891	.923	.080	.0659
One-factor model (Hamre et al., 2013)	92.925 (.000)	35	2.655	.793	.839	.110	.0830
Two-factor model (Hamre et al., 2013)	67.060 (.001)	34	1.972	.878	.908	.085	.0670
Revised two-factor model excluding Negative Climate	30.561 (.134)	23	1.329	.962	.976	.049	.0452

Note. TLI = Tucker-Lewis index; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root-mean-square residual.

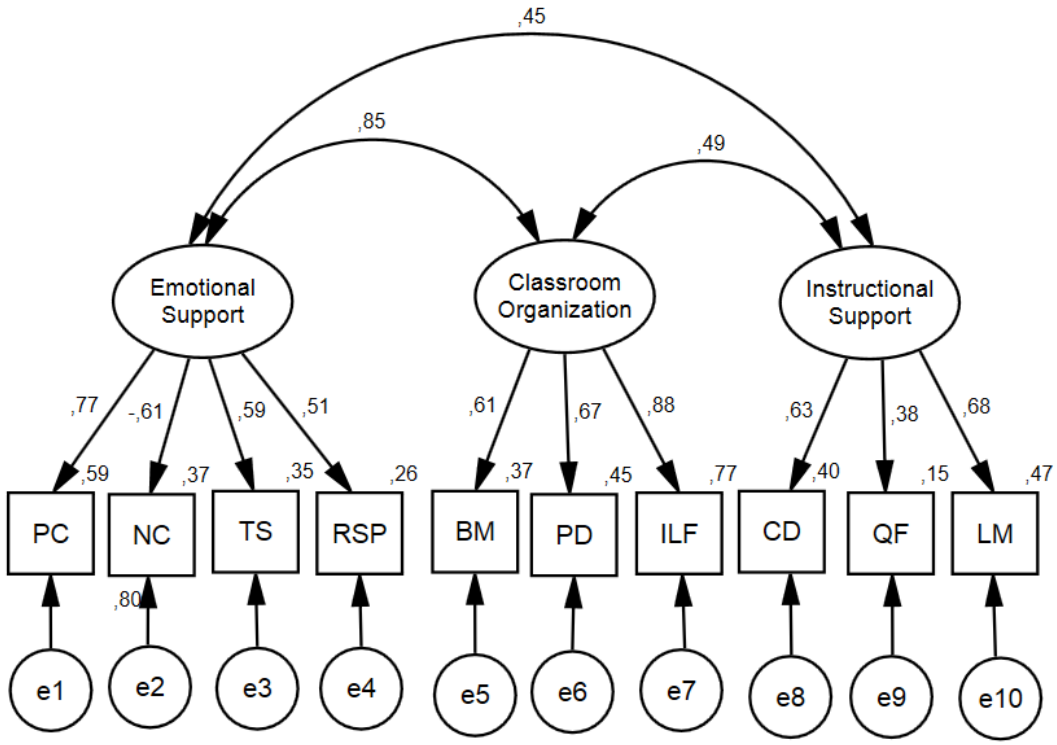


Figure 5.3

Original CLASS Pre-K model (Pianta et al., 2008; Hamre et al., 2013)

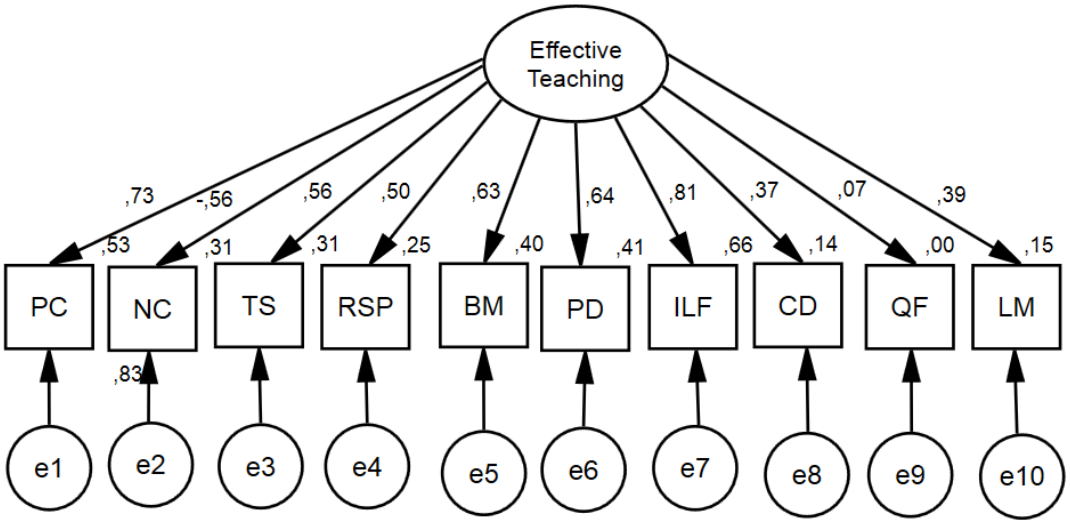


Figure 5.4

One-factor model (Hamre et al., 2013)

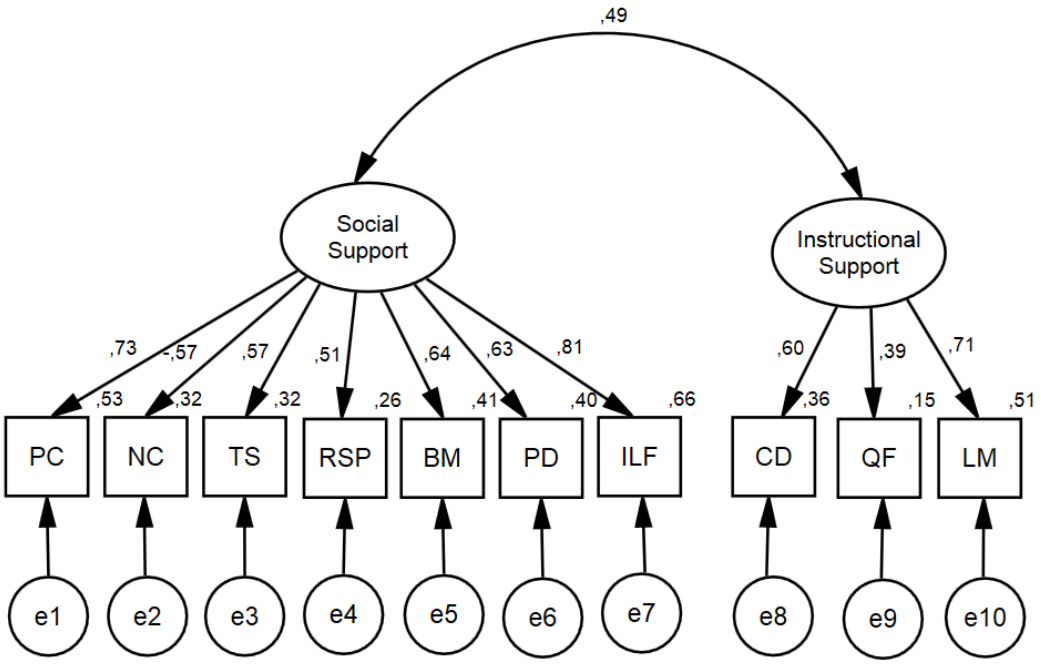


Figure 5.5

Two-factor model (Hamre et al., 2013)

Then, the solution suggested by the EFA was tested. Fit statistics suggested that the model had less than adequate fit: $\chi^2(26) = 55.434$, $p = .001$, CFI = .907, TLI = .871, RMSEA = .091, SRMSR = .0680.

Modification indexes suggested that the fit of the model would increase introducing some modifications. Specifically, allowing the residuals of the observed variables Positive Climate and Productivity, and Behavioral Management and Productivity to correlate created a slight improvement in fit: $\chi^2(24) = 40.089$, $p = .021$, CFI = .949, TLI = .923, RMSEA = .070, SRMSR = .0628. Moreover, introducing an additional substantive revision (i.e., allowing Quality of Feedback to load on both the two factors) besides these two minor modifications (i.e., correlating residuals) resulted in a revised two-factor model (illustrated in Figure 5.6) that fit the data well: $\chi^2(23) = 30.561$, $p = .134$, CFI = .976, TLI = .962, RMSEA = .049, SRMSR = .0452. The present model provided statistically significant parameter estimates.

The results seemed to support the two-factor structure excluding the Negative Climate dimension, as indicated by the EFA, as a better solution to fit the current sample than the original three-factor model.

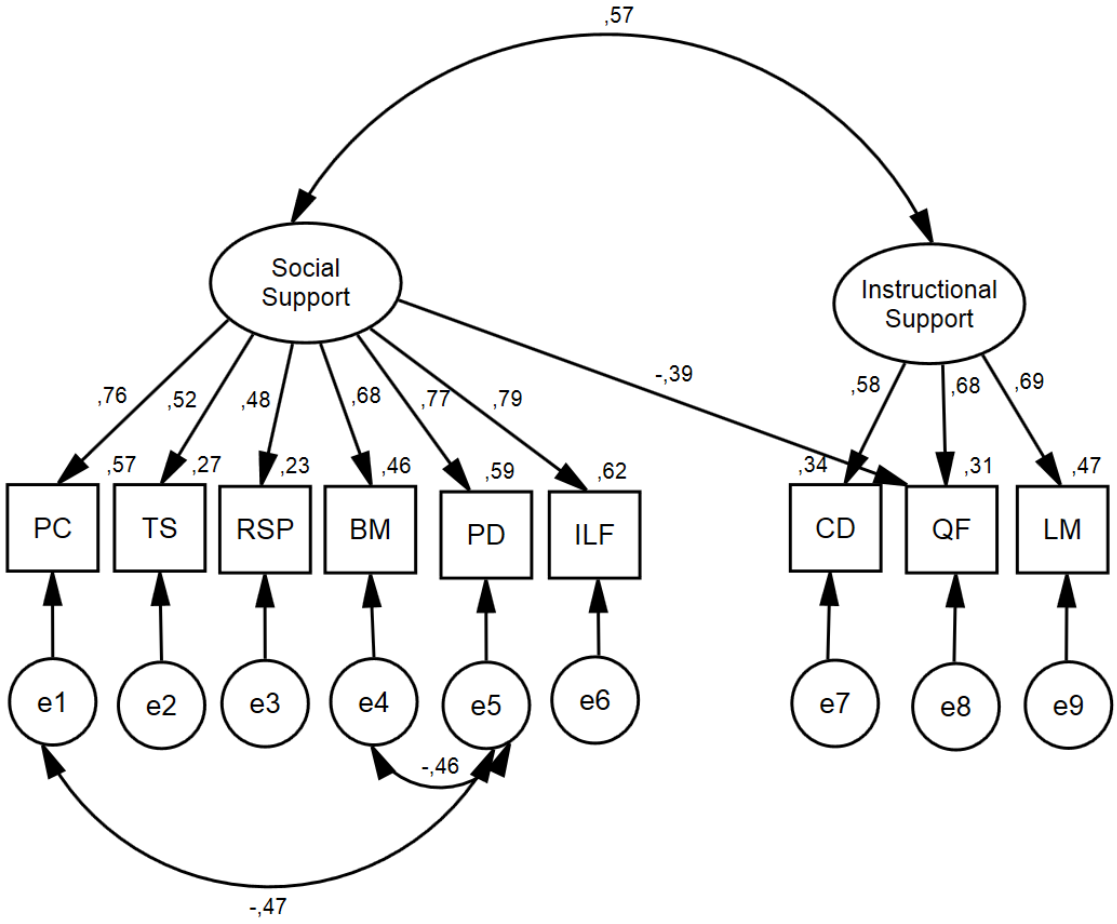


Figure 5.6

Revised two-factor model excluding Negative Climate dimension

5.2 The application of the CLASS Toddler in the Italian infant-toddler centers

The small number of infant-toddler center classrooms video-observed to date limited the quantitative analysis carried out on the sample to examining descriptive statistics of the CLASS Toddler domains and dimensions and comparing them with results reported at international level.

5.2.1 Descriptive statistics and comparisons with other countries

Table 5.6 and Figure 5.7 show means, standard deviations and score distribution ranges for the individual CLASS Toddler dimensions in the Italian sample.

At the domain level, the overall level of Emotional and Behavioral Support (M= 5.75, SD = 0.64) was moderately high, with two dimensions in the high range (Positive Climate and Reversed Negative Climate), two dimensions in the middle to high range (Teacher Sensitivity and Behavior Guidance, scoring between 5 and 6), and one dimension in the middle range (Regard for Child Perspectives).

The overall level of the second domain, Engaged Support for Learning, was in the middle range, as well as the means for its three constituent dimensions.

The standard deviations for the dimensions ranged from 0.33 to 1.29, and for most dimensions were approximately one scale point.

Negative Climate was the dimension registering the lowest variation (SD = 0.33). Score distribution ranged between 1 and 2 and the vast majority of the observational cycles – 87.8% – received a score of 1, revealing a highly skewed distribution (skewness = 2.40, kurtosis = 3.95). This result is in line with findings in the Italian Pre-K sample and further indicates that this dimension poorly differentiated classroom quality among Italian ECEC classrooms.

Table 5.6

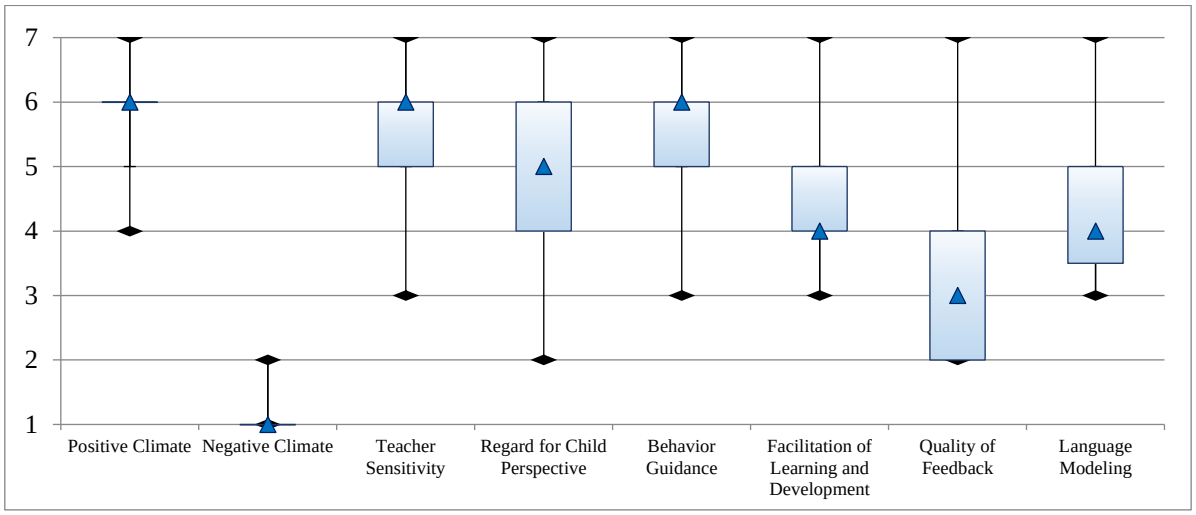
Means, standard deviations and ranges for teacher-child interactions in infant-toddler center classrooms

DOMAIN AND DIMENSION	M	SD	MIN	MAX
Emotional and Behavioral Support	**5.75**	**0.64**		
Positive Climate	6.00	0.78	4	7
Negative Climate	1.12	0.33	1	2
(Reversed Negative Climate)	(6.88)			
Teacher Sensitivity	5.68	0.79	3	7
Regard for Child Perspectives	4.71	1.23	2	7
Behavior Guidance	5.49	0.90	3	7
Engaged Support for Learning	**3.89**	**0.94**		
Facilitation of Learning and Development	4.32	1.01	3	7
Quality of Feedback	3.02	1.29	2	7
Language Modeling	4.32	1.01	3	7

Note. Each scale ranges from 1 to 7 points.

Figure 5.7

Box plot and score distribution ranges for teacher-child interactions in infant-toddler center classrooms



Note. The top and bottom of each rectangular box denote the 75th and 25th percentiles, respectively, with the triangle inside the boxes showing the median. Vertical bars extending from each box represent the extreme values.

The overall picture portrayed by the observations with the CLASS seems to suggest that process quality in Italian infant-toddler center was, on average, moderate to high across the eight dimensions considered by the tool, with Emotional and Behavioral Support registering higher scores than Engaged Support for Learning. In more detail, the dimension presenting the lowest score was Quality of Feedback, suggesting that Italian infant-toddler center teachers, despite creating a generally supportive and warm atmosphere in the classroom and offering children intentional opportunities and guidance for development and learning, were less effective in providing feedbacks that promoted thought processes and understanding. This result is consistent with findings from preschool observations, indicating that this area is likely to be one of the most challenging for Italian ECEC teachers.

Overall, this general pattern – higher levels of Emotional and Behavioral Support, with lower scores of Engaged Support for Learning dimensions – is in

line with those reported in previous studies conducted at international level, as shown in Table 5.7 and Figure 5.8.

Table 5.7

Comparison between the Italian infant-toddler center sample and samples from two other countries. Descriptive statistics

DIMENSION/DOMAIN	ITALY (N = 7)	USA (N = 93)	THE NETHERLANDS (N = 276)
Dimension			
Positive Climate	6.00 (0.78)	5.03 (1.22)	5.42 (1.17)
Negative Climate	1.12 (0.33)	2.70 (0.90)	2.16 (0.38)[a]
Teacher Sensitivity	5.68 (0.79)	4.33 (1.16)	5.34 (1.08)
Regard For Child Perspectives	4.71 (1.23)	4.36 (1.05)	4.24 (1.34)
Behavior Guidance	5.49 (0.90)	4.07 (1.29)	5.01 (1.12)
Facilitation of Learning and Development	4.32 (1.01)	3.43 (1.20)	3.73 (1.35)
Quality of Feedback	3.02 (1.29)	–[b]	2.91 (1.20)
Language Modeling	4.32 (1.01)	2.22 (1.07)	3.22 (1.29)
Domain			
Emotional and Behavioral Support	5.75 (0.64)		
Engaged Support for Learning	3.89 (0.94)		

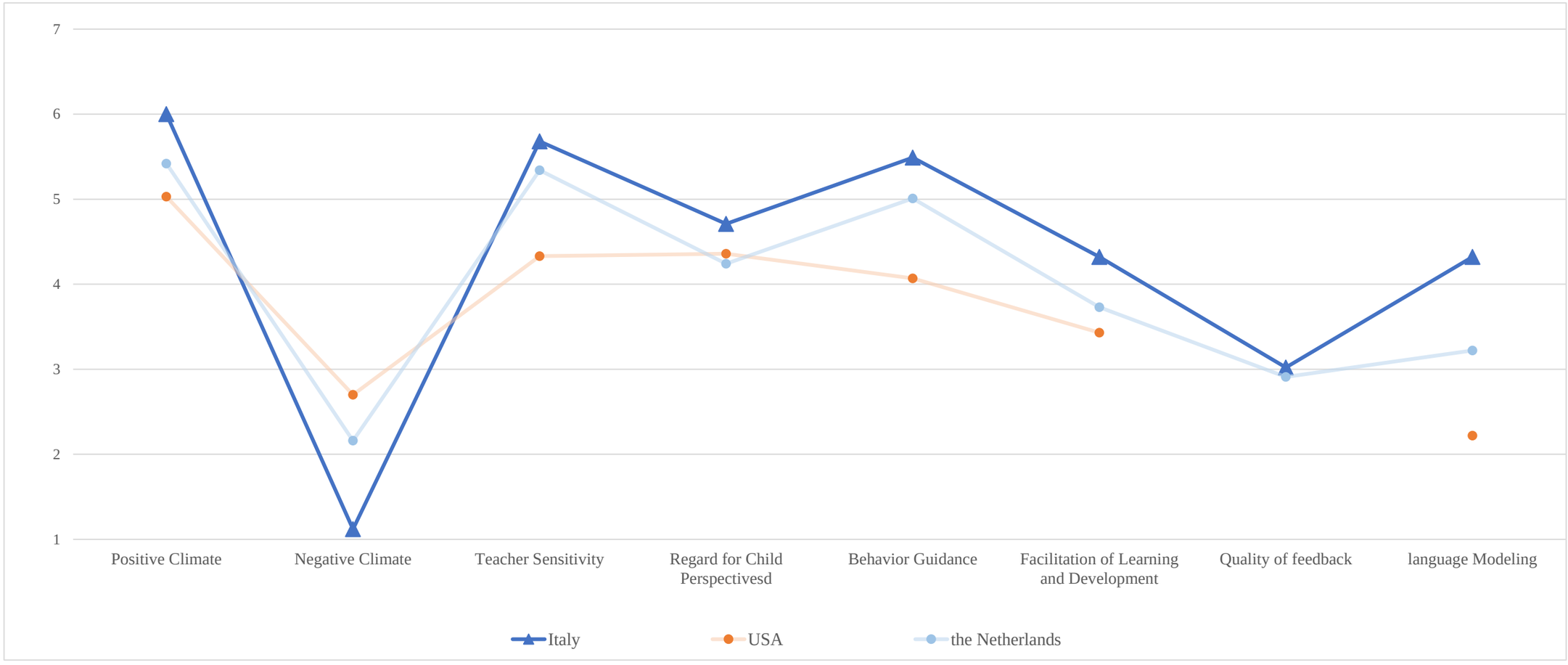
Note. Data are M (SD). USA sample: La Paro, Hamre & Pianta (2012), Dutch sample: Slot et al. (2015).

[a]In the original article, the mean score (5.84) for this dimension (with a negative connotation) was reversed. In this table, for consistency with the presentations of other study results (i.e., a higher score indicates less positive classroom interactions), the original mean score of 5.84 is not reversed to become 2.16.

[b]The version of the CLASS Toddler used in the present study did not include Quality of Feedback.

Figure 5.8

Comparison between the Italian infant-toddler center sample and samples from two other countries



Note. Data are M. USA sample: La Paro, Hamre & Pianta (2012), Dutch sample: Slot et al. (2015).

Chapter 6: Discussion

I'm always asking questions – not to find 'answers',
but to see where the questions lead. Dead ends sometimes? That's fine.
New directions? Interesting. Great insights? Over-ambitious.
A glimpse here and there? Perfect.
Lesley Hazleton

This is the first study to investigate, adopting a mixed-methods research design, the implications – both at cultural and methodological level – of applying the CLASS tool in the Italian ECEC context.

The qualitative findings suggest that, despite the presence of elements of continuities between the conceptualization of effective teaching rooted in the Italian pedagogical tradition and the one embedded in the tool, there are nonetheless significant differences and discrepancies. This result is yet more significant considering that these discrepancies mirror inconsistencies emerging at statistical level.

Moreover, evidence was found to support the possible fruitful integration of a standardised assessment with a participatory-reflective approach to assessment.

In the following paragraphs, the empirical findings will be discussed and elaborated upon, as will the limitations of the study. Lastly, suggestions for subsequent research will be provided.

6.1 The CLASS outside of its cultural cradle: insights provided by the integration of qualitative and quantitative results

A mixed-methods design was adopted in order to gain a more thorough and comprehensive understanding of the research problem than either a qualitative or a quantitative approach alone could provide.

In fact, some interesting insights were offered by combining and comparing the two set of data.

6.1.1 Challenging the universality of the CLASS framework

A first interesting highlight concerns the assumption, suggested by the CLASS authors (Hamre, Goffin & Kaft-Sayre, 2009; Hamre et al., 2013; Pianta et al., 2009; Vitiello, 2013), that the dimensions of classroom quality assessed by the tool are relevant across cultures. This assumption is called into question by the results of the present study.

The analysis of qualitative data showed that, although the CLASS framework contains many elements of interest and continuity, it also presents noteworthy discrepancies with the Italian pedagogical perspective. Specifically, Italian practitioners pointed out several elements of teacher-child relationships that are differently interpreted by the CLASS, key-features not captured by the tool whilst considered crucial in their definition of ECEC quality, and aspects of the CLASS conceptualization that do not mirror their idea of what effective teaching is.

As illustrated in detail in Chapter 5, the discrepancies emerged discussing the tool with Italian teachers and pedagogical coordinators were reflected at statistical level – at least in respect of the CLASS Pre-K.

The exploratory factor analysis (EFA) conducted on the Pre-K dataset (i.e., observational cycles of Italian preschool classrooms coded using the CLASS Pre-K) suggested that a two-factor solution better described Italian data rather than the three-domain latent structure posited in the Teaching through Interactions framework (Hamre & Pianta, 2007; Hamre et al., 2013).

These results were paralleled by those found conducting a confirmatory factor analysis (CFA). In fact, on the one hand, CFA findings did not support the original three-factor model, since fit indices did not meet the recommended criterion thresholds (Browne & Cudeck, 1993; Schermelleh-Engel, Moosbrugger & Müller, 2003) in the present sample. On the other hand, the revised model with two factors, which excluded Negative Climate dimension and had several minor modifications (i.e., correlating residuals), was determined to be the best fitting model with the current sample.

Moreover, the increased model fit resulting from introducing a substantial modification (i.e., allowing Quality of Feedback to load on both the two factors) further questioned the structural validity of the tool when applied to the Italian context, and provided additional evidence of the need to reconceptualise the model in order to properly assess relational quality in the Italian ECEC.

Overall, these qualitative and quantitative evidences combined strongly challenge the taken for granted universality of the CLASS framework. It invites a reflection on the potential biases and validity threats to which disregarding the cultural complexities of its international use may lead. They suggested that an uncritical application of this instrument not only may not fully encompass the quality of the service assessed, but it may also provide a partial picture, incapable of accounting local-cultural peculiarities and respecting pedagogical values and interpretations in which practitioners identify themselves.

6.1.2 The close link between classroom organization and emotional support

A second highlight regards the two-factor structure that, according to both EFA and CFA, seemed the best solution to describe Italian Pre-K data.

In this model, all the CLASS Pre-K dimensions from Emotional Support domain (excluding Negative Climate) and from Classroom Organizations domain loaded together on the first factor, whereas all the dimensions from Instructional Support loaded together on the second factor.

A similar organization by means of two factors features also in the CLASS Toddler, which presents two broad domains (Emotional and Behavioral

Support, and Engaged Support for Learning – see Table 2.2. in Chapter 2) to describe classroom quality. Moreover, this twofold structure presents analogies with some of the organizational frameworks that have been proposed in narrative reviews (Brophy, 1999; Elias, 2003). These are frameworks which emphasise the presence of two main components on which researchers must focus, namely academic knowledge and socio-emotional skills, in order to fully encompass classroom processes, and with which teachers must deliberately engage in order to effectively promote children's development and learning.

Although this two-factor solution is not entirely new in literature, it is not consistent with the conceptualization posited by the Teaching through Interactions framework (Hamre & Pianta, 2007; Hamre et al., 2013) which divides classroom socio-emotional features into two separate domains. On the one hand, Emotional Support assesses how effectively teachers create a welcoming atmosphere and a sense of security that allows for exploration of novel, give children individualized attention, and support their need to feel competent and autonomous – all crucial aspects that contributes to lay the foundations for a motivation to learn (Downer, Sabol & Hamre, 2010; Hamre et al., 2007; Pianta & Hamre, 2009b). On the other hand, Classroom Organization emphasises the role of organization and management of students' behavior, time, and attention in creating a smoothly functioning classroom and providing external regulations, which are considered crucial to help children develop behavioural and attentional self-regulatory skills (Downer, Sabol & Hamre, 2010; Hamre et al., 2007; Pianta & Hamre, 2009b).

However, the distinction between classroom emotional and organizational features proposed by the Teacher through Interactions framework is unsupported by Italian data and not merely at statistical level: the supposed duality between these two CLASS Pre-K domains is not even reflected in Italian practitioners' voices.

In fact, according to Italian teachers and mirroring the holistic perspective typical of Italian pedagogical tradition (Mantovani, 2007a, 2014), the demarcation line between Emotional Support and Classroom Organization, valid at theoretical level, actually fades in their daily educational practices, revealing instead how closely these two domains are interwoven.

As discussed in Chapter 4, there are several connections that Italian teachers identified between dimension belonging to these two domains that the CLASS Pre-K keeps separate: for instance, just to mention a few, the attention paid by Italian teachers to recognize and value children as competent subjects involving them in discussing and defining – rather than predetermining and imposing – classrooms rules; the promotion of self-regulatory skills allowing children to enforce rules and even learn to accept their peers' mutual correction; the active role assigned to children in co-constructing projects, activities and routines and actively shaping the classroom curriculum; the effort to offer children a significant pressure-free experience in which even time spent waiting can become an opportunity to learn rules of social life and respect for one another.

In this light, the distinction between classroom emotional and organizational features dissolves because the emotional support provided by teacher ceases to be conceived only in terms of how effectively they create a warm, secure, supportive climate in the classroom, and broadens to embrace also how teachers deliberately support children in learning emotional skills – a kind of learning that necessarily takes place in classroom social life and through the definition of its underlying rules, times and routines.

6.1.3 The lack of emphasis on promoting children's cognitive skills

This holistic vision, which, while comprising under a single domain both classroom organizational features and socio-emotional support, does not include teachers' interventions aimed at supporting children's understanding and reasoning, introduces a third interesting highlight.

Quantitative data presented a picture of Italian ECEC services in which cognitive skills and conceptual learning did not seem particularly emphasized. In fact, the low mean scores on Quality of Feedback (both in the Pre-K and Toddler sample) and Concept Development seem to suggest that often throughout the observations conducted teachers focused more on basic skills than on promoting children's reasoning and thinking.

In the international landscape, this result does not represent an exception, since these dimensions receive scores in the low range in almost all the studies

conducted in the USA and in other countries (see Chapter 5). This common cross-cultural pattern paves the way for two possible interpretations.

On the one hand, this pattern may be a result of inherent limitations of the tool itself. CLASS criteria for these dimensions may be too challenging or improperly operationalized resulting in a majority of classrooms obtaining scores that are clustered together in the low range, with very few extending across the full possible range of scores. The resulting floor effect, which at statistical level implies a skewed distribution with limited variability, may also greatly limit the utility of the assessment tool, due to its inability to discriminate teachers who differ in terms of how effectively they support children's cognitive development.

On the other hand, the low scores received at international level by these dimensions may reflect an actual common fragility, indicating a teachers' general undervaluation of or ineffectiveness in relation to these aspects.

In this respect, joint analysis of qualitative and quantitative data can contribute to shed light on this issue, providing some evidences to support the second hypothesis.

At qualitative level, Italian practitioners pointed out some limits in the conceptualization of learning proposed by the CLASS. Specifically, they argued that the tool embraces a rather narrow vision of what learning is, which – despite accentuating its cognitive component – leaves out many elements considered crucial from the Italian perspective (such as socio-emotional learning, learning to respect diversity and differences, learning through peer interactions…). Therefore, it underestimates Italian teachers' competences in promoting a broader range of learning. However, despite this significant gap, the majority of practitioners also recognized that many occasions to foster children's reasoning are missed in their own classroom, and agreed with the feedback provided by the tool in this regard. Thus, they appreciated the CLASS lens precisely because it draws and focus their attention on this often undervalued aspect and prompts them to reflect on how they can more effectively seize opportunities to foster children's cognitive skills during different activities and moments of the school day.

Moreover, an in-depth analysis of quantitative data suggest that, unlike Negative Climate scores which distribution is actually extremely skewed, these

CLASS dimensions present a higher variability and, albeit only in a limited number of observation cycles, even achieve the high range.

In this regard, two aspects should be stressed. First, in these observations not only were teachers effective in stimulating children's reasoning and understanding, but, in line with the principles of Italian early childhood pedagogy, they also often involved the group in these moments, expanding further child's cognitive learning through peers' contributions and comments. Secondly, in the majority of these cases, the observations were conducted in the ECEC centres selected within the CARE project as examples of 'good practices' and, thus, recognized by Italian researchers as high-quality services.

Therefore, on the one hand, these combined evidences seem to suggest that the CLASS, despite addressing a limited conceptualization of learning, is able to capture properly variations in teachers' effectiveness in stimulating children's higher-order thinking skills, discriminating and identifying those classrooms that stand out because of the more effective cognitive stimulation and feedback provided by their teachers.

On the other hand, they point out a general, widespread weakness in supporting children's cognitive learning that characterize Italian teachers – albeit probably not confined only to Italian ECEC services.

This tendency seems to reflect an idea of early childhood education that, despite the formal, well-established recognition of ECEC services as an environment of learning, relationship and life (DM 3/6/1991), in practice continues to emphasize mainly the socio-emotional dimension at the expense of attention to cognitive skills – skills perhaps still viewed to be primarily the prerogative of the higher educational levels.

The presence of this tendency in the Italian ECEC context and the concurrent need to overcome it are clearly outlined by Mantovani (2014, p. 27, translation by the author):

> «We must prevent the pressure on results, but also the stifling cocoon. For this purpose, we need to abandon once and for all the idea that children should be *protected* against learning and that ECEC services are meant only for care and for a wellbeing that disregards the pleasure that comes from knowledge and from feeling competent. […] We need to rethink our holistic approach to revitalize it. […] In fact, it is necessary to go beyond – in a new and dialectic

way – the outdated distinction between cognitive and socio-emotional [aspects], reintroduced today by the literature on *curriculum*. Today's wellbeing and not being excluded tomorrow are two equally important rights. We must find a way to put them back together».

6.2 Opportunities offered by combining a standard-based assessment and a participatory-reflective approach

As abovementioned, the emphasis placed by the tool on supporting children's cognitive skills was one aspect of the CLASS framework most appreciated by Italian teachers, since it was considered a valuable stimulus to reflection and improvement of their own educational practices in an area often undervalued.

These considerations draw the attention to a further issue addressed by the present study that deserves to be discussed in greater depth: the possibility of integrating fruitfully two apparently opposing perspectives on quality evaluation, namely the standard-based assessment provided by the CLASS and the participatory-reflective approach rooted in Italian ECEC services.

The Italian early childhood pedagogical tradition assigns great importance to practitioners' reflectivity, recognizing that it is a key-feature to defining and enhancing service quality. «Educative profession is a difficult profession, never guaranteed, always "under examination", constantly questioning its own practices and its own self-understanding» (Cambi et al., 2003, p. 45, translation by the author). Thus, as Sarsini and Di Bari (2015, p. 76-77, 82, translation by the author) explain citing Schön (1993):

> «Reflectivity allows educators to learn from their professional experience, comparing it with their colleagues', and to rethink their actions as well as their underpinning assumptions, while learning from mistakes, in a virtuous circle of feedback that becomes the key element of educational practice. It is not possible to get away from this constant and critical return on actions, knowledge, and projects because it is impossible to linearize or standardize the educational process due to the complexity and dynamism that constitute

it. […] Precisely this *sub judice* condition obliges the educator to become a reflective practitioner».

As a consequence, evaluation itself becomes the culmination of this reflective process:

> «Evaluation is primarily a reflective activity, [an activity of] confrontation, research and analysis; evaluation is neither aimed at stressing the deviation from standards and procedures, nor at rigidly indicating what are the best standards and procedures to follow. Evaluation asks "why?", looks for reasons behind mistakes and successes» (Parente, 2015, p. 56, translation by the author).

Within this general frame of reference, it can be understood why a standard-based assessment, such as the one proposed by the CLASS, may seem so distant and even incompatible with the reflective approach to evaluation predominant in Italian ECEC services. Besides this, the concept of standard itself may appear better suited to providing answers rather than stimulating questions, more focused on indicating how to *do things right* rather than inviting to reflect on how to *do the right things* (Coussée et al., 2010; Peeters & Vandenbroeck, 2011) in accordance with the complexities and peculiarities of each situation.

However, voices of Italian practitioners involved in the present study illustrated a diverse perspective to look at the relationship between these two evaluation approaches, suggesting that, more than the tool characteristics themselves, what really makes the difference is the way in which it is used and proposed to services.

When a tool – albeit an instrument structured and not negotiated like the CLASS – rather than rigidly imposing standards to be attained, instead becomes an *interlocutor* to dialogue and compare with, it ceases to seem 'foreign' or even 'scary'. When practitioners are involved in critically discussing and questioning meanings underpinning the assessment instrument, from a transformative and empowering rather than a prescriptive perspective, the tool becomes particularly valuable and useful to give fresh impetus to teachers' reflectivity – especially due to the cultural and methodological differences that it conveys. It becomes an occasion to explicit and become more aware of local definitions of quality,

to understand how deeply they influence their own educational practices, and to gain new insights and suggestions to improve these practices.

It should be stressed that a similar, critical approach to standard-based instruments, respectful of the local evaluation culture, has been already experimented in Italian ECEC services. For instance, as Savio (2008, p. 207, translation by the author) explains in regard to the SVANI:

> «[The critical evaluation of the SVANI] makes you ask yourself: how do I position myself in relation to the SVANI proposal? Do I agree with it? If not, how does my idea differ from it? And ultimately, what is in detail my idea of "good" nido? It is precisely this process of comparison, by analogy or by contrast, that lets emerge in a clearer and more precise way your own idea of nido. In such a way, the SVANI becomes a guide, a highlighter of blind spots in educational choices, stimulating even at individual level questions that uncover, situation by situation, your latent pedagogy and bring into focus your educational identity».

6.3 Limitations of the study

Several potential limitations need to be considered when interpreting the results from the current study.

First, the sample size of the present study was small. This may have arisen problems especially at quantitative analysis level. The reliability of factor analysis is dependent on sample size. Many authors recommend higher values for minimum necessary sample size in structural equation modelling (e.g., N = 200, Guilford, 1954; N = 300, Comrey & Lee, 1992), since small sample sizes might lead to inaccurate and unstable estimation of the parameters and sometimes inflated values for goodness-of-fit indices (e.g., Gagné & Hancock, 2006; Muthén & Muthén, 2002). However, further research, pointing out limitations of guidelines based only on sample size (MacCallum et al., 1999; Velicer & Fava, 1998), has suggested other criteria to refer to (MacCallum et al., 1999 – see Note 19 in Chapter 3), according to which the current Pre-K sample can be considered adequate to perform factor analysis. Therefore,

despite the evident limitation of a small sample size, the results of the present study can be assumed to be sufficiently valid.

Second, although data were collected in different Italian Provinces, the sample cannot be considered representative of Italian ECEC services nationally. For instance, the sample did not include infant-toddler centers and preschools in the southern regions, whereas Region Lombardy and especially the Province of Milan were over-represented. Moreover, it did include only public ECEC services[21], while private services – which account for, respectively, over two-fifths and almost a third of all children enrolled in infant-toddler centers and preschools (Istat, 2012; OECD, 2014) – were not represented.

Third, all the CLASS observational cycles were coded by the author, who went through the rigorous CLASS reliability process and was certified as a reliable CLASS observer for both the Toddler and Pre-K versions of the instrument. However, the presence of a single observer did not allow for double-coding and, consequently, for calculating inter-observer agreement – a condition necessary to monitor observer drift and further verify reliability of the ratings assigned.

Finally, another potential limitation lies in the use of video for coding, instead of live observation of classroom teaching. Although other published studies have used the video-observation procedure in the USA (Allen et al., 2013; Mashburn et al., 2007) and at international level (Araujo et al., 2014; Hu et al., 2016; Leyva et al., 2015), it should be taken into account that videotaping might result in missing some crucial cues observable only during a live observation.

6.4 Suggestions for further research

The results of the present study, besides offering interesting stimuli for reflections and contributing to shed light on the cultural and methodological issues regarding the international application of the CLASS tool, arise several future research suggestions.

[21] The only exception was represented by a single infant-toddler center that was run by a Consortium (involving collaboration between public institutions, no profit organizations and private companies).

With relevance to the above-described limitations, the conduction of further research could potentially be directed at extending the existing study, both qualitatively and quantitatively, by enrolling a larger sample group. It is recommended that future studies sample from more Italian Regions and Provinces and include ECEC services run by different providers (both public and private) in order to be more representative of the remarkable diversity that characterizes Italian ECEC landscape nationally. Moreover, a larger sample could allow new observations and comments on the tool to emerge from the qualitative discussions with practitioners. Furthermore, it could enhance the reliability and validity of the statistical findings, allowing also for extending factor analysis to the observations conducted with the CLASS Toddler.

Additionally, involving a greater number of certified observers in the coding phase could allow to monitor observer drift and to increase reliability of the ratings assigned.

Ultimately, collecting more qualitative and quantitative data could also provide a broader, sounder basis to verify the findings obtained by the current study. It could also provide indications on the next research steps to be taken, for instance, by suggesting the feasibility and usefulness of working on an adaptation of the tool to the Italian ECEC context. Alternatively, additional data may determine whether it would instead be more profitable to develop a unique, distinct instrument to assess the quality of teacher-child interactions, better-adapted to incorporate Italian peculiarities.

Furthermore, in the light of the rich contents offered by qualitative data and of the valuable insights provided by the combined interpretation of qualitative and quantitative results, it may be worthwhile to replicate in other countries the methodology applied in the present study[22]. Specifically, a mixed-methods approach involving a critical-cultural discussion on the tool with ECEC teachers could be valuable in those countries in which statistical analysis have pointed out some inconsistencies in the CLASS framework, despite being unable to provide adequate explanations for the reasons behind them. In fact, as Fenech (2010, p. 294) argues:

[22] An effort in this direction was initiated within the CARE project, which undertook a qualitative cross-country analysis of the CLASS involving groups of teachers from a limited number of selected preschools and infant-toddler centers in Italy, Portugal and the Netherlands (Pastori et al., 2016; Pastori & Mantovani, 2016).

«These quantitative measures […] can give conflicting messages about quality. They also generate questions about the elements that underpin existing and changing levels of quality, questions that these measures alone cannot answer. Complementing more comprehensive quantitative measures that incorporate structural elements of quality with qualitative explorations has the potential to provide more nuanced understandings of high quality, and thus a stronger platform to inform policy and practice».

Conclusion

Once you allow for some diversity
and recognize the possibility of multiple perspectives,
where do you draw the line?

The recent, increasing *globalization* of assessment tools can confer undeniable advantages (Grammatikopoulos, Gregoriadis & Zachopoulou, 2015; Limlingan, 2011). However, as literature review clearly shows, disregarding the underlying cultural and methodological complexities of the international use of these tools may also lead to possible pitfalls.

As the present study has pointed out, even tools with a solid theoretical and empirical background and widespread use internationally, such as the CLASS, cannot be considered *culture-free*. Being *children* of the contexts where they were developed, they still unavoidably reflect cultural values and methodological assumptions typical of their cultural *cradle* and not necessarily shared in the real contexts to which they are *exported* (Pastori & Pagani, *forthcoming*).

Therefore, their *migration* may have consequences at various levels. At statistical level, it may lead to potential biases and validity threats. More significantly, it may imply applying instruments in which local practitioners cannot recognize themselves and that are not able to fully capture – in all its complex variations – quality of the services that they intend to assess.

These considerations, based on the present discussion on the CLASS tool but likely to result in a broader reflection on the international use of standard-based assessment instruments, relate back to the opening question:

> «Once you allow for some diversity and recognize the possibility of multiple perspectives, where do you draw the line?» (Dahlberg, Moss & Pence, 1999, p. 104).

In the present case, «as the possibility of standardization fades in the face of diversity and complexity» (Dahlberg, Moss & Pence, 1999, p. 104), could renouncing the cross-cultural application of evaluation measures be the solution?

As Limlingan effectively notes (2011, p. 45), this is unlikely to be the answer:

> «As the world becomes smaller, it will become more important to find the most effective ways to organize and share information. Cross cultural comparisons using a common instrument, *so long as it is composed and utilized in the right way*, provides a good method to facilitate discussions which allow us to learn from one another» (Limlingan, 2011, p. 45, emphasis added).

These words introduce a key-element that plays a pivotal role throughout the present discussion. Refraining from applying internationally standard-based tools for fear of potential pitfalls would be as blind as their uncritical use. What is crucial is not to debate whether or not to implement these tools cross-culturally, but to reflect instead on how to use them in a *proper*, *responsible way*.

In fact, while it is true that each tool is vessel of culturally-bound values and assumptions about quality and pedagogy that may not fully reflect viewpoints and interpretations of different groups in different places; nonetheless it is not necessary to assume an extreme relativistic position. As Mantovani and Rogoff point out:

> «[Quality] shall have both a common, recognizable basis and specific cultural and local variations, according to early childhood pedagogical tradition and up to date with what cultural pedagogy asserts: educational universals are expressed through cultural regularities and, thinking about education and educational contexts, universalistic and cultural perspectives should dialogue with one another» (Mantovani, 2014, p. 25, translation by the author).

> «[The necessity to recognize cultural variation] does not mean that each community has a unique set of values and goals. There are regularities among the variations» (Rogoff, 2003, p. 23).

This scenario reveals and highlights the importance and the need to adopt a *critical-cultural* approach to evaluation measures – an approach attentive to the cultural and methodological complexities when these instruments reach contexts different from the original ones.

The inherent risk of the uncritical use of assessment tools is that they inadvertently be allowed to become the main criterion for defining what quality is rather than serving as means to measure quality. As a result, the evaluation process itself is likely to be reduced to a sort of screen that obscures cultural peculiarities and silences meanings and interpretations of quality that underpin local cultures of each ECEC service.

Conversely, adopting a critical-cultural approach means questioning a rigid universalistic idea of educational standards of quality, recognizing at the same time the value of cultural differences and the existence of continuities and similarities despite cultural variation.

It means ceasing to intend the relationship between the tools and the services they evaluate as a top-down, unidirectional, prescriptive process. Rather, it invites to consider instruments as interlocutors with which to dialogue and compare and even to question.

Ultimately, adopting a critical-cultural approach means acknowledging that what really makes the difference is the way in which tools are used and proposed to services, since every *tool* is not useful in itself, but – as the word itself implies – becomes valuable only in the capable, mindful hands of its users, whether they be researches, evaluators or practitioners.

From this perspective, it can be understood how distant and even apparently opposing perspectives on quality evaluation – such as a standard based assessment and a participatory-reflective approach – can nevertheless be revealed to share a common ground and can be integrated in an enriching way. Thus, even structured, standardized tools like the CLASS can become a valuable *trigger* of teachers' reflectivity – enabling teachers to develop new ways of looking at their own actions, to allow a fruitful *métissage* of their practices with new perspectives, to gain new awareness or understanding of their own assumptions and behaviours. Sustaining teachers' reflectivity can, therefore, foster their professional development, enhance their effectiveness in the

classroom, and, ultimately, improve quality (Day, 1999; Ferraro, 2000; Lazzari, Picchio & Musatti, 2013; Osterman, 1990). After all, that is, in essence, the heart of every evaluation process and the vital purpose for which such tools are developed.

References

Abbot-Shinn, M., & Sibley, A. (1992). *Assessment profile for early childhood programs: Research version*. Atlanta, GA: Quality Assist.

Aikens, N., Moiduddin, E.M., Xue, Y., Tarullo, L.B., & West, J. (2012). *Data tables for Child outcomes and classroom quality in FACES 2009 report* (OPRE Report 2012-37b). Washington, DC: U.S. Administration for Children and Families, Office of Planning, Research and Evaluation.

Alexander, R.J. (2000). *Culture and pedagogy: international comparisons in primary education*. Oxford: Blackwell.

Allen, P., Gregory, A., Mikami, A., Lun, J., Hamre, B.K., & Pianta, R.C. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the Classroom Assessment Scoring System-Secondary. *School Psychology Review*, 42, pp. 76-98.

Allulli G. (2008), *Autonomia e valutazione*. Working paper Fondazione Agnelli, 7. Retrieved from: http://www.fga.it/uploads/media/G._Allulli__Autonomia_e_valutazione_-_FGA_WP7.pdf.

Araujo, M.C., Carneiro, P., Cruz-Aguayo, Y., & Shady, N. (2014). *A Helping Hand? Teacher Quality and Learning Outcomes in Kindergarten*. Working paper, version: October 4, 2014. Retrieved from: http://economics.yale.edu/sites/default/files/carneiropedroahelpinghand_final.pdf.

Arnett, J. (1989). Caregivers in day-care centers: does training matter?. *Journal of Applied Developmental Psychology*, 10, pp. 541–552.

Athey, C. (1990). *Extending thought in young children: a parent-teacher partnership*. London: Paul Chapman.

Barrett, M.D., Huber, J., & Reynolds, C. (2014). *Developing intercultural competence through education*. Strasbourg: Council of Europe Publishing.

Becchi, E. (2000). La qualità educativa: punti di vista e significati. In A. Bondioli & P.O. Ghedini (Eds.), *La qualità negoziata. Gli indicatori per i nidi della Regione Emilia Romagna* (pp. 39-54). Bergamo: Edizioni Junior.

Becchi E., Bondioli, A., & Ferrari, M. (1999). ISQUEN: indicatori e scala di valutazione della qualità educativa al nido. In L. Cipollone (Ed.), *Strumenti e indicatori per valutare il nido. Un percorso di analisi della qualità di un servizio educativo in Umbria* (pp. 134-180). Bergamo: Edizioni Junior.

Birch, S.H., & Ladd, G.W. (1998). Children's interpersonal behaviors and the teacher-child relationship. *Developmental Psychology*, 34, pp. 934–946.

Bondioli, A. (2002). Dagli Indicatori alle condizioni del Progetto educativo: un percorso pedagogico-politico di definizione e di assicurazione della qualità degli asili nido della regione Emilia-Romagna. In E. Becchi, A. Bondioli & M. Ferrari (Eds.), *Il progetto pedagogico del nido e la sua valutazione. La qualità negoziata* (pp. 11-25). Bergamo: Edizioni Junior.

Bondioli, A. (2013). Prospettive sulla qualità: il caso degli asili nido. *RELAdEI. Revista Latinoamericana de Educación Infantil*, 2, pp. 19-33.

Bondioli, A., & Ferrari, M. (Eds.) (2004). *Verso un modello di valutazione formativa. Ragioni, strumenti e percorsi*. Bergamo: Edizioni Junior.

Bondioli, A., Ferrari, M. (Eds.) (2008). *AVSI - Autovalutazione della scuola dell'infanzia. Uno strumento di formazione e il suo collaudo*. Bergamo: Edizioni Junior.

Bondioli, A., & Ghedini, P.O. (Eds.) (2000). *La qualità negoziata. Gli indicatori per i nidi della Regione Emilia Romagna*. Bergamo: Edizioni Junior.

Bondioli, A., & Nigito, G. (2008). *Tempi, spazi, raggruppamenti. Un Dispositivo di Analisi e Valutazione dell'Organizzazione Pedagogica della Scuola dell'Infanzia (DAVOPSI)*. Bergamo: Edizioni Junior.

Bondioli, A., & Savio, D. (Eds.) (2010). *Partecipazione e qualità. Percorsi di condivisione riflessiva nei servizi per l'infanzia di Modena*. Bergamo: Edizioni Junior.

Bondioli, A., & Savio, D. (2015). La valutazione di contesto nei servizi per l'infanzia italiani: riflessioni, ricerche, esperienze e prospettive per il futuro. In A. Bondioli & D. Savio (Eds.), *La valutazione di contesto nei servizi per l'infanzia italiani. Riflessioni ed esperienze* (pp. 7-22). Parma: Edizioni Junior-Spaggiari Edizioni.

Bottani, N., & Cenerini, A. (2003). *Una pagella per la scuola: la valutazione tra autonomia e equità*. Trento: Erickson.

Bouchard, C., Cantin, G., Charron, A., Crépeau, H., & Lemire, J. (2014). *Validation du Classroom Assessment Scoring System (CLASS) en maternelle 4 ans à mi-temps au Québec*. Paper presented at the 82th Congres de l'Acfas, Colloque 519, Montréal, Canada, May 15.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.

Boyatzis, R.E. (1998). *Transforming qualitative information: thematic analysis and code development*. Thousand Oaks, CA: Sage.

Bracci, E. (2009). Autonomy, responsibility and accountability in the Italian school system. *Critical Perspectives on Accounting*, 20, pp. 293-312.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), pp.77-101.

Brilli, Y., Del Boca, D., & Pronzato, C.D. (2016). Does child care availability play a role in maternal employment and children's development? Evidence from Italy. Review of Economics of the Household, 14(1), pp. 27-51.

Brophy, J. (1999). *Teaching*. Educational Practices Series 1. Paris: International Academy of Education and International Bureau of Education.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Bruner, J.S. (1980). *Under Five in Britain*. London: Grant McIntyre.

Buyse, E., Verschueren, K., Doumen, S., Van Damme, J., & Maes, F. (2008). Classroom problem behavior and teacher-child relationships in kindergarten: The moderating role of classroom climate. *Journal of School Psychology*, 46(4), pp. 367-391.

Cadima, J., Leal, T., & Burchinal, M. (2010). The quality of teacher-student interactions: Associations with first graders' academic and behavioral outcomes. *Journal of School Psychology*, 48(6), pp. 457-482.

Calzolari, N., D'Ugo, R., & Vannini, I. (2010). La qualità educativa e didattica nella scuola dell'infanzia. Dalle procedure di valutazione alla formazione in servizio. *Ricerche di Pedagogia e Didattica*, 5, pp. 1-19.

Cambi, F., Catarsi, E., Colicchi, E., Fratini, C., & Muzi, M. (2003). *Le professionalità educative*. Roma: Carocci.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W.S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *The Teachers College Record*, 112(3), pp. 579-620.

Canevaro, A. (Ed.) (2007). *L'integrazione scolastica degli alunni con disabilità. Trent'anni di inclusione nella scuola italiana*. Trento: Erickson.

Cerini, G. (2012). *Una certa idea di valutazione: apprendimenti, insegnanti, scuole, sistema*. Faenza: Homeless Book.

Cicchetti, D., & Cohen, D.J. (1995). *Developmental psychopathology*. New York: Wiley.

Cittadinanzattiva observatory of prices and tariffs (2015). *C'è un nido? Rilevazione annuale sui costi e numero degli asili nido in Italia*. Roma: Cittadinanzattiva. Retrieved from: www.cittadinanzattiva.it.

Cochran, M. (2011). International perspectives on early childhood education. *Educational Policy*, 25(1), pp. 65-91.

Comrey, A.L., & Lee, H.B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Cosmai, L., & Caggio, F. (2008). Cercare e mantenere tempi distesi: un Nido con tempi di "ozio" per i bambini. In R. Zerbato (Ed.), *Infanzia: tempi di vita, tempi di relazione: Bambini e servizi educativi 0/6 tra continuità e cambiamento* (pp. 378-384). Bergamo: Edizioni Junior.

Coussée, F., Bradt, L., Roose, R., & Bouverne-De Bie, M. (2010). The emerging social pedagogical paradigm in UK child and youth care: Deus ex machina or walking the beaten path?. *British Journal of Social Work*, 40(3), pp. 789-805.

Creswell, J.W., & Plano Clark, V.L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

Dahlberg, G., Moss, P., & Pence, A. (1999). *Beyond quality in early childhood education and care: postmodern perspectives*. London: Falmer Press.

Dahlberg, G., Moss, P., & Pence, A. (2007). *Beyond Quality in Early Childhood Education and Care: Languages of Evaluation* (2nd ed.). London: Falmer Press.

Darder, P., & Mestres, J. (1994-2000). *ASEI. Autovalutazione dei servizi educativi per l'infanzia*. Italian translation and adaptation by M.P. Gusmini. Milano: FrancoAngeli.

Day, C. (1999). *Developing Teachers: The Challenges of Lifelong Learning*. London: Falmer Press.

Declercq, B., & Laevers, F. (2015). *Measuring 'pedagogical' quality with an observation tool: Insights from the pilot study*. Paper presented at the 25th EECERA Conference, Barcelona, Spain, September 7-10.

Dessus P., Cosnefroy, O., & Joët, G. (2014). *Evaluer la qualité des interactions enseignant-élèves en début de scolarité: Qualités psychométriques d'une adaptation française du CLASS*. Paper presented at the 82th Congres de l'Acfas, Colloque 519, Montréal, Canada, May 15.

Dewey, J., & Bentley, A.F. (1949). *Knowing and the Known*. Boston: Beacon Press.

Di Giandomenico, I., Musatti, T., & Picchio, M. (2008). Il ruolo della valutazione nella costruzione di un sistema integrato di servizi per l'infanzia. *Rassegna Italiana di Valutazione XII*, 40, pp. 89-106.

Douglas, F. (2004) *A critique of ECERS as a measure of quality in early childhood education and care*. Paper presented at the Questions of Quality: CECDE International Conference, Dublin, Ireland, September 23-25.

Downer, J., Sabol, T.J., & Hamre, B.K. (2010). Teacher–Child Interactions in the Classroom: Toward a Theory of Within and Cross-Domain Links to Children's Developmental Outcomes. *Early Education and Development*, 21(5), pp. 699-723.

D'Ugo, R. (2013). *La qualità della Scuola dell'infanzia. La scala di valutazione PraDISI*. Milano: FrancoAngeli.

D'Ugo, R., & Vannini, I. (2014). Osservare le prassi didattiche nella scuola dell'infanzia. Il PraDISI come opportunità di formazione e sostegno della professionalità. *Giornale Italiano della Ricerca Educativa*, 7(13), pp. 99-116.

Eccles, J.S., & Roeser, R.W. (1999). School and community influences on human development. In M.H. Bornstein & M.E. Lamb (Eds.), *Developmental Psychology: An advanced textbook* (4th ed., pp. 503-554). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Elias, M.J. (2003). *Academic and social-emotional learning*. Educational Practices Series 11. Paris: International Academy of Education and International Bureau of Education.

European Commission Childcare Network (1990). *Quality in Childcare Services – Seminar Report*. Brussels: European Commission Equal opportunities Unit.

European Commission Childcare Network (1996). *Quality Targets in Services for Young Children – Seminar Report*. Brussels: European Commission Equal opportunities Unit.

European Commission/EACEA/Eurydice/Eurostat (2014). *Key Data on Early Childhood Education and Care in Europe*. Luxembourg: Publications Office of the European Union.

Everitt, B.S. (1975). Multivariate analysis: The need for data, and other problems. *The British Journal of Psychiatry*, 126(3), pp. 237-240.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, 4, pp. 272-299.

Farquhar, S. (1999) *Research and the Production of 'Worthwhile' Knowledge about Quality in Early Years Education*. Paper presented at the AARE-NZARE (Australian Association for Research in Education-New Zealand Association for Research in Education) conference, Melbourne, Australia, November 29-December 2.

Fenech, M. (2011). An Analysis of the Conceptualisation of 'Quality' in Early Childhood Education and Care Empirical Research: Promoting 'Blind Spots' as Foci for Future Research. *Contemporary Issues in Early Childhood*, 12(2), pp. 102-117.

Ferrari, M. (1994). La qualità del contesto educativo: strumenti e metodologie valutative. In M. Ferrari (Ed.), *La valutazione di contesti prescolari: strumenti e realtà* (pp. 16-33). Bergamo: Edizioni Junior.

Ferrari, M. (2002). La qualità negoziata: il percorso di elaborazione delle Idee guida del nido d'infanzia. In E. Becchi, A. Bondioli & M. Ferrari (Eds.), *Il progetto pedagogico del nido e la sua valutazione. La qualità negoziata* (pp. 27-40). Bergamo: Edizioni Junior.

Ferrari, M. (2004a). ISQUEN. In M. Ferrari & A. Bondioli (Eds.), *Verso un modello di valutazione formativa. Ragioni, strumenti e percorsi* (pp. 157-167). Bergamo: Edizioni Junior.

Ferrari, M. (2004b). Restituire. In M. Ferrari & A. Bondioli (Eds.), *Verso un modello di valutazione formativa. Ragioni, strumenti e percorsi* (pp. 59-91). Bergamo: Edizioni Junior.

Ferrari, M. (2013). Dispositivi di analisi della qualità nei contesti educativi per l'infanzia: per un'esperienza di valutazione formativa "interculturale" e dialogica. *RELAdEI. Revista Latinoamericana de Educación Infantil*, 2(2), pp. 35-51.

Ferraro, J.M. (2000). *Reflective practice and professional development*. ERIC Digest, No: ED449120.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage publications.

Ford, D.H., & Lerner, R.M. (1992). *Developmental systems theory: An integrative approach*. Newbury Park, CA: Sage.

Fortunati, A. (2007). Quality. In R.S. New & M. Cochran (Eds.), *Early childhood education: An international encyclopedia*, vol. 4 (pp. 1122-1125). Westport: Praeger.

Fortunati, A., & Tognetti, G. (2003). Storie di vita quotidiana. Bambini, educatori e genitori nei nidi e negli altri servizi educativi per l'infanzia e la famiglia di San Miniato. In L. Gandini, S. Mantovani & C.P. Edwards (Eds.), *Il nido per una cultura dell'infanzia*. Bergamo: Edizioni Junior.

Franchi, L., & Caggio, F. (Eds.) (1999). *Per una cultura della qualità nell'asilo nido e nella scuola materna. Riflessioni, criteri, parametri di verifica e valutazioni*. Bergamo: Edizioni Junior.

Fyfe, B. (1994). Images from the United States: Using ideas from Reggio Emilia experience with American educators. In L.G. Katz & B. Cesarone (Eds.), *Reflections on the Reggio Emilia approach* (pp. 19-32). Urbana, IL: ERIC/EECE.

Gagné, P., & Hancock, G.R. (2006). Measurement model quality, sample size, and solution propriety in confirmatory factor models. *Multivariate Behavioral Research*, 41(1), pp. 65-83.

Gandini, L. (1995). Uno spazio che riflette una cultura dell'infanzia. In C. Edwards, L. Gandini & G. Forman (Eds.), *I cento linguaggi dei bambini* (pp. 235-244). Bergamo: Edizioni Junior.

Gariboldi, A., Babini, M.P., & Vannini, L. (Eds.) (2014). *La qualità nella scuola: un percorso di valutazione formativa*. Parma: Edizioni Junior-Spaggiari Edizioni.

Gay, G. (2002). Preparing culturally responsive teaching. *Journal of Teacher Education*, 53, pp. 106-116.

Gillen, J., Cameron, C.A., Sombat Tapanya, C., Pinto, G.D., Hancock, R.A., Young, S., & Accorti Gamannossi, B. (2007). 'A day in the life': advancing a methodology for the cultural study of development and learning in early childhood. *Early Child Development and Care*, 177(2), pp. 207–218.

Gorsuch, R.L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

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

Gregory, A., & Weinstein, R.S. (2004). Connection and regulation at home and in school: Predicting growth in achievement for adolescents. *Journal of Adolescent Research*, 19, pp. 405-427.

Guba, E., & Lincoln, Y. (1989). *Fourth Generation Evaluation*. Newbury Park, CA: Sage.

Guilford, J.P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw-Hill.

Gusmini, M.P. (2004). Riflettere. In M. Ferrari & A. Bondioli (Eds.), *Verso un modello di valutazione formativa. Ragioni, strumenti e percorsi* (pp. 93-143). Bergamo: Edizioni Junior.

Hafen, C.A., Hamre, B.K., Allen, J.P., Bell, C.A., Gitomer, D.H., & Pianta, R.C. (2015). Teaching through interactions in secondary school classrooms revisiting the factor structure and practical application of the classroom assessment scoring system–secondary. *The Journal of Early Adolescence*, 35(5-6), pp. 651-680.

Halle, T., Vick Whittaker, J.E., & Anderson, R. (2010). *Quality in Early Childhood Care and Education Settings: A Compendium of Measures* (2nd ed.). Washington, DC: Child Trends.

Hamre, B.K., Goffin, S., & Kraft-Sayre, M. (2009). *Classroom Assessment Scoring System (CLASS) Implementation Guide: Measuring and improving classroom interactions in early childhood settings*. Charlottesville: Teachstone.

Hamre, B.K, La Paro, K.M., & Pianta, R.C. (2014). *Classroom Assessment Scoring System Infant Manual*. Baltimore: Brookes Publishing.

Hamre, B.K., & Pianta, R.C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, 72, pp. 625-638.

Hamre, B.K., & Pianta, R.C. (2006). Student-teacher relationships. In G.C. Bear & K.M. Minke (Eds.), *Children's needs III: Development, prevention, and intervention* (pp. 59-71). Washington, DC: National Association of School Psychologists.

Hamre, B.K., & Pianta, R.C. (2007). Learning opportunities in preschool and early elementary classrooms. In R.C. Pianta, M. Cox & K. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49-84). Baltimore: Brookes.

Hamre, B.K., Pianta, R.C., Downer, J.T., DeCoster, J., Mashburn, A.J., Jones, S.M., Brown, J.L., Cappella, E., Atkins, M., Rivers, S.E., Brackett, M.A., & Hamagami, A. (2013). Teaching through Interactions: Testing a Developmental Framework of Teacher Effectiveness in over 4,000 Classrooms. *The Elementary School Journal*, 113(4), pp. 461-487.

Hamre, B.K., Pianta, R.C., Mashburn, A.J., & Downer, J.T. (2007). *Building a science of classrooms: Application of the CLASS framework in over 4,000 U.S. early childhood and elementary classrooms*. New York: Foundation for Child Development.

Harms, T., & Clifford, R.M. (1980). *Early Childhood Environment Rating Scale*. New York: Teachers College Press.

Harms, T., Clifford, R.M., & Cryer, D. (1998). *Early Childhood Environment Rating Scale – Revised Edition*. New York, NY: Teachers College Press.

Harms, T., Clifford, R.M., & Cryer, D. (2005). *Early Childhood Environment Rating Scale - Revised Edition*. New York, NY: Teachers College Press.

Harms T., Cryer, D., & Clifford, R.M. (1990-1992), *Scala per la Valutazione dell'Asilo Nido*. Italian translation and adaptation by M. Ferrari & P. Livraghi. Milano: Franco Angeli

Harms, T., Cryer, D., & Clifford, R.M. (2003). *Infant/Toddler Environment Rating Scale - Revised Edition*. New York, NY: Teachers College Press.

High/Scope Educational Research Foundation (2003). *Preschool Program Quality Assessment* (2nd ed.). Ypsilanti, MI: High/Scope Press.

Howes, C., Burchinal, M., Pianta, R.C., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*, 23, pp. 27-50.

Hu, B.Y., Fan, X., Gu, C., & Yang, N. (2016). Applicability of the Classroom Assessment Scoring System in Chinese preschools based on psychometric evidence. *Early Education and Development*, 27(5), pp. 714-734.

Hu, L., & Bentler, P.M. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, pp. 1-55.

Hyson, M.C., Hirsh-Pasek, K., & Rescorla, L. (1990). The classroom practices inventory: an observation instrument based on NAEYC's guidelines for developmentally appropriate practices for 4- and 5-year-old children. *Early Childhood Research Quarterly*, 5, pp. 475-494.

Ishimine, K., & Tayler, C. (2014). Assessing Quality in Early Childhood Education and Care. *European Journal of Education*, 49(2), pp. 272-290.

Istat (2012). *La scuola e le attività educative*. Retrieved from http://www.istat.it/it/archivio/71706.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, pp. 141-151.

Kass, R.A., & Tinsley, H.E.A. (1979). Factor analysis. *Journal of Leisure Research*, 11, pp. 120−138.

Kline, P. (1979). *Psychometrics and psychology*. London: Academic Press.

Kline, R.B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Koomen, H.M.Y., Verschueren, K., & Thijs, J.T. (2006). Assessing aspects of the teacher-child relationship: A critical ingredient of a practice-oriented psycho-diagnostic approach. *Educational and Child Psychology*, 23, pp. 50-60.

La Paro, K.M., Hamre, B.K., & Pianta, R.C. (2012). *Classroom Assessment Scoring System Toddler Manual*. Charlottesville: Teachstone.

La Paro, K.M., Pianta, R.C., & Stuhlman, M. (2004). The classroom assessment scoring system: findings from the pre-kindergarten year. *The Elementary School Journal*, 104, pp. 409-426.

Lazzari, A., Picchio, M. & Musatti, T. (2013). Sustaining ECEC quality through continuing professional development: systemic approaches to practitioners' professionalisation in the Italian context. *Early Years*, 33(2), pp. 133-145.

Lerner, R.M. (1998). Theories of human development: Contemporary perspectives. In W. Damon & R.M. Lerner (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (5th ed., pp. 1-24). New York: Wiley.

Leyva, D., Weiland, C., Barata, M., Yoshikawa, H., Snow, C., Treviño, E., & Rolla, A. (2015). Teacher–child Interactions in Chile and Their Associations with Prekindergarten Outcomes. *Child development*, 86(3), pp. 781-799.

Limlingan, M.C. (2011). On the Right Track: Measuring Early Childhood Development Program Quality Internationally. *Current Issues in Comparative Education*, 14, pp. 38-47.

López Boo, F., Araujo, M.C., & Tomé, R. (2016). *How is child care quality measured? A toolkit*. Washington, DC: Inter-American Development Bank (IDB).

MacCallum, R.C., Widaman, K.F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods*, 4(1), 84-99.

Malaguzzi, L. (1993). For an Education Based on Relationships. *Young Children*, 49(1), 9-12.

Malaguzzi, L. (1994). Your image of the child: Where teaching begins. *Child Care Information Exchange*, 96, pp. 52-61.

Malaguzzi, L. (1998). History, ideas, and basic philosophy: An interview with Lella Gandini. In C. Edwards, L. Gandini & G. Forman (Eds.), *The hundred languages of children: The Reggio Emilia approach – advanced reflections* (2nd ed., pp.49-98). Westport, CT: Ablex.

Maldonado-Carreño, C., & Votruba-Drzal, E. (2014). *Early childhood development and preschool quality in Bogota, Colombia*. Unpublished manuscript, Universidad de los Andes, Bogota, Colombia.

Mantovani, S. (2007a). Early Childhood Education in Italy. In R.S. New & M. Cochran (Eds.), *Early childhood education: An international encyclopedia*, vol. 4 (pp. 1110-1115). Westport: Praeger.

Mantovani, S. (2007b). Pedagogy in Italian Early Childhood Education. In R.S. New & M. Cochran (Eds.), *Early childhood education: An international encyclopedia*, vol. 4 (pp. 1115-1118). Westport: Praeger.

Mantovani, S. (2014). I nidi e le scuole dell'infanzia in Italia. Patrimonio antico e parole nuove. *Bambini*, 3, pp. 21-29.

Marcuccio, M., & Zanelli, P. (Eds.) (2013). *Sguardi sul nido... Strumento per lo Sviluppo di Processi Riflessivi e Indagini valutative nei Nidi da parte dei*

Gruppi di lavoro educativi (SPRING). Parma: Edizioni Junior-Spaggiari Edizioni.

Mashburn, A.J., Hamre, B.K., Downer, J.T., & Pianta, R.C. (2007). Teacher and classroom characteristics associated with teachers' ratings of prekindergartners' relationships and behaviors. J*ournal of Psychoeducational Assessment*, 24(4), pp. 367-380.

Mashburn, A.J., Pianta, R.C., Hamre, B.K., Downer, J.T., Barbarin, O.A., Bryant, D., Burchinal, M., Early, D., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79, pp. 732-749.

Mathers, S., Linskey, F., Seddon, J., & Sylva, K. (2007). Using quality rating scales for professional development: Experiences from the UK. *International Journal of Early Years Education*, 15(3), pp. 261-274.

Montessori, M. (1912) *The Montessori Method*. New York: Frederick A. Stokes Company.

Montessori, M. (1949). *The absorbent mind*. Madras: Adyar.

Montessori, M. (2004). *The discovery of the child: Revised and Enlarged Edition of The Montessori Method*. Delhi: Aakar books.

Montie, J.E., Xiang, Z., & Schweinhart, L.J. (2006). Preschool experience in 10 countries: Cognitive and language performance at age 7. *Early Childhood Research Quarterly*, 21, pp. 313-331.

Morse, J.M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing research*, 40, pp. 120-123.

Moss, P. (1994). Defining quality: values, stakeholders and processes. In P. Moss & A. Pence (Eds) (1994) Valuing quality in early childhood services. London: Paul Chapman (pp 1-9).

Musatti, T. (2007). Socioemotional Development in Italian Early Childhood Education. In R.S. New & M. Cochran (Eds.), *Early childhood education: An international encyclopedia*, vol. 4 (pp. 1133-1136). Westport: Praeger.

Musatti, T., & Picchio, M. (2010). Early education in Italy: Research and practice. *International Journal of Early Childhood*, 42 (2), pp. 141-153.

Muthén, L.K., & Muthén, B.O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9, pp. 599-620.

New, R.S. (1995). Differenze culturali nella definizione di una prassi educativa appropriata allo sviluppo infantile: Una sfida per la teoria e per la pratica. In C. Edwards, L. Gandini & G. Forman (Eds.), *I cento linguaggi dei bambini* (pp. 275-281). Bergamo: Edizioni Junior.

NICHD Early Child Care Research Network (2000). The relation of childcare to cognitive and language development. *Child Development*, 71, pp. 960-980.

OECD (2012). *Starting Strong III: A Quality Toolbox For Early Childhood Education and Care*. Paris: OECD Publishing.

OECD (2013a). How Do Early Childhood Education and Care (ECEC) Policies, Systems and Quality Vary Across OECD Countries?. *Education Indicators in Focus (EDIF)*, 11. Retrieved from http://www.oecd.org/edu/skills-beyond-school/EDIF11.pdf.

OECD (2013b). *Literature review on Monitoring Quality in Early Childhood Education and Care (ECEC)*. Retrieved from http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/EDPC/ECEC(2013)3anddoclanguage=en.

OECD (2014). *Education at a glance*. Paris: OECD Publishing.

OECD (2015). *Starting Strong IV: Monitoring Quality in Early Childhood Education and Care*. Paris: OECD Publishing.

Osterman, K.F. (1990). Reflective practice. *Education and Urban Society*, 22(2), pp. 133-153.

Pakarinen, E., Lerkkanen, M.K., Poikkeus, A.M., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., & Nurmi, J.E. (2010). A validation of the Classroom

Assessment Scoring System in Finnish kindergartens. *Early Education and Development*, 21(1), pp. 95-124.

Parente, M. (2015). Orientamenti di qualità nello sviluppo dei servizi educativi per la prima infanzia. In A. Mariani (Ed.), *Cultura della qualità nei servizi educativi per la prima infanzia* (pp. 49-69). Trento: Erickson.

Pastori, G., & Mantovani, S. (2016). Study 5: A Cultural Analysis of ECEC Quality in 7 Countries. In P.L. Slot, J. Cadima, J. Salminen, G. Pastori & M.K. Lerkkanen (Eds.). *D2.3: Multiple case study in seven European countries regarding culture-sensitive classroom quality assessment*. Working paper, version: March 31, 2016.

Pastori, G., Mantovani, S., Braga, P., & Pagani, V. (2016). *WP2 Case Study. Qualitative and Cultural Study and Analysis*. Working paper, version: February 18, 2016.

Pastori, G., & Pagani, V. (*forthcoming*). Is validation always valid? Cross cultural complexities of standard-based instruments migrating out of their context. *European Early Childhood Education Research Journal*.

Pastori, G., Pagani, V., & Mantovani, S. (2015). *Is validation always valid? Cross cultural complexities of standard-based instruments migrating out of their context. A Study on CLASS Pre-K in Italy (3-6 ECEC services)*. Paper presented at the 25th EECERA Conference, Barcelona, Spain, September 7-10.

Pastori, G., Pagani, V., & Mantovani, S. (2016). *A critical cultural approach to CLASS. The voice of Italian ECEC teachers*. Paper presented at the EARLI Conference on SIG5 (Learning and Development in Early Childhood), Porto, Portugal, June 29-July 1.

Peeters, J., & Vandenbroeck, M. (2011). Child care practitioners and the process of professionalization. In L. Miller & C. Cable (Eds.), *Professional issues, leadership and management in the early years* (pp. 62-76). London: Sage.

Perry, L.B., & Southwell, L. (2011). Developing intercultural understanding and skills: Models and approaches. *Intercultural Education*, 22(6), pp. 453-466.

Pianta, R.C. (1994). Patterns of relationships between children and kindergarten teachers. *Journal of School Psychology*, 32(1), pp. 15–31.

Pianta, R.C. (1997). Adult-Child Relationship Processes and Early Schooling. *Early Education and Development*, 8(1), p. 11-26

Pianta, R.C. (1999). *Enhancing relationships between children and teachers*. Washington, DC: American Psychological Association.

Pianta, R.C. (2001). *Student–Teacher Relationship Scale. Professional manual*. Lutz, FL: Psychological Assessment Resources.

Pianta, R.C., Barnett, W.S., Burchinal, M., & Thornburg, K.R. (2009). The Effects of Preschool Education What We Know, How Public Policy Is or Is Not Aligned With the Evidence Base, and What We Need to Know. *Psychological Science in the Public Interest*, 10(2), pp. 49-88.

Pianta, R.C., & Hamre, B.K. (2009a). A Lot of Students and Their Teachers Need Support: Using a Common Framework to Observe Teacher Practices Might Help. *Educational Researcher*, 38(7), pp. 546-548.

Pianta, R.C., & Hamre, B.K. (2009b). Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity. *Educational Researcher*, 38(2), pp. 109-119.

Pianta, R.C., Hamre, B.K., & Stuhlman, M.W. (2003). Relationships Between Teachers and Children. In W.M. Reynolds & G.E. Miller (Eds.), *Educational Psychology, Vol. 7 of Handbook of Psychology* (pp. 199-234). New York: Wiley.

Pianta, R.C., La Paro, K.M., & Hamre, B.K. (2008). *Classroom Assessment Scoring System Pre-K Manual*. Baltimore: Brookes Publishing.

Picchio, M., Di Giandomenico, I., & Musatti, T. (2014). The use of documentation in a participatory system of evaluation. *Early Years*, 34(2), pp. 133-145.

Pontecorvo, C. (1991). Scuola dell'infanzia: un ambiente per l'apprendimento. In B. Astorino, G. Cerini & M. Mozzetti (eds.). *Cominciare da tre. La nuova scuola del bambino* (2nd ed., pp. 71-81). Milano: Franco Angeli.

Pressley, M., Roehrig, A., Raphael, L., Dolezal, S., Bohn, C., Mohan, L., Wharton-McDonald, R., Bogner, K., & Hogan, K. (2003). Teaching processes in elementary and secondary education. In W. Reynolds & G. Miller (Eds.), *Handbook of Psychology: Volume 7 Educational Psychology* (pp. 153-176). New York: Wiley.

Region Tuscany & Istituto degli Innocenti (2006). *La qualità dei servizi educativi per la prima infanzia: il nuovo sistema di valutazione dei nidi e dei servizi educativi integrativi*. Firenze: Regione Toscana.

Rinaldi, C. (1993). The emergent curriculum and social constructivism. In C. Edwards, L. Gandini & G. Forman (Eds), *The Hundred Languages of Children* (pp. 101-111). Norwood, NJ: Ablex.

Rogoff, B. (2003). *The cultural nature of human development*. Oxford: University Press.

Salminen, J., Lerkkanen, M.K., Poikkeus, A.M., Pakarinen, E., Siekkinen, M., Hännikäinen, M., Poikonen, P., & Rasku-Puttonen, H. (2012). Observed classroom quality profiles of kindergarten classrooms in Finland. *Early Education and Development*, 23(5), pp. 654-677.

Sameroff, A.J. (1995). General system theory and developmental psychopathology. In D. Cicchetti & D.J. Cohen (Eds.), *Developmental psychopathology: Vol. 1. Risk, disorder, and adaptation* (pp. 659–695). New York: Wiley.

Sandilos, L.E., DiPerna, JC., & Family Life Project Key Investigators (2014). Measuring Quality in Kindergarten Classrooms: Structural Analysis of the Classroom Assessment Scoring System (CLASS K–3). *Early Education and Development*, 25(6), pp. 894-914.

Sarsini D., & Di Bari, C. (2015). Competenze e professionalità degli educatori/insegnanti. In A. Mariani (Ed.), *Cultura della qualità nei servizi educativi per la prima infanzia* (pp. 71-90). Trento: Erickson.

Sartorio, S., & Nigito, G. (2005). Valutare nei servizi per l'infanzia. *Bambini*, 21, pp. 32-39.

Savio, D. (2008). Promuovere autovalutazione formativa: il percorso degli Asili Nido della Valle d'Aosta per la definizione di linee guida. In R. Zerbato (Ed.), *Infanzia: tempi di vita, tempi di relazione: Bambini e servizi educativi 0/6 tra continuità e cambiamento* (pp. 202-211). Bergamo: Edizioni Junior.

Savio, D. (2015). Un'inchiesta sulla cultura della valutazione nei servizi educativi per l'infanzia. In A. Bondioli & D. Savio (Eds.), *La valutazione di contesto nei servizi per l'infanzia italiani. Riflessioni ed esperienze* (pp. 77-104). Parma: Edizioni Junior- Spaggiari Edizioni.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, 8, pp. 23-74.

Schleicher, A. (2011). *Building a high-quality teaching profession: Lessons from around the world*. Paris: OECD Publishing.

Schön, D. (1993). *Il professionista riflessivo*. Bari: Dedalo.

Sheridan, S. (2007). Dimensions of pedagogical quality in preschool. *International Journal of Early Years Education*, 15(2), pp. 197-217.

Slot, P.L. (2014). *Early Childhood Education and Care in the Netherlands: Quality, Curriculum, and Relations with Child Development*. PhD diss., Universiteit Utrecht.

Slot, P.L., Cadima, J., Salminen, J., Pastori, G., & Lerkkanen, M.K. (2016). *D2.3: Multiple case study in seven European countries regarding culture-sensitive classroom quality assessment*. Working paper, version: March 31, 2016.

Snow, C.E., & Van Hemel, S.B. (2008). *Early Childhood Assessment: Why, What, and How*. Washington, DC: The National Academies Press.

Staccioli, G. (2008). I tempi dell'impermanenza: ritmi di scuola e ritmi di vita. In R. Zerbato (Ed.), *Infanzia: tempi di vita, tempi di relazione: Bambini e servizi educativi 0/6 tra continuità e cambiamento* (pp. 378-384). Bergamo: Edizioni Junior.

Stipek, D., & Byler, P. (2004). The early childhood classroom observation measure. *Early Childhood Research Quarterly*, 19, pp. 375-397.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2004). *The Effective Provision of Pre-School Education (EPPE) Project: The Final Report: Effective Pre-School Education*. London: Institute of Education.

Sylva, K., Roy, C., & Painter, M. (1980). *Child Watching at Playgroup and Nursery School*. London: Grant McIntyre.

Tabachnick, B.G., & Fidell, L.S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn and Bacon.

Tanaka, J.S. (1993). Multifaceted conceptions of fit in structural equation models. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 10-39). Newbury Park, CA: Sage.

Tayler, C., Ishimine, K., Cloney, D., Cleveland, G., & Thorpe, K. (2013). The quality of early childhood education and care services in Australia. *Australasian Journal of Early Childhood*, 38(2), pp.13-21.

Teruggi, L. (2007). Literacy in Italian Early Childhood Education. In R.S. New & M. Cochran (Eds.). *Early childhood education: An international encyclopedia*, vol. 4 (pp. 1130-1133.). Westport, CT: Praeger.

Tobin, J.J. (2005). Quality in early childhood education: An anthropologist's perspective. *Early Education and Development*, 16(4), pp. 421-434.

Tobin, J.J., Hsueh, Y., & Karasawa, M. (2009). *Preschool in three cultures revisited. China, Japan and the United States*. Chicago, IL: The University of Chicago Press.

Tobin, J.J., Wu, D.Y., & Davidson, D.H. (1989). *Preschool in three cultures: Japan, China, and the United States*. Newhaven, CT: Yale University Press.

Treviño, E., Toledo, G., & Gempp, R. (2013). Calidad de la educación parvularia: las prácticas de clase y el camino a la mejora. Pensamiento

Educativo. *Revista de Investigación Educacional Latinoamericana*, 50(1), pp. 40-62.

UNICEF (2008). *The Child Care Transition. A league table of early childhood education and care in economically advanced countries*. Innocenti Report Card 8. Florence: UNICEF Innocenti Research Centre.

Vandell, D., & Wolfe, B. (2000). *Child care quality: Does it matter and does it need to be improved?* Madison, WI: Institute for Research on Poverty.

Vandenbroeck, M. (2010). Participation in ECEC Programs. Equity, Diversity and Educational Disadvantage. In E. Baker, B. McGaw, & P. Peterson (Eds.). *International Encyclopedia of Education* (3rd ed., pp. 81-85). Oxford: Elsevier.

Vandenbroeck, M., & Peeters, J. (2014). Democratic experimentation in early childhood education. In G. Biesta, M. De Bie & D. Wildemeersch (Eds.) *Civic Learning, Democratic Citizenship and the Public Sphere* (pp. 151-165). Dordrecht: Springer.

Vannini, L. (2014). Il percorso di Riflessione della Qualità – Ri.Qua. In A. Gariboldi, M.P. Babini & L. Vannini (Eds.), *La qualità nella scuola: un percorso di valutazione formativa* (pp. 41-104). Parma: Edizioni Junior-Spaggiari Edizioni.

Varin, D. (2007). L'esperienza precoce ed estesa di asili nido: fattori di facilitazione per lo sviluppo e aspetti di rischio. *Psicologia Clinica dello Sviluppo*, 11(3), pp. 359-384.

Velicer, W.F., & Fava, J.L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 3, pp. 231-251.

Vitiello, V.E. (2013). *Dual Language Learners and the CLASS Measure: Research and Recommendations*. Charlottesville, VA: Teachstone.

von Suchodoletz, A., Fäsche, A., Gunzenhauser, C., & Hamre, B.K. (2014). A typical morning in preschool: Observations of teacher–child interactions in German preschools. *Early Childhood Research Quarterly*, 29(4), pp. 509-519.

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

Wood D.J., McMahon, L., & Cranstoun, Y. (1980). *Working with Under Fives*. London: Grant McIntyre.

Zanelli, P. (1998). *La qualità come processo. L'esperienza dei nidi forlivesi*. Milano: FrancoAngeli.

Zanelli, P., Sagginati, B., & Fabbri, E. (Eds.). (2004). *Autovalutazione come risorsa*. Bergamo: Edizioni Junior.

Ziv, Y., & Aviezer, O. (2014). Using the Classroom Assessment Scoring System (CLASS) in Israeli Kindergarten for evaluating and shaping classroom climate. *Hed Hagan*, 78(3), pp. 22-29.