LITERATURE REVIEW ON MONITORING QUALITY IN EARLY CHILDHOOD EDUCATION AND CARE (ECEC)

Revised and adapted version prepared for presentation and discussion at the 4th meeting of the Trans-Atlantic Forum on Inclusive Early Years, June 30th – July 2nd, 2014 - Amsterdam, the Netherlands.

Author:

Ineke Litjens
Analyst
Early Childhood and Schools Division
Directorate for Education and Skills
ineke.litjens@oecd.org || www.oecd.org/edu/earlychildhood

2, rue André Pascal - 75775 Paris Cedex 16 || Tel: +33 1 45 24 86 19

FOREWORD

This literature review has been prepared by Ineke Litjens, Analyst at the OECD Education and Skills Dircetorate, with the support of Kelly Makowiecki. The version presented at the Trans The paper has been sent out for feedback and comments to the bureau of the OECD ECEC Network whose feedback has been implemented in this document. In addition, this document reflects feedback from the following researchers, policy-makers and experts (in alphabetical order): Sandra Antulic (Croatia); Tijne Berg-Le Clercq (Netherlands); Elisabeth Bjornestad (Norway); Megan Carolan (United States); Stephanie Currenton (United States); Ramon Flecha (Spain); Ruben Fukkink (Netherlands); Olivia Maria Herrera (Chile); Kari Jacobsen (Norway); Heajin Jang (Korea); Christa Japel (United States); Bente Jensen (Denmark); Leslie Kopf-Johnson (Canada); Maura Lantrua (United Kingdom); Lasse Lipponen (Finland); Thomas Moser (Norway); Michal Perlman (Canada); Ingrid Pramling Samuelsson (Sweden); Sonja Sheridan (Sweden); Dongju Shin (Korea); Michel Vandenbroeck (Belgium); Petr Varmuza (Canada); Annemiek Veen (Netherlands); Malva Villalon (Chile); Jissok Yeom (Korea); Henrik Daae Zachrisson (Norway); Zhizong Zuo (China).

TABLE OF CONTENTS

Tables

Boxes

LITERATURE REVIEW ON MONITORING QUALITY IN EARLY CHILDHOOD EDUCATION AND CARE (ECEC)

1. This paper is part of the OECD Early Childhood and Schools (ECS) project '*OECD Review of Policies and Practices for Monitoring and Evaluating Quality in Early Learning and Development*' and contributes to the knowledge base supporting the project. The project aims to encourage dialogue on monitoring quality in early childhood education and care (ECEC); to stimulate further thinking and action at policy level about the best ways to monitor different quality aspects; and to explore how countries can put in place policies to enhance monitoring systems to improve quality and children's early development.

2. For the purpose of this overview and project, the definition of ECEC is, in co-ordination with the Starting Strong reports (OECD 2001, 2006 and 2012a), defined as all arrangements providing care and education for children under compulsory school age, regardless of setting, funding, opening hours, or programme content (OECD, 2001).

3. The definition of staff, as used in this paper, refers to all staff members working directly with children in ECEC centres as well as leaders and managers of ECEC provisions. Staff includes the terms practitioners, child carers, care givers, teachers, professionals, instructors etc. These terms are used interchangeably in this document. When referring to managers and leaders, both refer to those people who are in charge of running an ECEC centre. The terms 'service', 'provision' and 'setting' all have the same meaning in this report and refer to any centre providing ECEC services. These terms are used interchangeably as well.

1. ABSTRACT

This report discusses the different monitoring policies and practices in place. The focus in this review lies on monitoring four topics: i) service quality; ii) staff quality; iii) child development and outcomes and; iv) curriculum implementation. Initially, the report provides an overview of what monitoring practices are in place in OECD countries, based on the Starting Strong III report and OECD Early Childhood Education and Care (ECEC) Network discussions, for what purposes these practices are used and how stakeholders use the monitoring results. Country examples are included when these were available to clarify practices and users. The paper then reviews the literature related to the effects or impacts of monitoring on quality and child development, and discusses design and implementation aspects based on the literature findings.

Monitoring and evaluating quality is, in general, perceived as an important component of enhancing quality in ECEC systems. It can point to strengths and weaknesses of systems and programmes, provide incentives for improvement of standards and practices, and can assist staff in enhancing children's early development. Research findings on the effects of monitoring in ECEC indicate some positive effects of monitoring on staff behaviour and practices as well as curriculum implementation, which can positively impact child development, and monitoring can lead to an overall increase in service quality. Findings differ according to practice or method used whereby narrative assessment techniques were indicated to be more appropriate than testing for a holistic evaluation of child outcomes since it does not only include outcomes of development, but also inputs by staff and parents. However, research also points out it is difficult to separate and identify the impact of a single tool or method since quality is a result of different aspects and experiences within and beyond the ECEC institution. It is difficult to attribute causality between a monitoring process or practice and quality.

When designing and implementing monitoring policies and practices of quality, literature indicates that it is relevant to define quality, ensure the practices are appropriate to meet the purpose(s), and that they are useable and are linked to actual practices. In addition, the involvement of stakeholders is regarded as important; training evaluators and ECEC staff and management on implementation of monitoring systems is key to success; and monitoring practices of child development should be designed to be developmentally appropriate. Lastly, piloting the system or practice and the dissemination of monitoring results with stakeholders and the general public can support successful implementation of monitoring systems.

2. INTRODUCTION

4. With increasing investments, both public and private, in early childhood education and care (ECEC), it is becoming critical to know whether ECEC systems are delivering high-quality services. To understand how an ECEC system performs is not only important for accountability purposes, but also for (improvement of) policy design and implementation and for informing parents of what level of quality is being offered (Levitt, Janta and Wegrich, 2008). By evaluating inputs or outputs, monitoring systems intend to create incentives for improved quality and performance, and possibly identify ‘‘underperforming’’ settings for remediation (Booher-Jennings, 2007). Increasingly, countries are developing a range of tools and techniques for monitoring and evaluating the quality of ECEC systems as part of their efforts to enhance quality as well as early child development.

5. Quality encompasses all the features of children’s environments and experiences that are presumed to be beneficial to their well-being. This includes the use of a curriculum, staff characteristics, teacher or caregiver behaviours and practices, and staff-child interactions which form the core of children's ECEC experiences, often referred to as process quality. In addition, quality involves structural features of the setting such as space, group size, and safety standards (NCES, 1997; OECD, 2006; OECD, 2012a). However, quality is not a static concept. Definitions of quality may differ between countries since this is a value- and cultural-based concept and any definition of quality is subject to change over time (Kamerman, 2001).

6. Different aspects related to "quality" can be monitored, ranging from service quality in general to staff performance (or staff quality). This report focuses on the monitoring systems of four aspects of quality, as in line with the scope of Strand 1 of the new programme of work and budget for ECEC, 2013/14:

1. Service quality
2. Staff quality
3. Child outcomes and development
4. Implementation of curriculum

7. Some countries monitor all of the above, or one (or more) of the aspects are integrated into the monitoring tool of another one. For example, curriculum implementation can be monitored when evaluating quality at a more general service level, or when assessing staff performance. Therefore, monitoring one aspect of ECEC quality does not exclude another.

8. The focal point of this review are the monitoring systems and tools used in OECD countries for the four aspects mentioned above, and what research tells us about the effectiveness of these systems and tools. More specifically, this report aims to answer three broad questions:

- What monitoring systems are in place in OECD countries, how are these organised, who are the users, and how do the users use the results?

- What is the empirical evidence on the effects of monitoring on quality and child development?

- What are the implications for policy design and implementation?

9. The literature depicted has been derived from four main sources: the Starting Strong reports (OECD 2001, 2006 and 2012a); empirical research literature which is primarily from Anglo-Saxon countries; policy analysis literature; and other research and policy documents relevant to monitoring quality and child development, much of which comes from Western European and Anglo-Saxon countries. In addition, documents produced by the OECD ECEC Network's thematic working groups have been used to list country examples or provide background information. A number of government and government-sponsored websites were surveyed to identify monitoring practices and policies. The literature used was found through a broad search in bibliographic databases and search engines such as *ERIC, Google Scholar, JStor,* and *ScienceDirect,* and through "snow-balling" of references. The search was conducted in English, and to a lesser extent in French. We aimed at including studies on the effects of monitoring that meet the following criteria: having research methods, analysis and findings sufficiently detailed to provide a sound basis for judgement about the robustness of the conclusions, and information about the characteristics of the context.

10. This paper does not intend to provide a comprehensive overview of *all* existing research findings or policy papers, neither is it a general overview of how OECD countries conduct monitoring practices. There are aspects of quality which might be monitored in countries that are not included in this review, such as monitoring parental satisfaction, since these aspects are beyond the scope of the new project. Most research on monitoring has been conducted on ECEC services for preschool-age children. While research has been conducted on ECEC services for infants and toddlers, the availability of papers, research and information on monitoring in settings for young children is less exhaustive, although we attempted to address both fields in this paper. Additionally, most papers and reports focus on centre-based care and education, so family-based settings are therefore under-represented in this report.

11. Chapter 3 defines the different monitoring practices and policies and how monitoring methods are organised. In Chapter 4, an overview is given on the different users and uses of monitoring. Chapter 5 reviews the empirical evidence on the effects or impacts of monitoring practices for accountability and improvement purposes. Chapter 6 discusses the implications for policy design and implementation. A final conclusion is offered in Chapter 7.

3. MONITORING PRACTICES AND POLICIES

12. The term *monitoring* can refer to both *evaluation* and *assessment* practices. According to education literature, assessment is used to refer to the process of deciding, collecting and making judgments about evidence relating to the achievement of children and staff. Evaluation is used for the process of deciding, collecting and making judgments about systems, programmes, materials, procedures and processes (Harlen, 2007). Consequently, assessment encompasses child development assessments as well as standardised tests and observations of staff performance, while evaluation encompasses practices such as inspections, self-evaluations and targeted programme evaluation. This report focuses on monitoring, including assessment *and* evaluation practices. Similar practices may be in place to monitor different aspects of quality.

3.1 Service quality

13. Evaluation of the level of quality of an ECEC setting (here referred to as service quality) is often conducted for accountability purposes. This can be for either external accountability purposes (also referred to as bureaucratic or hierarchical accountability) or internal accountability purposes (also referred to as professional accountability) (Adams and Kirst, 1999; Firestone, 2002; Levitt et al., 2008). The latter focuses on internal staff processes and practices and is therefore described under 3.2 "Performance of staff".

14. When quality is evaluated for external accountability, ECEC settings are understood as an instrument for implementation of family, labour market and education policies on national, regional and local levels. Information on the ECEC setting is provided to policy makers and the general public about compliance with regulations and the general quality of services provided. ECEC settings and staff are held accountable for the quality of care and early education they provide, most often measured by quality indicators.

15. It differs among countries who has the responsibility for conducting the monitoring practices for service quality. There may be a central authority in charge with a national agency conducting monitoring practices. But in many OECD countries, monitoring of service quality has been decentralised to local authorities, as is the case in the Scandinavian countries for example. ECEC staff and managers are often involved in the monitoring process or can be the evaluators of service quality through self-evaluations.

16. Additionally, parents or other stakeholders might evaluate the level of service quality. In Korea for example, Parent Monitoring Groups have been set up in 2005. These groups are managed and overseen by local governments since 2005. They visit child care centres, observe and monitor the on-going activities and provide child care policy recommendations to the local government. Children are not frequently involved in monitoring and evaluating quality, although several OECD countries have pointed out to a need for involvement of children in monitoring processes so their voices can be heard (OECD Network on Early Childhood Education and Care, 2012b).

3.1.1 Inspections

17. *Inspections*, conducted by external evaluators which have the primary role of inspecting services, are the most common tool for monitoring service quality in OECD countries, including Australia, the Flemish Community and Wallonia-Brussels Federation of Belgium, Canada, Czech Republic, Estonia, England, Germany, Hungary, Ireland, Italy, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Scotland, Slovak Republic, Spain, and the United States (at state level). Observation is the most common method of monitoring quality in inspections although the functions of inspections may differ among countries. A checklist or the use of rating scales (see Box 3.1) is a common tool used by evaluators during inspections. In some countries, such as Ireland and Norway, inspectoral observations are combined with interviews with managers and staff (OECD, 2012a; OECD Network on Early Childhood Education and Care, 2012a). Inspections may focus on procedural aspects, processes or a combination of both.

18. Procedural inspections focus primarily on monitoring compliance with regulations and standards. These are usually concerned with quality aspects such as staff-child ratios, indoor and outdoor space, health and safety regulations, the use of infrastructure, working conditions, staff's educational background, planning and workload of staff, and management of human and financial resources, rather than the actual work and tasks of staff.

19. When inspections focus on processes, they are intended to examine the relations between staff and children and staff and parents, collaboration between colleagues and other services or organisations (such as primary schools), the quality of teaching (instruction) and caring, and the implementation of a curricular programme and administrative tasks. Process inspections examine staff performance on a general level. Individual staff evaluations or assessments are generally not part of process inspections. The aspect of processes in inspection is receiving increasing international attention due to the growing interest in evaluating process quality (OECD, 2012a).

20. In general, inspections rarely focus on processes solely: process inspection is most often combined with an inspection of compliance with regulations (procedural). In addition, the results of formal child outcome or development assessments may be a part of the evaluation of service quality. However, since formal or standardised assessment of child outcomes or development is not common in ECEC, the adoption of an inspection approach including outcomes is rare in the ECEC sector. 3.1.2 Self-assessments

21. An ECEC setting can conduct an *internal evaluation* as well. This is done, for example, in Sweden where each ECEC setting prepares an annual evaluation report based on an internal assessment exercise (OECD, 2012a). An ECEC setting may use several tools to evaluate its own level of quality, including observations, interviews, rating scales (see Box 3.1), checklists (a list of items, tasks or steps to be taken in a specific order to be checked or consulted), and surveys (see 3.1.3).

22. The aim of internal evaluation is to establish whether settings have achieved national and local ECEC objectives as well as their own aims. It can also be used to reflect on what aspects need improvement and result in proposals for improving the level of quality. Self-assessments can be used by external evaluators as the basis of their inspection, but can also provide information to policy makers and the general public.

23. As an example, the Dutch Consortium for Child Care (NCKO) which studies the effects and levels of quality of child care in the Netherlands has developed a 'quality monitor', an instrument with which child care centres can assess their own quality. The results of the monitor provide an overview of what the weaker and stronger points are of a provision - with the goal to enhance the level of quality. The monitor assesses the interactions of all pedagogical staff, the quality of the care environment, as well as

structural aspects of the provision and makes use of checklists and rating scores. Special training modules have been developed to train staff and managers of child care centres in using the monitor. Besides, there is training available on analysing and improving staff-child interactions which are found to be key for early child development (NCKO website).

Box 3.1. The use of rating systems in external and internal evaluations.

Rating systems, such as the Infant/Toddler Environment Rating Scale (ITERS), the Early Childhood Environment Rating Scale (ECERS), the Classroom Assessment Scoring System (CLASS), or the Quality Improvement Rating System (QRIS), are used for external evaluations as well as for instruments of self-evaluation and quality improvement. These scales and guidelines are used most widely in the United States, but many other countries, including Australia, Canada, Chile, Germany and the United Kingdom, have also adopted these instruments.

Generally, rating scales focus on different aspects of quality, including standards and/or regulations as well as staff quality or performance. Rating scales are sometimes criticised as normalising, decontextualised instruments that may undermine more participatory and democratic reworking of quality by local communities (Dahlberg and Moss, 2005). When designing, adapting and implementing rating scales, it is important that staff are trained sufficiently to understand the theory and cultural assumptions behind these scales as to ensure they have sufficient knowledge on how to use the tool, what the rating scales measure, and know how to interpret the findings correctly. It is also highly important to adapt rating systems to a country's, state's or setting's own needs and circumstances.

Sources: OECD, 2006; 2012a.

3.1.3 Surveys

24. Surveys or questionnaires can be used by ECEC staff and management, as well as parents, to evaluate the quality of service provision. It establishes an opportunity for them to give their opinion on the level of quality of the ECEC setting and surveys can therefore also be an indication of parental or staff satisfaction. Surveys can be individual evaluation tools or part of a larger evaluation exercise, such as inspections or self-assessment exercises.

25. The survey or questionnaire may include a rating scale in which the evaluators rate certain aspects of the ECEC setting, such as "provision of information to parents" or "size of the room". Several countries, including Canada, Denmark, Finland, the Flemish Community of Belgium, Italy, Korea, Norway, Slovenia and Sweden give parents and/or staff a chance to evaluate service quality on a regular basis (OECD, 2012a).

3.2 Staff quality

26. Assessment and evaluation of staff performance is mainly conducted for internal accountability purposes and to improve staff practices and skills. It differs from external inspections of service quality (process inspections; see 3.1.1) since it relates to *individual* staff performance rather than general staff performance.

27. When assessing staff performance, staff is held accountable for how they implement their knowledge and skills, i.e., their interaction with colleagues and children. Staff can be held accountable to pre-defined criteria (at national, regional, local or setting levels) or professional standards, or codes of conduct (as set by a professional association and/or the government) (Rosenkvist, 2010). Monitoring staff quality most often involves observations or self-evaluations in combination with the use of rating scales, checklists or portfolios, and can be part of monitoring service quality as well (Isoré, 2009).

3.2.1 Observations

28. Observation practices taking place in an ECEC setting are one of the most common practices in monitoring individual staff quality in OECD countries (OECD, 2012a). This includes observing whether the staff member (teacher, instructor, caregiver or pedagogue) applies adequate practices for young children in the ECEC setting. This might be assessed as in accordance with existing standards or not (Isoré, 2009). Observations mostly take place in the form of inspections or peer reviews (OECD, 2012a).

29. Countries and practitioners have an increasing interest in monitoring and assessing the quality of staff-child interactions, since research indicates that process quality is highly important in providing high quality ECEC. High quality staff-child communication and interplay benefit children's learning and participation in activities (Sheridan, 2011). The Classroom Assessment Scoring System™ (CLASS) is an observation instrument that assesses the quality of teacher-child interactions in center-based preschool classrooms (see Box 3.2).

30. Another tool for monitoring staff quality through observation is the use of video feedback, whereby the staff member is being filmed and the video is being watched by the staff member and possibly other evaluators with the purpose to (self-) evaluate the performance and provide inputs for improvement.

Inspections

31. *Inspections* are widely used in OECD countries to observe and evaluate a setting's performance in general, taking staff's overall performance into account. This has been described in 3.1.1. National inspections rarely assess staff's individual performance, although few countries administer inspections when assessing staff quality. Newly qualified teachers in Ireland, for example, are registered with conditions, pending an evaluation by the Department of Education and Skills Inspectorate. During two unannounced visits, the Inspectorate evaluates the professional competence of probationary teachers by observing teaching and learning, examining preparation and progress records and evaluating samples of pupils' work. The outcome determines whether the teacher can move to full registration (OECD Network on Early Childhood Education and Care, 2012a).

Peer reviews

32. The use of *peer reviews* is a more common practice in assessing and evaluating individual staff performance. In peer reviews, other ECEC professionals (either within the same setting or from outside) and/or the manager observe a person and give constructive feedback to the person under evaluation. Reviewing staff performance allows staff to enhance their skills, adapt their practices to the needs of children, and informs settings of professional development needs of staff. *Portfolios* prepared by the staff member can also be used in peer reviews. Portfolios are a collection of staff member's current work, including day plans, videotapes of practices and exercises (video feedback), samples of children's work, etc. In addition to observations, these provide the evaluator with information on staff's capacities and skills.

Box 3.2. The Classroom Assessment Scoring System (CLASS)

The Classroom Assessment Scoring System™ (CLASS) is an observation instrument that assesses the quality of teacher-child interactions in center-based preschool classrooms. Evidence suggests that improving children's' academic achievement and social skill development requires attention to the nature and quality of teacher-student interactions; and enhancing the effectiveness of teacher-student interactions requires knowing how to assess them. CLASS does not measure other important components of high quality teaching and learning such as the curriculum used, the process of the ongoing assessment of child progress, or individualised teaching.

The CLASS observation tool can help teachers and schools improve the quality of their interactions with students. CLASS includes three domains or categories of teacher-child interactions that support children's learning and development: Emotional Support, Classroom Organisation, and Instructional Support. Within each domain are dimensions which capture more specific details about teachers' interactions with children. Following observations, observers rate each dimension on a 7-point scale, from low (score of 1-2) to high (score of 6-7).

In addition to defining and assessing these interactions, the tool also provides aligned professional development supports to give targeted feedback to districts, schools and teachers with the overarching goal of improving student outcomes. All reviewers and observers who work with the CLASS system are trained on implementation of the system and certified. In addition, they are provided with professional development that includes activities such as refresher trainings, webinars with CLASS experts, mentor coaching, and access to phone and email support. Empirical evidence from experimental evaluations demonstrates that teachers receiving the professional development support engage in more effective interactions with students, especially in classrooms that serve higher proportions of students in poverty.

Source: Center for Advanced Study of Teaching and Learning (2011).

3.2.2 Self-evaluations

33. Another common source of assessing staff performance is the use of self-evaluations where staff members evaluate their own performance (OECD, 2012a). These can take place as part of peer reviews or as an individual exercise. Self-evaluations are most often done through the use of self-reported questionnaires (surveys about a person's teaching and caring skills and practices filled in by staff themselves), self-reflection reports, journals or sheets, portfolios, and/or video feedback. The self-reflection process enables a professional to be aware of their own strengths and weaknesses, and to identify their needs for improvement, professional development or coaching (Isoré, 2009). Self-evaluations of educational practice were found to be a powerful tool for continuous support to the professionalism of early childhood education practitioners in Italy (Picchio et al., 2012).

Box 3.3. Self-evaluation tools for staff in ECEC

A major purpose of this type of evaluation is to raise the awareness of staff concerning different aspects of quality. Various participatory evaluation instruments of this nature have been developed in OECD countries for this purpose, such as the United Kingdom's EEL (Effective Early Learning) instrument. This instrument supports centres performing self-evaluations by encouraging discussion and reflection by staff on their programme, their attitudes and practice towards children and parents, as well as on the more technical aspects of administration, finance and planning. Through the EEL process, which normally takes place over several months, centres define their own programmes and activities, which are appropriate to their community circumstances.

The Flemish Community of Belgium has a process oriented self-evaluation instrument for care settings, named SICS (Self-evaluation Instrument for Care Settings), in place. The instrument was developed by the Research Centre for Experiential Education (Leuven University – Belgium), and takes the child and his or her experience of the care environment as the main criterion of quality. At the centre of this framework stand the process variables, well-being and involvement. Well-being and involvement are regarded as critical indicators for quality. The first is seen as a condition to secure mental health, the second as the condition for deep level learning and development. The procedure for self-evaluation starts with an assessment of the actual levels of well-being and involvement of the children by staff. For both indicators a five-point scale serves as a tool to code observations conducted in each of the groups that are part of the provision. To collect these scores, a preliminary scanning procedure is used by the head of the setting or by an internal co-ordinator. Individual children are observed one after the other, each for two minutes. In a second stage, the scores are analysed in order to identify the conditions that explain the recorded levels of well-being and involvement of the children. This analysis is guided by a framework in which five dimensions of the pedagogical approach are distinguished: the infrastructure and offer of activities, group climate, child initiative, adult style and organisation. Each of these dimensions is further defined by a series of items that can be rated on a three-point scale ("this is o.k." – "this could be improved" – "this needs to be addressed"). The analysis of the pedagogical approach is a shared activity in which the practitioners as a group work towards setting priorities for action. These serve as a guide to define possible interventions or practices they should adapt, to implement them and to reflect on their impact. This whole cycle of observation, analysis and action can be repeated several times in a year.

Sources: Bertram and Pascal, 1997; OECD, 2006; 2012a.

3.2.3 Tests

34. Test results of child outcomes, or children's developmental progress, are rarely used as a method in monitoring staff quality. Child development may reflect staff performance but is seldom used as a measurement of staff quality primarily because of the difficulties in using child assessments to draw conclusions about the quality of an ECEC setting the child attended, including the quality of the staff who worked with the child. Since early child development is not merely formed in ECEC but happens largely outside of ECEC provisions and is influenced by parents, family, friends, peers, and others, developmental tests for children are often regarded as an inappropriate or incomplete tool to assess staff quality.

35. Staff knowledge and pedagogical skills may be evaluated through testing professionals, although this method is rarely applied in OECD countries. Tests for staff are more frequently used for new pre-primary education teachers, as is currently the case in Chile, Denmark, France, Israel, Korea, Luxembourg, Norway, Spain, Turkey and the United States where a (competitive) examination is in place to enter the teaching profession. These tests often involve exams on subject-related matters as well as on pedagogical practices (OECD, 2012b). In Chile for example, pre-primary teachers take an initial professional test (Evaluación Inicia) before joining the labour force. This test comprehends disciplinary and pedagogical themes and is aligned with professional standards. Test results provide diagnostic information to universities and schools forming preschool teachers and provides data on what aspects of the studies need improvement (Chile Education Ministry website).

3.2.4 Interviews

36. Interviews are seldom used for assessing staff performance. They are more often conducted with the purpose of identifying needs for professional development, asking staff in which ways they think they need or would like to improve (Isoré, 2009). Performance management interviews or reviews can be used to identify staff's individual needs. However, the usefulness of such interviews highly depends on whether the staff member feels confident discussing their weaknesses and needs, and the possibilities to receive relevant and constructive feedback from this process.

3.2.5 Surveys

37. Surveys about staff practice can be completed by principals, managers or parents. They can assess staff quality through the continuous interaction they have with the staff member(s). Surveys on staff quality for parents are often conducted to analyse parental satisfaction. Surveys and questionnaires are tools for evaluating and assessing staff quality in several OECD countries, including e.g. the Slovak Republic and Sweden (OECD, 2012a).

38. While the responses to such surveys might provide interesting insights, it is important to take into account that when parents fill out these surveys, they are (in most cases) not ECEC experts and might value different qualities than the aspects which have been found to enhance child development. Besides, research indicates that parental satisfaction and pedagogical quality are not correlated although parental feedback might point to areas for improvement or lead to actions (Peterson et al., 2003).

3.3 Child development and outcomes

39. The assessment of child development and outcomes is a broad term that encompasses several methods and techniques for assessing how children have developed or what they have learned (and not learned). Child developmental assessments in ECEC are conducted to identify whether children have developmental delays or special needs that should be addressed (OECD Network on Early Childhood Education and Care, 2012b). Child assessments may use benchmarking tools, such as predefined outcomes for different ages, learning standards, developmental goals or curriculum objectives. But child assessments can also be applied without benchmarking children's development.

40. A distinction can be made between *formative* and *summative* assessment (OECD, 2005; Harlen, 2007). Summative assessment is used to measure what children have learned at the end of a certain time period and can be used for different purposes. Ministries can use summative assessments as a way to hold staff and settings accountable for providing quality ECEC. Or they can be used as a method to identify whether children are ready for entry into primary education. Formative assessments refer to frequent, interactive assessments of child development and progress with the purpose of understanding and identifying learning needs and adjust instruction and teaching methods accordingly (OECD, 2005). Formative assessments are more commonly used in ECEC than summative assessments (OECD, 2012a), summative assessment are more frequently applied in higher levels of education.

41. In ECEC, information and data on children's development or outcomes *over time* may be used to define the effectiveness of a setting and its staff. Such assessments are called "value-added assessments". It involves the comparison of child development or performance with its past assessments (Rosenkvist, 2010). An overview of the most frequently applied child assessment practices are described below.

3.3.1 Tests

42. A test is a formal assessment, often administered on paper or on the computer, intended to measure children’s knowledge, skills and/or aptitudes. Tests can be either standardised or not. A

standardised assessment is a test designed in such a way that the questions, conditions for administering, scoring procedures and interpretations are consistent. The goal of this consistency is to make the results as objective as possible so that they can be used to compare the assessed outcomes (OECD, 2012a; Zucker, 2004). Standardised assessments are usually administered to large groups of children and mainly for the purpose of measuring academic achievement and/or comparing members of a cohort (Rosenkvist, 2010). In most OECD countries, testing children in early education and care is not a common practice (OECD Network on Early Childhood Education and Care, 2012b).

3.3.2 Observations

43. Informal assessments of children's development and progress are more common in ECEC. Observations are the most widely used informal assessment method of child outcomes and their development. Observations mostly take place on a regular basis and involve the continuous and careful examination of a child's behaviour and skills by an evaluator. The evaluator is usually an ECEC professional working directly with the child (OECD Network on Early Childhood Education and Care, 2012b). Observations might be accompanied by rating scales and checklists (further explained below) or result in narrative assessments (see 3.3.3). Most OECD countries apply some form of observation practices in assessing children's developmental progress. These can be conducted in collaboration or in addition to, observations by parents or other early childhood professionals such as doctors or speech therapists. Information on child development by other actors than ECEC staff provide relevant and complementing data on children's skills, behaviour and growth.

44. In Estonia, for example, the national curriculum is used as a benchmarking tool when observing child development. Estonia's curriculum describes a child's development and required knowledge, skills and experience. Teachers carry out observations in the classroom as children engage in undirected play and in activities directed by the teacher. Through these observations, the teacher assesses whether the child is developing expected general skills in the subjects and areas articulated in the national curriculum. The Czech Republic has a similar approach to child assessment (OECD Network on Early Childhood Education and Care, 2012b).

Rating scales

45. Rating scales and/or checklists may be used in the assessment of children's development or outcomes and can be used by both professionals and parents, although parents are less frequently asked to fill out a rating scale exercise or a checklist. Both techniques are applied through observing the child. Within a rating system, different early child development skills or abilities are defined for children of different ages or in different development stages. For each skill or ability, the evaluator rates the child's competence in numbers or steps. This can range from, e.g., zero to three with three being an excellent rating (the child has the skill or ability to count to five for example) and zero being unable to do the task. Steps in between (such as one for a child that can count to two, and two for a child that can count to three or four) provide the evaluator with more refined and child-specific assessment possibilities than checklists.

Checklists

46. Checklists may also include a series of tasks, skills and abilities to assess children's development or knowledge, such as "child can count to five" or "child is able to play independently". However, unlike a rating system or scale, checklists only indicate whether a child can do a certain task or has a certain skill. The results of a checklist are often less specific and detailed than the outcomes of a rating system.

47. Lastly, checklists might also serve the purpose of guiding staff in observation methods. A checklist for staff can consist of a range of points the evaluator should take into account when observing a child's development ("observe a child's interaction skills with peers") so as to ensure the professional addresses all aspects relevant for assessing child development during its observation.

3.3.3 Narrative assessments: portfolios and storytelling

48. An increasingly popular method of assessing children's development is the use of narrative assessment practices. It can also be referred to as pedagogical documentation of development (Katz and Chard, 1996). Narrative assessments describe the development of a child through narratives/stories. Narrative assessment is a more inclusive approach to assessing child development, as it involves not only professionals but also the children, and can also include inputs or feedback from parents. It is a combination or package of what a child has done and learned, such as examples of drawings and exercises, and feedback from staff. Portfolios or storybooks of children's development are well known examples of narrative assessment practices. Narrative assessments are particularly popular in monitoring early development since these practices take place within the child's natural environment, taking into account its activities and relationships, and does not involved testing since children in ECEC are often too young for paper-and-pencil tests (OECD Network on Early Childhood Education and Care, 2012b).

Box 3.4. Narrative assessments to monitor child development

An increasing number of OECD countries and ECEC programmes are using narrative assessments for children's development. Narrative assessments form the basis of monitoring child development in the well-known Reggio Emilia programmes. Their pedagogical documentation is not only used to follow children's learning process but is also a tool for quality improvement. The narrative assessments or pedagogical documentations of Reggio Emilia include samples of children's work at several different stages of completion. This monitoring method does not only show the final product, but also tells staff and parents something about the way the child carried out the assignment or work, planned it, and completed it.

The Flemish Community of Belgium uses "story booklets" in which staff write and/or show what the child did and learned at the centre each day. In New Zealand, children's experiences are described in a Learning Story Framework by both staff and children. The Framework focuses on assessment in a narrative form as a story, a connection between the individual learner and the environment. It takes the view that children leave the early childhood setting for further education with some well-established learning narratives or working theories: packages of inclination, knowledge and skills to do with being a learner. This has provided a useful way for children and practitioners to reflect on ways to implement curriculum and assessment practices.

Finland and Norway have growth portfolios in place. This is a record of each child's life and growth at the centre, and is often the basis for assessment discussions with parents during the year. In addition to teacher comments and records, the child also contributes to the portfolio by entering photos, drawing and memories of significant moments. Through the portfolio, programme aims are explained to parents with the objective of mobilising parental follow-up and of achieving a shared understanding of education. Children can take their portfolio with them when they change to a new centre or school which can help smooth their transition from one setting to the other although this is not practiced everywhere.

Sources: Katz and Chard, 1996; OECD, 2006; 2012a; Taguma, Litjens and Makowiecki, 2012.

3.3.4 Screening

49. Some countries conduct screening tools to monitor children's development. Developmental screening allows a professional to keep track of a child's health or cognitive or social developmental process. This most often includes screening of language and/or reading development. Developmental screening of health development is done by health care professionals such as doctors, nurses or specialists in e.g. hearing development or autism.

50. Screening usually involves a short test to tell if a child is learning basic skills when he or she should, or if there are delays. It can include some questions the professional asks to a child or parent (depending on a child's age) or talk and play with the child during an examination to see how he or she plays, learns, speaks, behaves, and moves. The practice of screening is normally not used for measuring child outcomes but mostly conducted to detect any developmental arrears as early as possible and provide the child with additional support for development and learning if needed.

3.4 Curriculum implementation

51. Monitoring curriculum implementation is not only relevant to evaluate the usefulness of a curriculum, analyse the need for changes in the curriculum, or define staff development needs, but also in assessing process quality. Curriculum implementation is often part of a broader monitoring practice and incorporated in staff performance assessments or general service quality evaluations. While there is very little information available on monitoring practices regarding curriculum implementation specifically, observations seem to be the most common monitoring method (OECD, 2012a).

3.4.1 Inspections

52. Monitoring curriculum implementation is often part of the general inspection process, in which the use of a curriculum or learning standards is being evaluated as well as the adaptation of a national or state-level curriculum to local or setting circumstances and needs. Curricula are monitored as a part of the general inspection process in Australia, the French Community of Belgium, Korea, Japan, Norway, Poland, Slovak Republic, Spain, Turkey and the United Kingdom. In most cases, monitoring focuses on compliance with national legislation and the national curriculum or curriculum standards in place. For example, in Korea, the daily activity plans are analysed by the inspector as a method to check if these are in line with the national curriculum (Faubert, 2009; OECD, 2012a).

3.4.2. Peer reviewing

53. Curriculum implementation by ECEC staff can be monitored through the use of peer reviewing where a professional reviews the implementation of a curriculum or curriculum subjects of another professional. Peer reviewing may be part of a professional's performance interview or review. Only a few examples were found of the use of peer reviewing in monitoring curriculum implementation. As an example, beginning ECEC staff in New Zealand gain provisional teacher registration. They then embark on a two-year induction process during which a mentor teacher oversees and reviews their programme, including the implementation of the New Zealand curriculum and its content. Only when new teachers meet the Satisfactory Teacher Dimensions according to the mentor will the new teacher be recommended to the professional leader of the early childhood service. The professional leader then recommends the teacher to the New Zealand Teacher Council for full registration (Taguma, Litjens and Makowiecki, 2012).

3.4.3 Self-assessments

54. Staff can also be their own evaluators of curriculum implementation. A self-evaluation of the practical use of a curriculum framework, the knowledge on curriculum content, and a reflection on a person's strengths and weaknesses on the curriculum can provide useful insights into how staff can enhance practices and skills and better implement the curriculum. Tools for self-assessment may include checklists, rating scales and surveys. Self-assessments of curriculum implementation are most often applied as part of staff performance assessments.

55. In Korea, in 2012, surveys were used at a more large-scale level to evaluate the implementation of the new Nuri curriculum for five year olds. 7 822 teachers and directors of ECEC provisions participated in the survey. Results showed that most teachers needed more support in the implementation

of the curriculum, besides the existing guide and materials available, and pointed to the specific needs and areas for support.

3.4.3 Tests

56. Tests to assess staff's knowledge of the curriculum or curriculum subjects and its implementation are rarely used in ECEC. In some countries, including Denmark and Luxembourg, new pre-primary teaching staff need to pass a competitive examination to enter the teaching profession, which usually tests the knowledge on curriculum subjects as well (OECD, 2012b). In Chile, new staff take a voluntary test which examines their knowledge, as well as strengths and weaknesses on curriculum subjects including natural and social sciences, mathematics, language, visual arts and music (Chile Ministry of Education website).While tests on a regular basis might identify strengths and weaknesses of staff regarding the curriculum, as well as their professional development needs, peer reviews and performance reviews are more commonly used for this purpose.

4. USERS AND USES OF MONITORING PRACTICES AND RESULTS

4.1 Uses of monitoring practices and results

57. While the uses, or purposes, of monitoring practices differ depending on the subject that is being monitored, monitoring is most commonly applied for accountability and/or improvement reasons. Other purposes include identifying learning needs, monitoring for receiving rewards or sanctioning, and informing policy-makers and/or the public. Monitoring may attach high or low stakes, indicating that the results of the monitoring practice might have (larger or smaller) implications for provisions, staff, and children. The idea of attaching high stakes to monitoring is that it can provide incentives for services and their staff to work better and more effectively.

4.1.1 For accountability purposes

58. Monitoring practices are frequently conducted for accountability reasons. This refers to an account-giving (i.e. information-sharing) action from ECEC provisions to policy makers, stakeholders, or the general public. It involves the sharing of information of a provision's (past or future) actions and decisions, and justification for these. These actions can refer to financial spending patterns, compliance with rules and regulations, the use of human resources and other inputs, outputs of a centre's or system's practices, and other actions the government feels ECEC centres should be held accountable for. Increasing public spending levels in ECEC, as well as a trend in decentralising authority regarding ECEC to local governments and provisions, contribute to a growing demand for accountability and efficiency in ECEC systems (Morris, 2011).

59. Both high and low stakes can be attached to monitoring for accountability purposes. High stakes may involve receiving rewards or putting sanctions in place (see 4.1.4). Publication of monitoring results may add to the stakes, as was found to be the case in schools (OECD, 2012b) since directors of schools and teachers will work to avoid the stigma of low performance or quality (McDonnell and Choisser, 1997).

4.1.2 For improvement purposes

60. There can also be a great emphasis on use of monitoring results for improvement. The emphasis is on using the information gathered in monitoring practices as a means to enhance performance of provisions and centres, professionals, or improve children's development. The results can provide feedback on what works and help identify areas of improvement. Monitoring for improvement is usually associated with low stakes since results are used to guide practices for improvement.

Improvement of quality

61. Centres can use information from monitoring and evaluations to identify weak areas, better understand its strengths, and strengthen aspects which need improvement. Actions can be taken based on the monitoring results, such as providing appropriate professional development for staff to enhance staff quality, securing extra financial or human resources, ensuring the appropriate use of tools and materials or provision of support materials, improving the content of curriculum to better meet needs of children,

parents and society, and other practices to improve the overall quality of ECEC provision. In the Slovak Republic for example, results are used to improve the curriculum (National Educational Programme), provide materials, and make amendments to legal regulations (OECD Network on Early Childhood Education and Care, 2012b).

Enhancing staff performance

62. Using monitoring results to improve practices and instructions is a key purpose of monitoring. Monitoring results help staff to continually evaluate the effectiveness of their instructions and make better informed instructional decisions (Safer and Fleishman, 2005). Monitoring practices provide important feedback to professionals on their abilities and skills, as well as knowledge on specific curriculum subject areas. Based on the results, staff can improve their instruction techniques, analyse their needs for development and strengthen their skills and knowledge through professional training (see 4.1.3), identify needs for support (either in additional human resources or extra materials), and improve implementation of the curriculum.

Improving child outcomes

63. An important part of monitoring children is screening for possible developmental delays in order to initiate intervention services as early as possible for remediation. Monitoring can help identify which children are at risk of having developmental arrears. The results can be used to inform curriculum and remedial interventions. In addition, practices can be adjusted to better meet these children's needs, or support materials can be provided to the staff and children to help improve children's development and diminish any learning arrears (Allen, 2007).

64. Monitoring is not only beneficial for children 'at risk' but for all children in general. Continuous monitoring of every child can provide staff and parents with useful knowledge on what and how a child learns and grows, and creates a better understanding of every child's unique development (Horton and Bowman, 2002; Morris, 2011). Such increased understanding can be used to further stimulate children's development and can be helpful in identifying useful strategies in enhancing child outcomes.

4.1.3 To identify learning needs

65. Aside from accountability and improvement purposes, monitoring can be conducted in order to identify learning needs of staff and children. This also referred to as 'monitoring for learning' or 'assessment for learning' since it aims at identifying strengths and weaknesses of ECEC professionals and of children, and providing adequate development opportunities for the areas in need of improvement. The role of monitoring for identifying learning needs is to underline professional ways to improve and adapt current practices of programmes and professionals to either enhance their own performance or support children in enhancing their development. Most often, low stakes are attached to the results of monitoring to identify learning needs. Rather, centres and professionals and encouraged to use to results to guide their practices.

66. The results of monitoring practices allow ECEC centres to adapt their professional development programmes to the needs of ECEC professionals in accordance to their own objectives. Provisions can learn from the strengths of effective professionals and implement development programmes that respond to staff members' weaknesses. Linking monitoring practices to professional development opportunities for staff ensure that staff quality is prioritised and that professionals receive the appropriate training that fits their needs. For children with certain weaknesses, curricula and programmes can be adapted, support tools may be provided, or specific programmes to help these children can be implemented (OECD Network on Early Childhood Education and Care, 2012b).

67. Institutions in charge of pre-service (initial) education of ECEC professionals can also benefit from the results of monitoring and adapt their programmes to professional and practical needs of future ECEC staff.

68. As an example, Caregiver Interaction Profile (CIP) scales in the Netherlands can serve as a starting point for education and training to improve the quality of caregiver–child interactions. The CIP rates six key skills of caregivers for interacting with 0- to 4-year-old children in child care centres: sensitive responsiveness, respect for autonomy, structuring and limit setting, verbal communication, developmental stimulation, and fostering positive peer interactions. Each interactive skill is rated on a 7-point scale based on observation of video-recorded caregiver–child interactions. Together, the 6 scale scores constitute an Interaction Profile for individual caregivers that can indicate for what aspects caregivers need training (Helmerhorst et al., 2014).

4.1.4 For reward or sanction purposes

69. Monitoring can be linked to rewards or sanctions. This is most often the case when monitoring is conducted for accountability purposes. Actions or decisions regarding rewards or sanctions are often made based on whether ECEC centres meet certain standards or targets. Rewards for centres can include additional funding or other financial contributions, whereas sanctions can include intervention, financial resource limitations, or in rare cases closure of the provision.

70. Monitoring can also be used as a basis for recognition of a professional's performance. Monitoring staff quality for example, provides opportunities to recognise and reward staff competences, which is found to be essential in retaining effective staff members as well as to make the sector more attractive (OECD, 2005). The results of monitoring can be used in making consequential decisions concerning ECEC professionals such as promotion opportunities or salary increases but can also be used in responding to ineffective professionals by, for example, offering them training or extra schooling(Isoré, 2009; Morris, 2011).

4.1.5 To inform policy-making

71. Monitoring practices are often conducted to inform policy-makers on the effectiveness of their policies and the results are used in developing appropriate policies. As an example, in the Netherlands, data on child outcomes and development are collected of a sample of all young children attending ECEC every two years. In this way, it is possible to track trends in child development, ECEC participation, and school career. This data can be used by policy makers to analyse whether their policies target the right group of children, to inform them on the intensity of participation, and the outputs of their policies. Based on the results, policies can be revised or programmes adapted (Kohnstamm Instituut, 2012). Evidence shows that ongoing dialogue between researchers, policy-makers and stakeholders raises the impact of research findings and successful implementation of policy recommendations (Munté, Serradell and Sordé, 2011; Puigvert, Christou and Holdford, 2012).

4.1.6 To inform the public

72. Another purpose of monitoring is to inform the general public about the performance of centres or the ECEC sector more broadly. This refers to the practice of making monitoring results publically available for use by parents, stakeholders, and the media. This is most often done through the web, or by publishing monitoring reports. The provision of information to the public is an important mechanism for parents and stakeholders to be able to hold the ECEC system accountable and to use this information to point out needs for improvement or changes. Not only does this serve the purpose of providing information

on ECEC system performance to the general public, but the results may also be used by stakeholders to take action (OECD, 2012a). For example, parents may use the monitoring results to make decisions on their child's ECEC participation or the media can use the results to put ECEC on the public agenda.

4.2 Users of monitoring results

73. There are also a number of users that use monitoring results. In this paper, the users are categorised into seven broad categories. An overview of these is presented in Table 4.1.

Table 4.1. Users of monitoring results

Category	Users
Central/ National authorities	Ministries, Directorates, Government-related agencies (*e.g.* Inspectorates)
Regional/ Local authorities	Regional and Local Governments, Municipalities
ECEC settings	Managers, principals, leaders, owners and administrators
ECEC staff	Professionals working in ECEC settings directly with children
Parents	Parents with children in ECEC settings
Researchers	Research institutions, universities and colleges
Media	Internet websites, news papers, TV channels

4.2.1 Central authorities

74. Government administrations organise data collection and monitoring in the ECEC field and cover important areas of ECEC policy, such as demand, supply and utilisation of ECEC places; the volume and allocation of public financing; the status of children (demographic, health, socio-economic, etc.) within and outside ECEC services; and the recruitment and training levels of staff. By linking monitoring results on children, programme characteristics and workforce across multiple programmes and governance structures, policy makers can acquire a holistic understanding of the system (OECD 2012a; OECD, 2006).

75. Monitoring helps policy-makers in establishing facts and evidence about the ECEC sector, for example, whether children have equitable access to high-quality ECEC, and it can contribute to evidence-based policy making. Besides, it provides evidence of particular strengths and weaknesses of the ECEC system, and whether government objectives and goals are being met. Central authorities may also perceive monitoring results as a strategy for making the ECEC system more accountable. This can be done through, for example, making monitoring results publically available for stakeholders, parents and citizens and inform them on how the ECEC system is performing (OECD, 2012a).

Examples of uses

76. In Canada, the Child Care Human Resources Sector Council (CCHRSC) collected information and data on wages, working conditions, and human resource issues in regulated child care and early education centres through a survey. The findings from this survey, which were published in a report in 2013, provided policy makers and employers with essential data that will be used to address ECEC

workforce policy challenges such as recruitment and retention, training and professional development, opportunities for career advancement and job satisfaction (CCHRSC, 2013).

77. Norway also uses collected data and information on the workforce for evidence-based policy-making. Statistics Norway regularly collects data on ECEC sector staff, working conditions and workforce supply. Additionally, standardised annual reports from all kindergartens indicate their number of staff and their qualifications. Based on the collected information, policy areas in need of improvement or challenges in the ECEC sector are identified. The collection of this data indicated a need for more qualified staff and, more specifically, which regions had difficulties with workforce supply. As a result, a general action plan for the recruitment of preschool teachers in targeted regions was launched by the Ministry of Education and Research (OECD, 2012a).

78. In the United Kingdom, the Office for Standards in Education, Children's Services and Skills (Ofsted) produces an annual report presenting evidence from inspection and regulatory visits across the full range of Ofsted's statutory remit which includes early years and child care, provision for education and skills and schools, colleges and adult learning, children's social care and local authority services for children. Information is collected about provision of children's social care, education and skills, resulting in a national perspective on provision. The report sets out an overview of the quality of provision and of the progress and experience of the children concerned; and when appropriate, it draws comparisons with inspection findings from previous years. The purpose of the report is to establish a candid picture of the quality of services so that informed choices can be made about them, both by those who use them and by those who fund them. In addition, such reports are publicised to provide incentives for services to raise standards (Ofsted, 2011).

79. In the United States, the result of a study on the use of monitoring results found that the most common application of ECEC data collected across the country is to guide teacher professional development (85%), followed by provide technical assistance (70%), corrective actions or sanctions (67%) and changes to ECEC policy (59%). Data and monitoring are crucial to evidence-based policy making, which can contribute to greater policy effectiveness and efficiency (Barnett et al., 2012).

4.2.2 Local authorities

80. Local authorities exercise responsibilities in the field of ECEC in many OECD countries, particularly in countries with decentralised ECEC systems at the municipal/centre level such as e.g. the Nordic countries. Local policymakers use monitoring data and information to make informed decisions, as they seek solutions and supports to promote high-quality services within their jurisdictions. Assessing children's school readiness, for example, can help policy makers document population trends, track children's progress over time and determine if public ECEC expenditures are making a difference (CCSSO, 2011).

Examples of uses

81. In New Jersey (United States), classroom quality is measured through structured observations using three instruments: the first instrument measures general classroom quality with an emphasis on classroom environments for health, safety and provision of education materials; the second assesses both general and specific materials and teacher activities and interactions that have been found to lead to increased oral language and literacy skills; and the third provides information on the materials and teaching interactions across all the types of mathematical thinking and skills. Together, these criterion-referenced instruments provide information that is easily understood and used for programme improvement at the classroom, district and state levels (OECD, 2012a).

82. In Denmark, all day care institutions are required by law to evaluate their pedagogical curricula annually, documenting the extent to which the applied methods and activities have achieved the learning outcomes set out in these curricula. The legislation then requires local authorities to consider the monitoring results and assess whether action is required in order to ensure optimal service quality.

83. In Norway, the Country Governor inspects municipalities to ensure that ECEC is in accordance with the Kindergarten Act. Municipalities are responsible for the development and supervision of both private and municipal institutions and for ensuring that institutions are run according to goals set by the national government. Results of inspections can be used to draw attention to areas in need of improvement and may lead to actual improvements in provisions (OECD, 2012a).

84. An assessment of the quality of daycare services in Quebec (Canada), based on the Quebec Longitudinal Study of Child Development (QLSCD), raised many issues related to child care practices in the various types of settings in Quebec. The assessment findings were presented to the Social Affairs Committee of the Quebec National Assembly as well as to the "Association Québécoise des centres de la petite enfance", and led to several regional initiatives to improve the quality of the services (Japel, Tremblay and Côté, 2005).

85. Local governments in Australia and in Vancouver (Canada) use the results of child development assessments, such as the Early development Index (EDI), to support applications for funding. Mapping child development outcomes by region indicates whether regions are good performers, or whether there are needs for improvement. Such information is used by local governments when applying for funding (Early Years Institute, 2012).

4.2.3 ECEC settings

86. Leaders of ECEC settings, such as managers, principals and administrators, can use monitoring results from various levels (national, local, centre, staff) for numerous purposes: to make improvements to their ECEC setting, to account for child development and outcomes, to hold staff accountable for their performance, etc. Studies show that monitoring quality can lead to increased programme quality, as reflected by the adoption of higher standards and improved ratings of the ECEC setting environment (Office of Child Development and Early Learning, 2010; Rand, 2008).

Examples of uses

87. In the Slovak Republic, monitoring results of inspections are used by ECEC settings to improve the quality provision. The State School Inspectorate carries out inspections for kindergartens, focussing on the quality of education and related child outcomes; quality of ECEC provision and its management; co-operation with parents and other sectors (e.g., primary school, special pedagogues, psychologists); working conditions of ECEC staff; and supplementary activities of the kindergarten. Based on inspection results, kindergarten services share updated goals and information with relevant stakeholders; establish appropriate decision-making standards for kindergarten directors; collaborate with advisory bodies to address professional issues; and co-operate with various educational institutions to enhance child development (OECD, 2012a).

88. ECEC services in the United States using Head Start assess child development and use the results and feedback from these assessments to improve their programmes. Head Start programmes must assess children at least three times per year in alignment with the Head Start Child Development and Early Learning Framework, which specifies progress children should make on ten domains ranging from literacy development to physical development. Each programme determines how to assess children, and then uses the feedback from the assessment for programme improvement and to set early learning goals for each

child. Methods employed by individual programmes may include standardised tests, teacher observations and developmental checklists aligned with curriculum. The national government inspects the programmes once every three years to determine whether the programme is conducting an assessment of child outcomes three times per year and is using the information appropriately to set early learning goals. The results of the inspection are linked to a programme's funding: failure to analyse child assessment data three times per year and establish school readiness goals results in a programme's funding not to be automatically renewed, and they will need to compete with other ECEC providers in the area for a new grant. Federal reviewers do not analyse the results of the child assessment and the results of the assessments are not reported to the national government (OECD Network on Early Childhood Education and Care, 2012b).

89. Also Australia uses monitoring and evaluation results to improve the level of quality. In 2012, Australia introduced a National Quality Standard (NQS), which is regulated by State/Territory Regulatory Authorities. Within this regulatory framework, it is recommended for professionals to regularly assess their practice, in order to recognise their strengths and identify areas that can be improved. Therefore, services are expected to have in place a comprehensive process of reflection, self assessment and evaluation, and it is recommended that it is in line with NQS as to ensure that all standards and elements outlined in the NQS are addressed. This enables services to gain an informed picture of current practice and the quality of education and care experienced by children and families. Based on the assessment of quality results, the service can determine where quality improvements need to be made and plan effectively to implement them (OECD Network on Early Childhood Education and Care, 2011).

90. New Zealand's Education Review Office publishes national evaluation reports, which are publically available online; and some reports are published in booklet form and sent to all early childhood services. Feedback indicates that the findings of these reports are useful and used to inform practice and as a basis for self review in early childhood services (OECD Network on Early Childhood Education and Care, 2012b).

4.2.4 ECEC staff

91. By using findings from a variety of monitoring methods (i.e. evaluation of staff performance, staff observations of children, child assessments, parental surveys), staff are equipped with tools to provide more effective supports and learning opportunities for children. ECEC professionals working directly with children generally want to know the status of children's early learning and development in order to plan high-quality programme services and practices and determine whether these are effective. However, it is important that staff receive training to understand and interpret findings of monitoring practices, and it is essential that staff are aware of precautions specific to the assessment of young children (Barnett et al., 2004).

Examples of uses

92. In the Wallonia-Brussels Federation of Belgium, teachers evaluate children on a day-to-day basis to determine whether children have delays or special needs that should be addressed. Preschool is not compulsory, so children's performance is generally not evaluated from a summative perspective. Nevertheless, at the end of preschool, before the child starts compulsory school at age six, several cognitive and compartmental competences are evaluated by the teacher and by a psycho-medical centre. The results are used to detect any developmental arrears and in some cases, they can advise to direct the child in another form of education if disabilities are present (OECD Network on Early Childhood Education and Care, 2012b).

93. In Spain, an evaluation of the quality of the country's "Trilingual Early Stimulation Programme" underlined the strengths and weaknesses of this early education programme for their staff. It was found that

there is a need to assess the programmes on a more regular basis, and that a protocol is needed for this. Consequently, a protocol to assess the programmes has been developed and applied. This protocol, as well as regular assessment of their programmes and identification of needs for improvement, has allowed educators to refine and improve their programmes over time (Rivas and Sobrino, 2011).

94. In New Jersey (United States), the introduction of a quality rating score (among other things) in the Abbott Preschool Programme, allowed practitioners and management to improve their practices. A rating score exercise may point to any strengths as well as weaknesses in practices and programmes, which can be used by staff to improve their organisation and instruction (Frede et al., 2007; Frede et al., 2011).

4.2.5 Parents

95. Once parents decide to use ECEC services, they must decide which type of service to use. The quality of different services is often an important factor taken into consideration by parents when choosing a service. Practical considerations, such as opening hours, available spaces, flexible use and service location are also highly important, particularly for those who are working. In addition, finding a service that matches their child's age, needs and abilities is relevant for parents. For some parents, the philosophy of the provider is important, but other parents do not consider this an important consideration (Barnett et al., 2004).

96. Making monitoring results publically available to the public, including parents, is a manner to actively involve parents in ensuring quality at ECEC centres. Parents can, with public information on quality performance, challenge provision's weaknesses. This can result in adjustments in ECEC provisions and lead to improvements in quality of ECEC services. Furthermore, it may enable parents to choose a desirable setting for their children. Once the child is attending an ECEC service, information on developmental progress can and should be shared with parents to help them understand what and how their children are learning and how they can extend this learning at home (Barnett et al., 2004; Rosenkvist, 2010).

Examples of uses

97. Estonia's Education Information System provides the public with an opportunity to review performance indicators of preschool institutions. On their website, information on indicators regarding support for children with special needs, the number of teachers with required qualification, the average amount of in-service training for teachers, the teachers' ages, turnover rates, teacher gender and the staff-child ratio can be found for preschool programmes. Such data informs parents and the general public about the level of quality a programme offers, and serves as a tool for preschool programmes to track progress in quality over time (OECD Network on Early Childhood Education and Care, 2012b).

98. In the United States, about half of all states have a Quality Rating and Improvement System (QRIS) in place. These are voluntary systems that rate child care providers on a scale of 1-3 or 1-5 using stars or another consumer friendly method. The state establishes quality standards for each level in the QRIS and the rating is shared with the public, so that parents can identify the level of quality that they are buying and make informed decisions on what centre their child should attend (OECD Network on Early Childhood Education and Care, 2012b).

99. In British Columbia (Canada), the Human Early Learning Partnership works in collaboration with the provincial government and local communities to map Early Development Index (EDI) results, socioeconomic data and demographic characteristics for local regions across the province. EDI data are used extensively to inform communities about how their children are doing and what can be done to improve children's early learning environments (Forum for Early Childhood Development Monitoring).

4.2.6 Researchers

100. Having information on ECEC programme components and child outcomes allows researchers to draw clearer conclusions regarding who benefits and under what conditions. In many OECD countries, data collection and monitoring is outsourced to independent researchers, such as universities, government-sponsored research centres or private research firms. In turn, these researchers may be responsible for sharing findings with policy-makers, other researchers, policy-makers, and the general public. Research results are mostly shared with the goal of building the knowledge base, and use the findings to improve quality or child outcomes (Zaslow et al., 2009).

Examples of uses

101. In Germany, the NUBBEK Consortium, an affiliation of several institutes and individual researchers, has taken on the following tasks within the framework of a study: to make reliable, basic, empirical and practical knowledge available; to scientifically examine the existing and emerging conditions and problems; and to use this empirical knowledge to expand the basis for the design of good early childhood education, care and upbringing for children and to increase support for families with child-raising responsibilities (OECD, 2012a).

102. In the Netherlands, the Netherlands Consortium Kinderopvang (NCKO or Dutch Consortium on Child Care) conducts research on the pedagogical quality of Dutch child care provisions for children from birth to four years. The studies are commissioned by the Ministry of Social Affairs and Employment, who has the responsibility for child care. The work started in 2002 and is repeated every couple of years to provide an overview and trend analysis of the provision of quality in Dutch child care centres to the Ministry and the public. To ensure the results can be used in practice for improvement of quality, a sound board consisting of relevant stakeholders in the child care centre, has been established which provides feedback to the Consortium and the Ministry regarding action steps. The studies not only provide information on the level of quality of child care but also data on quality aspects (NCKO website, 2013). In the United States, the National Institute of Early Education Research works with state and national policy makers and other organisations to develop research and communication strategies to fill gaps in knowledge and to effectively apply scientific knowledge to early education policy (OECD, 2012a).

103. In Scandinavia, the Danish Clearinghouse of Educational Research (University of Aarhus) conducts an annual mapping and assessment of Scandinavian research in early childhood institutions since 2006. The purpose is to present and evaluate reviewed research publications on institutions for children aged 0–6 years that have been conducted in Sweden, Norway and Denmark and (partially) Finland. An open online accessible database has been set up which makes this research available for the international ECEC community: http://www.nb-ecec.no/skandinavisk-forskning-pa-dagtilbudsomradet-en?set_language=en

4.2.7 Media

104. ECEC research findings can be shared extensively and at a rapidly increasing pace via news media. In many OECD countries, newspapers, blogs, television, radio, etc. provide ECEC-related coverage on a variety of topics including the impact and benefits of ECEC, well-performing and under-performing ECEC programmes, investing in ECEC, ECEC policy developments and access to ECEC services. News sources range from national- and local-level monitoring results to academic research findings; and media coverage can be instrumental in raising awareness of ECEC issues as well as evoking a response, such as government support or parental engagement.

5. EVIDENCE REGARDING MONITORING PRACTICES

105. This section presents the evidence on the impact of different monitoring practices on quality and child development or outcomes. While there is little research and studies available on the effects of monitoring in ECEC, we aimed at providing an overview of the findings for each of the four monitoring aspects: service quality; staff quality; child development/outcomes; and curriculum implementation. The research depicted includes qualitative and quantitative evidence. Further, in-depth information and knowledge on country practices and content of monitoring practices and tools will be collected through a survey as part of the new project. The results of this survey will be presented in a final report on monitoring quality in ECEC, and is expected to provide additional analyses on challenges and strategies in monitoring early education and care.

5.1 Monitoring service quality

106. Research on the effects of monitoring service quality on the improvement of the level of quality is gradually emerging, but the existence of findings does not necessarily imply the ability of researchers to identify the impact of monitoring *per se*. The literature supports the idea that monitoring and evaluation are critical for high-quality ECEC services. Cubey and Dalli (1996) indicate that without evaluation, there can be no guarantee that services meet their aims and goals. Many countries monitor the service quality of ECEC settings using external evaluation tools (e.g., inspections, observations, surveys, questionnaires, rating scales, checklists) more often than internal evaluation tools (e.g., rating scales, checklists, self-assessments, evaluation reports, portfolios) (OECD, 2012a). Studies have been conducted, mainly in the United Kingdom and the United States, on the impact certain monitoring tools have had on the quality of ECEC services; but it is often challenging to separate and identify the impact of a single tool or method.

107. A study conducted by the RAND Corporation (Zellman et al., 2008) assesses the validity of a quality rating and improvement system (QRIS) as a tool for improving child care quality. The QRIS assessed was implemented in 1999, was one of the first of its kind and was created by Qualistar Early Learning, a Colorado (United States) based non-profit organisation. The rating system includes components generally agreed to contribute to high-quality care: classroom environment, child-staff ratios, staff and director training and education, parent involvement and accreditation. The study found that among providers using the QRIS, service quality did improve over time. However, it is not possible to unequivocally attribute improvements to the QRIS: improvements could have been a reaction to being monitored, for example. Difficulties in measuring the effect of this particular intervention include participant self-selection, the lack of a comparison group and limited data on the implementation of the intervention. The study notes the importance of validating a tool such as the QRIS, particularly as it is sometimes linked to rewarding higher quality services with, for example, higher per-child subsidies. Tout et al. (2009) find that while QRISs potentially serve as a hub for quality improvement, attaining this goal requires extensive co-ordination across agencies, services and data systems.

108. A study of a quality rating system implemented in Oklahoma (United States) found that the intervention improved the quality of individual child care centres as well as the overall quality of child care services throughout the state (Norris, Dunn and Eckert, 2003). However, when applied to family child care settings, the rating system was not found to elevate the level of quality of this type of service. The intervention simply served as evidence that family child care settings vary in terms of ratings, thus

validating the notion that the rating criteria represent different levels of quality (Norris and Dunn, 2004). The rating system articulates quality criteria beyond licensing requirements that providers may choose to meet in order to receive higher rates of reimbursement for the provision of services. Criteria focus on staff education and training, compensation, learning environments, parent involvement and programme evaluation. The child care centre study reports that, as a result of the rating system and increased financial support for highly rated services, more programmes are enrolling children subsidised by the state's Department of Human Services, and global quality ratings have risen. The family child care study indicates that aspects of family child care practice are missing from the rating system, and a challenge is to find beneficial criteria that can be operationalised and implemented by both policy makers and providers.

109. In the United Kingdom, the Office for Standards in Education, Children's Services and Skills (Ofsted) issued an evaluation report (Matthews and Sammons, 2004) on the impact of inspections carried out over the course of 10 years since Ofsted's inception in 1992 with the aim to improve education and care services. The report notes that Ofsted has little direct control over this aim, except regarding statutory provisions for identifying and monitoring schools and regulatory control of child care. Findings indicate that well-managed providers and those that cause concern are the most likely to benefit from inspections. Others show varying levels of improvement, but in the absence of external follow-up or something such as funding or prestige, tend to make incomplete use of their inspection results.

110. Another Ofsted report (2011) presenting evidence from inspection and regulatory visits undertaken from 2010-11 provides a more detailed look at the growing early years sector, as the Early Years Foundation Stage (EYFS) was introduced in 2008. The EYFS is the statutory framework that sets standards that all early years providers must meet to ensure that children learn and develop well and are kept healthy and safe. Upon inspecting against the requirements of the EYFS, the quality of child care has seemingly improved. For example, in 2008, just under two thirds of provision was found to be good or better, and 5% was inadequate; and in 2011, almost three quarters was good or better, and 3% inadequate. Providers that have left the sector are 10 times as likely to have been found inadequate as those who remained in the sector. Hence, the report suggests that inspection against the EYFS requirements has contributed to an overall increase in service quality.

111. Lee and Walsh (2004) stress that, in order to be valid and meaningful, evaluation of ECEC programmes should be based on in-depth understanding of dynamic programme processes and a range of stakeholders' perspectives on quality. When monitoring outcomes, or using outcomes in monitoring processes, it is important to monitor inputs as well since it are the inputs which directly or indirectly result in the outcomes. They warn that placing a high value on outcome-oriented evaluations of programme effectiveness and on standards-based evaluations for quality control may lead to under-diversified monitoring approaches and prevent the ability to look at service quality from various angles.

112. Additionally, Matthews and Sammons (2004) point out that while some researchers and policy makers assume it should be possible to demonstrate a causal link between a monitoring tool, such as inspection, and service quality improvement, such expectations may be too simplistic. Despite evidence of quality improvement, it is extremely difficult to attribute causality in the study of social and educational processes. Concurring interventions and initiatives with varying functions and levels of effectiveness make it near impossible to isolate the effect of one particular monitoring tool: improvements are most likely the result of a combined impact of numerous policy developments to monitor and improve service quality. However, when evidence on the impact of a particular intervention from various sources points in the same general direction, it is reasonable to infer a general association between the monitoring tool and its ability to stimulate quality improvement. This does not indicate that monitoring cannot have any benefits since monitoring makes it possible to analyse strengths and weaknesses of an ECEC service and can, through this, contribute to improvement.

113. Lastly, several studies point to the importance of family engagement in ECEC and education and it is argued that this should be a part of monitoring service quality systems (Edwards et al., 2008; Hidalgo, Epstein, and Siu, 2002; Weiss et al., 2008). Research has indicated that family involvement in early education has a great influence on children's learning and development. Hidalgo, Epstein, and Siu (2002) found evidence that family involvement activities are highly important in helping children succeed in education. This was found to be true for children with different backgrounds, regardless of parents' formal education, income level, family culture or language spoken at home. A case study in Spain for example, found that pre-primary schools that achieve better learning and developmental outcomes for all children are those with not only high quality staff-child interactions but also strong coordination between staff and the child's home environment and community services (Gatt, Ojala and Soler, 2011).

5.2 Monitoring staff quality

114. The literature widely acknowledges that quality of staff and their pedagogical activities, interactions, and knowledge have a large impact on children's well-being and development (Fukkink, 2011; OECD, 2012a). It also points out that effective monitoring of staff is central to the continuous improvement of ECEC services. Countries employ a mix of external and internal evaluations to monitor staff performance in kindergartens/preschools, child care centres and family day care. Commonly used tools for external evaluations include inspections, survey and observations; and for internal evaluations, self-assessments and rating scales (OECD, 2012a). The staff traits that research identifies as important in facilitating high-quality services and outcomes include such aspects as a good understanding of child development and learning, the ability to develop children's perspectives, age-appropriate communication and practices, leadership and problem solving skills, and development of targeted lesson plans (OECD, 2012a). However, there is little consensus among researchers regarding the effects or impacts of monitoring staff quality.

115. While it is difficult to measure the impacts of monitoring staff quality on, for example, the improvement of the level of service quality, staff performance and implementation of curriculum, and child outcomes/development, research has emerged (primarily from Anglo-Saxon countries) that begins to scratch the surface of this complex topic. This research tends to examine the impacts of specific monitoring tools (see Box3.2) or the effectiveness of certain monitoring methods.

116. Several studies indicate that self-evaluation among staff is an important tool in furthering the development of the practitioner and teaches staff to be more reflective. Self-evaluation may highlight those aspects of staff's practices that have been particularly effective (Cubey and Dalli, 1996), and it was found to lead to greater awareness of ongoing activities and pedagogical processes (Sheridan, 2001). Research in Italy found that systematic documentation and analysis of educational practice in self-assessments can be a useful tool for continuous support to the professionalism of early childhood education practitioners (Picchio et al., 2012).

117. A study in the United Kingdom examined the effectiveness of self-assessment as a method of monitoring, evaluating and enhancing the quality of service provision in day care settings (Munton, Mooney and Rowland, 1997). Self-assessment materials were produced, combining the views of different groups of stakeholders with an extensive review of day care research literature; and the materials were presented in a format consistent with a social learning theory analysis of experiential learning. Providers self-assessed aspects such as managing children's behaviour, helping children to learn and creating a warm and friendly atmosphere. Results of the evaluation study found no significant differences in the quality of day care provision between providers that had used the self-assessment materials and those that had not, with a small exception between control and intervention providers concerning staff-child interactions and staff skills, namely the tone of adult-child interactions, discipline and cultural awareness. There is the possibility that the quality of care provided by the intervention groups improved over the period of the

evaluation but in ways not assessed by the measures used. In general, the research concluded that a greater understanding of how providers implement self-assessment procedures and initiate changes in practice is required.

118. In New Jersey (United States), the introduction of a quality rating score allowed practitioners and management of the New Jersey Abbott Preschool Program to improve their practices, and statistically significant effects were found on children's literacy skills (Frede et al., 2007; Frede et al., 2009). To measure and assess programme progress, the Department of Education formed the Early Learning Improvement Consortium (ELIC) by bringing together a group of the state's top early childhood education faculty to collect and report on data on children and classrooms. ELIC hires and trains observers to conduct classroom observations on randomly selected Abbott preschool classrooms. Observation data have been systematically collected since the 1999-2000 school year, and results have been reported periodically since then. Classroom quality increased steadily each year, and by 2004-05, children were entering kindergarten with language and literacy skills closer to the national average than in prior years. This progress is attributed in part to rating tools used during observations, which provide staff with an indication of teaching practices that could require improvement and, hence, a basis for goal setting.

119. In the Flemish Community of Belgium, a process oriented self-evaluation instrument for staff in care settings **(**named SICS, see box 3.2) was introduced in 2004. Significant changes have been observed in the settings that use the self-evaluation instrument. Practitioners feel that the use of the instrument contributes to their professional development and teamwork. In their pedagogical approach, they indicated they learned to take into account the perspective of the child and because of this, to create optimal conditions for social-emotional and cognitive development (OECD, 2006). While these results are subjective, they indicate that monitoring can contribute to better perceived practices and knowledge.

120. With regarding to using test results (of those children old enough to be tested), researchers are not convinced that test results are sufficiently valid and reliable for making any fair conclusions on individual staff quality (Goe, 2007; Lockwood, Louis and McCaffrey, 2002; Waterman et al., 2012; Zaslow, Calkins and Halle, 2000). The fact teachers and caregivers, and staff-child interactions, matter for child outcomes and children's development does not necessarily indicate that child outcomes are the result of the instruction and activities of the professional. It is not merely the staff member that has an impact on children's learning outcomes, but also the home environment and environmental aspects such as noise and distractive behaviour of other children. Lastly, the impact of staff instruction is not limited to knowledge and skills which can be assessed through testing but also include transfer of psychological and lifelong learning skills (Barblett and Maloney, 2010; Isoré, 2009; Margo et al., 2008).

121. When monitoring is linked to professional development, it can have beneficial outcomes for staff performance and child outcomes. An evaluation of staff quality, involving 51 early childhood classes for preschool-age children throughout the United States, exposed weaknesses in the delivery of certain curriculum subjects, and as a result, staff training was developed and offered in those specific areas. Researchers found that training staff on curriculum subjects in which they were less competent and training them how to instruct children better in these subjects, led to better child outcomes in these subjects (Odom et al., 2010).

122. An analysis in the United States (Newark, New Jersey and Chicago, Illinois) found that classroom management can be a challenge for preschool teachers. As a response to this, an intervention was designed to train preschool teachers in managing children's behaviour and promote a more positive learning environment. It also included a classroom-level consultation component, which involved an external consultant monitoring implementation of the intervention through classroom observation. The consultant's data collection efforts were used to provide teachers with feedback on how to improve their practices. This intervention was found to reduce problem behaviour among children and improve their

social and emotional competence as measured by improvements in children's approaches to learning and executive function skills. However, no effects were found for children's mathematical and literacy development (Morris et al., 2013).

123. As shown in *Starting Strong III*, countries use a range of methods to monitor staff performance (e.g.., inspections, rating scales, checklists, surveys, self-assessments, observations), and these methods are administered by a range of actors (e.g., national/regional/local authorities, external inspectors, ECEC staff and/or management, parents) (OECD, 2012a). Hence, there are great differences in the design and implementation of monitoring approaches across and even within countries, which prevents drawing generic conclusions about the effects or impacts of monitoring staff quality *per se*. While there appears to be a general consensus on the need for effective monitoring of staff to improve ECEC services, perhaps the more complicated issue at hand is the design and implementation of the monitoring processes (see Chapter 6), as this seems to have the greatest effect on staff performance.

5.3 Monitoring child development/outcomes

124. Research in primary and secondary education has demonstrated that when teachers monitor children's development and progress, children learn more and teacher decision making improves (Safer and Fleishman, 2005). Literature points out that assessment of child development or outcomes is a critical component of the development and teaching/caring cycle (Barblett and Maloney, 2010). Evidence on the effects of monitoring child development in ECEC is scarcer, pointing to different impacts depending on the monitoring purpose or practice.

125. That there is little research available on the effects and impacts of monitoring children's development does not indicate that monitoring does not, or cannot, have a positive effect on the level of quality, staff performance, or child outcomes. Literature emphasises the importance of monitoring children's early development. Without a monitoring system, less information and knowledge on children's skills and progress would be available for staff and parents. This while the availability of such information and knowledge can improve staff practices and interactions with children, as well as lead to adaptations in curriculum or standards to better meet children's needs. Many OECD countries monitor children's development or outcomes through the use of portfolios, rating scales or checklist and narrative assessment tools. Testing of young children is not very popular among OECD countries (OECD, 2012a).

126. Literature indicates there are positive relationships between the use of non-formal monitoring practices such as observation, documenting through the use of portfolios or narrative assessments, and child outcomes (Bagnato, 2005; Meisels et al., 2003; Neisworth and Bagnato, 2004; Grisham-Brown, 2008). In the United States, a study was conducted to examine the impact of an outcomes-driven authentic monitoring model of child outcomes on classroom quality. The intervention approach in this study relied on the use of ongoing monitoring practices which were conducted in the children's natural contexts, provided information that is useful in planning for each child, and were aligned with the Head Start Child Outcomes Framework. Teachers in intervention and pilot classrooms implemented an approach (practice) that incorporated the use of a curriculum-based assessment tool, the development of portfolios aligned with the mandated Head Start Child Outcomes, and the integration of this child assessment information into individual and classroom instructional planning. These classrooms demonstrated significant improvements on some dimensions of classroom quality (as measured by the Early Language and Literacy Classroom Observation toolkit), whereas comparison classrooms exhibited no change (Hallam et al., 2007).

127. 128. The outcomes of the use of the QRIS are mixed, with some studies finding significant linkages between QRIS ratings and children's developmental outcomes and others not finding linkages. In Missouri (United States), children who participated in programmes with higher quality ratings and especially low-income children showed significantly higher gains on measures of social-emotional

development compared to children in programs with lower ratings (Thornburg et al., 2009). In contrast, in an evaluation of Colorado's Qualistar programme, linkages between the ratings and children's outcomes were not found (Zellman, Perlman and Setodji 2008).

129. Documented outcomes of the use of the Early Development Index (EDI, see box 5.1) in Australia include increased community awareness of the importance of early childhood development and more collaborative relationships between stakeholders. Staff reported that the results of EDI have informed their processes and practices in ECEC centres, and that they are better able to meet children's needs. In Vancouver, the outcomes of their own EDI led to development and implementation of support programmes for children and their parents to enhance early development. Parenting education and family literacy programmes have been set up as to improve literacy skills of young children, as well as inform parents on how they can stimulate early development at home. In addition, improved dental, vision and hearing assessment projects were developed as to track possible special needs at an early stage (Early Years Institute, 2012).

Box 5.1. The Early Development Index (EDI)

The Early Development Index is a population-level measure f children's development or well-being, originally developed in Ontario, Canada. Following Canada, other countries developed their own EDI based on their cultural and societal needs. As an example, Australia developed the Australian Early Development Index.

The Index consists of a teacher-completed checklist on children's development. The data is aggregated to a group level (school, neighbourhood, city, etc.) to provide a population-based measure of children's development. EDI is not reported at the child or class level and is not used as a diagnostic tool for individual children, or assessing individual children's school readiness. The checklist measures five key areas, or domains, of early childhood development:

- physical health and wellbeing
- social competence
- emotional maturity
- language and cognitive skills (school-based)
- communication skills and general knowledge

The EDI results allow local governments, communities or provisions to see how local children are doing relative to, or compared to other children in their community, and across a country (if implemented at country-level).

Sources: Early Years Institute website and Australian Government website (Department of Education, Employment and Workplace Relations)

130. However, monitoring child development to define 'school readiness' and with the purpose to postpone or deny kindergarten entry into school, can have negative impacts on early development. Retention labels children as 'failures' at the start of their formal educational career. There is no evidence in place which found that such practices have any academic benefits for children: by the end of the primary school level, children whose entry has been delayed do not perform better than peers who enter on time. Delaying a child's entry into school also appears to be a threat to children's social-emotional development since it denies them opportunities for cognitive growth through interaction with their age-mates. Those children were also found to show higher rates of behavioural problems (Bredekamp and Copple, 1997; Byrd, Weitzman and Auinger, 1997; NAEYC, 2001; Shore, 1998).

131. Regarding the practices, tools and instruments to monitor child development, a single monitoring practice at *one moment in time* is not a valid predictor of a child's potential either (CSSO, 2011). Because growth is more rapid in the period from birth to age eight then at other periods of development it is challenging to capture children's skills and abilities at any one point in time (Zaslow, Calkins and Halle,

2000). Monitoring children's outcomes should be based on multiple sources of information and should not be based on a single test score or other monitoring practice at one point in time (NAEYC, 2010). Most research points out that an ideal method of monitoring young children is through authentic, naturalistic observations that occur on an *ongoing* basis through the use of e.g. portfolios or narrative assessments. In addition, literature points more frequently to the importance of including the view of the child in assessing or evaluating their own development (e.g. in the form of self-confidence or satisfaction) or staff and service quality. While several countries, such as the Nordic ones, regard the view of the child as an important aspect of quality, this is an area which requires more research and reflection with regard to validity of instruments and results (CSSO, 2011; Meisels, 2007; NAEYC, 2010; Neisworth and Bagnato, 2004).

5.4 Monitoring curriculum implementation

132. There has been very little research conducted on the effects or impacts of monitoring curriculum implementation on the level of quality, staff performance, child development, or improvement in implementation. As with the other three topics described above, literature points out to the relevance of evaluating curricula and its implementation. A curriculum in place, and monitoring the curriculum implementation as well as progress in implementation, can assist ECEC leaders, managers, staff as well as policy-makers in strengthening curriculum implementation which may lead to better staff quality, better staff-child interactions and more appropriate activities and practices (OECD, 2012a; Danmarks Evalueringsinstitut, 2012). An evaluation of Norway's Framework Plan for the Content and Tasks of Kindergartens, initiated by the Ministry of Education and Research in Norway, showed that implementation of the Plan varies between settings (Østrem et al., 2009). Such information can contribute to a better understanding of what challenges settings face in implementation and point to needs of improvement of the curriculum, training needs, or other implementation supports.

133. There is little research available on this topic for ECEC. It is particularly difficult to study the impacts of monitoring one quality aspect (in this case curriculum implementation) and draw any conclusions on the relationship between a monitoring practice and quality or child outcomes. Further research on the importance of monitoring curriculum as well as its benefits or disadvantages for quality is needed to draw robust conclusions on its impacts.

134. When results of monitoring curriculum use and implementation are used by practitioners to design a child's learning environment, the child's success in ECEC is enhanced and a more stimulating, exciting learning environment is facilitated (Niemeyer and Scott-Little, 2001). Monitoring curriculum implementation seems particularly beneficial when combined with staff training or support. Monitoring results of the implementation of a comprehensive preschool curriculum targeting children's social-emotional competence, language, and emergent literacy, indicated that implementation quality increases over time if staff receives weekly coaching support in implementation (Domitrovich et al., 2010). And a better implementation of a curriculum subject, or subjects, results in better child outcomes in that area as was found in 51 preschools in the United States, where a better implementation of literacy led to improved literacy outcomes (Odom et al., 2010).

135. Research on monitoring curriculum implementation in third and fourth grades in the United States found that when staff monitor curriculum implementation through the use of checklists, portfolios, and feedback from parents as well as children, teaching enhanced and learning improved. Monitoring was conducted through self-analysis of curriculum implementation as well as planning, reviewing, and analysis of implementation in collaboration with their colleagues. Parents and children were also allowed to give feedback on staff practices. This resulted in better child literacy outcomes as well as improved math outcomes, although latter effects were rather small (Meisels et al., 2003).

6. DESIGN AND IMPLEMENTATION

136. This section discusses several aspects important for design and implementation of monitoring systems and practices based on the available literature. A range of design and implementation pointers which have been depicted in research are outlined below. In practice, additional factors to the ones indicated in research may be of importance when designing monitoring systems and practices, and implementing these. First-hand information from countries on designing and implementing monitoring systems, practices and tools will be collected in the next phase of the project and presented in an analytical report.

Defining quality

137. Literature indicates that to monitor quality, a country or programme needs to define quality. Any measurement or assessment of quality depends to a large extent on the way in which quality is defined (Woodhead, 1998). Definitions of quality may differ between countries since this is value- and cultural-based concept. Therefore, any definition of quality is subject to change over time and defining quality is an ongoing process. The definition of quality might therefore need revision at some point and should not be considered as a static concept (Kamerman, 2001).

Aligning purpose(s) and practice(s)

138. Policymakers and programme administrators need to be clear about the purpose or purposes for which they are developing monitoring systems and select the practices most appropriate for those purposes. The NAEYC (2010) noted that the purpose of monitoring, in general, should be to collect information that can be used in improving services and planning programmes and curriculum, as to ensure that children benefit from their early ECEC experiences. Monitoring practices should be chosen with great care since there is a risk of negatively affecting the validity and reliability when a monitoring practice designed for one purpose is used for other purposes (Ackerman and Barnett, 2005; Meisels, 2007). One single monitoring practice might not be reliable or valid for evaluating the level of quality of a provision, staff quality, child development, or the quality of curriculum implementation. This is especially important to consider when high-stakes decisions are involved (Meisels, 2007).

139. As an example, Sheridan et al. (2009) found that internal assessments of quality and external evaluations may lead to different quality results. Their analysis of 38 preschools in Sweden found that when preschools were externally evaluated as being of low or good quality tend to evaluate their own preschool quality as high (i.e. overestimate their own quality), while staff in preschools of high quality seem to underestimate their own quality (Sheridan et al., 2009). This research underpins the viewpoint that using one monitoring tool might not necessarily lead to reliable or valid monitoring results. Besides, when high stakes are involved, it is crucial that tools are validated in the context in which they are used, and not just translated and adapted from one country to another without validation studies.

140. Additionally, it is important to align monitoring practices that are used at different executive levels. In many countries, varying monitoring practices are in place at national, regional and municipal or centre level. When these practices are not aligned, varied purposes and content of monitoring can be in

place. Different views on what should be monitored and what constitutes quality may arise at different levels. Alignment of practices and purposes across different levels can ensure an equal understanding of what aspects should be monitored and lead to a more equal monitoring process by different evaluators.

Ensuring practical relevance

141. The NICHD (2002) recommends that monitoring practices have practical relevance, meaning that these are or can be linked to actual practices. The results of monitoring should lead to "usable" knowledge and policy-makers, managers or staff should be able to link the outcomes to practical activities to strengthen the ECEC system.

142. A common example often referred to in literature is linking monitoring practices and tools to the curriculum. This offers an opportunity to elevate the level of professionalism in the field of early education and care. When monitoring is linked to the curriculum content, the results can be used in practical manners: staff can adapt their instructions to be more in line with the curriculum or the curriculum can be revised to better meet the needs or capabilities of staff and children. An evaluation on monitoring among ECEC professionals also pointed out monitoring systems should forge a strong and useful connection with the curriculum as to ensure the practical relevance of monitoring (Horton and Bowman, 2002).

Involving stakeholders

143. Children's development depends greatly on the range of learning opportunities and developmental supports they receive. These opportunities occur in home and community-based environments, as well in early care and education programmes. Designing a monitoring system or practice should therefore include information from these environments as well. Literature emphasises the importance of parental involvement especially. To provide an accurate, useful picture of a child's knowledge and skills, a comprehensive monitoring practice should include information from the ECEC settings a child attends, parents, and community informants such as health and social services the child comes in contact with during its early years. Such collaborations also offer opportunities to establish common learning goals and priorities for the child in question. The more these stakeholders are included in the monitoring process the greater the ability of management and staff to make fully-informed decisions and adapt practices and curriculum to children's needs (Golan, Petersen and Spiker, 2008; National Association of School Psychologists, 2005).

144. Involvement of stakeholders is not merely useful when monitoring child development. The involvement of relevant stakeholders in monitoring service or staff quality, or even curriculum implementation, can contribute to greater parental engagement to ECEC and create a better understanding of the system as a whole or a service (OECD, 2012a). Besides, it gives stakeholders a sense of 'ownership' and it is a form of recognising and valuing their opinions. Their inputs can, as with involvement in monitoring child development, point to needs for improvement or give a better view on societal and parental expectations regarding ECEC. This can create greater buy-in and support for monitoring tools as was found in a review of the use of QRIS in five states in the United States (Oklahoma, Colorado, North Carolina, Pennsylvania, and Ohio). They attributed the success of implementation to stakeholder buy-in, political support, greater public awareness and adequate financing (Zellman and Perlman, 2008). However, stakeholders are not yet frequently involved in monitoring. An evaluation of ECE services in New Zealand found that in almost half of all services there was a lack of meaningful participation of parents and the community in monitoring practices (Education Review Office, 2007).

145. In addition, the involvement of researchers on monitoring may contribute to the success of a monitoring tool. A research in the United States on the design and implementation of a state-wide quality rating and improvement system (QRIS) in Indiana, showed that frequent communication between

evaluators of the tool (researchers) and programme implementers during the4-year evaluation project resulted in more appropriate fine-tuning of the programme, evaluators making needed adjustments in the design of the tool, and making more plausible interpretations of results (Elicker et al., 2013).

146. A challenge with involving stakeholders are the conflicting perceptions of what should be monitored and how. Policy-makers, researchers, ECEC managers, and ECEC staff may disagree on what aspects of quality should be monitored, how frequently, in what manner, and what stakes are attached to it. Policy-makers might want to impose certain monitoring practices, such as one inspection by external evaluators at least once every two years, while management and staff may oppose – especially when high stakes are involved with the implementation of a monitoring system. Although the presumption of monitoring is that the findings lead to improvement of quality, some critics indicate that adverse results are to be expected since inspections or other forms of assessment may cause stress to management and staff. In addition, opponents of monitoring argue that monitoring practices can lead to a narrowing of the curriculum which can decrease overall learning and development (Faubert, 2009).

Training evaluators on implementation of monitoring practices

147. Evaluators, those implementing monitoring practices, must be trained and monitored to apply monitoring practices and tools as to ensure tools are properly understood, used and to ensure practices render consistent and objective judgments (Waterman et al., 2012). Appropriate skills are required on the part of policy-makers, ECEC professionals and managers/leaders to use monitoring practices and collected information, and translate monitoring results into practice. There is evidence showing that staff that have been trained in implementation of monitoring practices, make less errors and their personal opinions influenced the monitoring results less frequently (Hoyt and Kems, 1999; Raudenbush et al., 2008). Besides, training on monitoring practices enhances the quality as well as the quantity of practices (Stuart et al., 2008), and staff are better able to use assessment for learning and development as was found in New Zealand (Mitchell, 2008).

Developmental appropriateness of monitoring child development practices

148. Literature points out to the importance of using tools that are developmentally appropriate (Meisels and Atkins-Burnett, 2000; NICHD, 2002; Sattler, 1998). Developmental appropriateness refers not merely to ensuring that practices are suitable for monitoring development and progress for a specific age, but also that these adequately monitor the multiple domains of child functioning and recognize children's language, culture, and learning needs (Raver, 2002; Snow, 2007; Espinosa and López, 2007).

149. Firstly, practices and tools for monitoring children's development should be adapted to capture children's learning needs, skills and abilities in accordance with their age. Young children are not generally capable of completing traditional paper-and-pencil tests; monitoring practices should be appropriate to either be used by children of a young age (if the child is to be involved in monitoring) or reflect the development domains important for early development. At different ages, children acquire and have different levels of knowledge and skills, which should be reflected in the monitoring practices (Waterman et al., 2012).

150. Secondly, child development is a compilation of numerous skills and capacities that vary within a population of children (Snow, 2007), and which are not limited to merely academic knowledge and cognitive skills (Barblett and Maloney, 2010). Physical well-being and motor development, social-emotional development, and approaches towards learning are development domains which are highly important for early child development. Moreover, children's progress in these developmental domains is both independent and highly interrelated (Snow, 2007). The social-emotional development of young children before, during, and after the transition to formal schooling is predictive of academic progress in

other domains. Unfortunately, many monitoring child development practices focus largely on language, literacy, and mathematics (Raver, 2002).

151. Moreover, children's linguistic and cultural differences must be considered in monitoring practices. Monitoring systems should be able to capture such contextual aspects of children's early learning and development (Espinosa and López, 2007). In addition, different expectations in different countries or regions might exist regarding what children should know or be able to do. The monitoring process should reflect societal values and beliefs regarding child development (CSSO, 2011). Making monitoring processes and tools more culturally relevant can be achieved through close cooperation between ECEC centres and family and community members with different cultural and linguistic backgrounds. This was found to be effective in improving intercultural relations (Oliver et al., 2011).

Piloting a system before full implementation

152. Conducting a pilot implementation is a cost-effective way to ensure viability and reliability of a monitoring system or practice before full implementation. A pilot phase allows the review of the system or practice and adjustments in light of potential flaws. Since further implementation problems might show up when the system is being fully implemented, which might lower the credibility of the system or practice, ECEC services and provisions for the pilot implementation should be chosen with caution (Isoré, 2009).

The dissemination of monitoring results

153. The dissemination and sharing of results with not merely the centre, management, or staff involved but also with the parents and broader general public may involve political and ethical issues (Visscher et al., 2000). Publishing monitoring results about quality might lead to labelling effects where centres, staff, or children with unfavourable monitoring results are labelled as bad performers. Supporters of making monitoring results of quality publically available indicate that this enables parents to make well-informed decisions on what provision their child should attend (Faubert, 2009).

154. In what way monitoring results are shared with the public should be carefully considered, based on the purpose of making such documents publically available. Publication of results often differs between monitoring results of quality or child outcomes and development. In most OECD countries, monitoring results of child development are not publically available, while monitoring and evaluation reports of centres may be offered online on the website of the national, regional or local government.

7. CONCLUSION

155. This paper investigated systems and practices in monitoring service quality, staff quality, child development and outcomes, and curriculum implementation. It provides an overview of the most common practices in place to monitor these quality aspects, explains purposes of monitoring systems and practices, lists who the users of these practices and their results are, and includes country examples. Research findings on the effects or impacts of the different monitoring practices and tools have been summarised, and implications for monitoring design and implementation are presented.

156. The report finds that different monitoring tools are in place, and that tools differ in accordance with the purpose or goal of monitoring. Monitoring can be a powerful tool in holding ECEC centres, management and staff accountable for their performance and may provide incentives for improvement of quality standards, staff practices and behaviour, and curriculum. Moreover, monitoring can be useful for identifying children's and staff's learning needs, as well as strengths and weaknesses of an ECEC system.

157. While literature supports the idea that monitoring and evaluation are critical for high-quality ECEC services, research on the effects of monitoring on quality or child development is only gradually emerging. The currently available research points to benefits of monitoring such as improved staff practices and better curriculum implementation, but the existence of such findings does not necessarily imply the ability of researchers to identify the impact of monitoring *per se*. It is often challenging to separate and identify the impact of a single monitoring practice, tool or method and further research on the impacts of monitoring is needed.

158. The challenge for policy-makers is to get the design and implementation right. Regardless of how a monitoring system is organised, monitoring practices and tools should be reliable, valid and fair. From the literature depicted, it can be concluded that one single monitoring practice is often not sufficient to draw conclusions about e.g. the level of quality of a provision or children's development. Monitoring practices and purposes need to be aligned and carefully chosen. Besides, ensuring the monitoring tool has practical relevance, that different stakeholders are involved in the monitoring system and that staff are properly trained to implement the monitoring practices were found to be important pointers in developing and implementing monitoring systems. Piloting before full implementation as well as in what way monitoring results should be shared with stakeholders and the general public, should be carefully considered.

159. Due to increasing public (and private) investments in ECEC, and increasing participation rates, monitoring is becoming more important in many OECD countries. This enhances the need for knowledge on how countries organise and design their monitoring systems, as well as how they implement them, a careful analysis on what challenges they have encountered in this, and what strategies they developed to overcome the challenges. The survey for the OECD Early Childhood and Schools (ECS) project '*OECD Review of Policies and Practices for Monitoring and Evaluating Quality in Early Learning and Development*' will collect information on these, and will contribute to the knowledge base on countries' monitoring practices. In addition to this, there is a need for further research on the benefits and disadvantages of monitoring practices, and its impacts on child development and staff performance.

BIBLIOGRAPHY

Adams, J. E. and M. Kirst (1999), "New Demands for Educational Accountability: Striving for Results in an Era of Excellence", in Murphy, J. and K.S. Louis (eds.), *Handbook of Research in Educational Administration*, Jossey-Bass, San Francisco, CA.

Ackerman, D.J. and W.S. Barnett (2005), *Prepared for kindergarten: What does "readiness" mean?*, National Institute for Early Education Research , New Burnswick, NJ.

Allen, S.F. (2007), "Assessing the Development of Young Children in Child Care: A Survey of Formal Assessment Practices in One State", *Early Childhood Education Journal*, Vol. 34, No. 6, pp 455-465.

Australian Government – Department of Education, Employment and Workplace Relations website on Australian Early Development Index: http://deewr.gov.au/australian-early-development-index accessed March 7, 2013.

Bagnato, S. (2005), "The authentic alternative for assessment in early intervention: An emerging evidence-based practice", *Journal of Early Intervention*, Vol. 28, No.1, pp 17–22.

Barblett, L. and C. Maloney (2010), "Complexities of Assessing Social and Emotional Competence and Wellbeing in Young Children", *Australasian Journal of Early Childhood*, Vol. 35, No. 2, pp 13-18.

Barnett, W.S., M.E. Carolan, J. Fitzgerald and J.H. Squires (2012), *The state of preschool 2012: State preschool yearbook*, National Institute for Early Education Research, New Brunswick, New Jersey.

Barnett W. S., D. J. Epstein, M. E. Carolan, J. Fitzgerald, D. J. Ackerman and A. H. Friedman (2010), *The state of preschool 2010: State preschool yearbook*, National Institute for Early Education Research, New Brunswick, NJ.

Bertram, A.D. and C. Pascal (1997), "A Conceptual Framework for Evaluating Effectiveness in Early Childhood Education", in M. Karlsson Lohmander (ed.), *Researching Early Childhood*, Vol. 3, Göteborg, Suécia, Universidade de Göteborg, pp 125-150.

Center for Advanced Study of Teaching and Learning (2011), *Measuring and Improving Teacher-Student Interactions in PK-12 Settings to Enhance Students' Learning*, Charlottesville, VA.

Flanagan, K., J. Beach and P.Varmuza (2013), *You bet we still care! A Survey of Centre-Based Early Childhood Education and Care in Canada - highlights report*, Child Care Human Resources Sector Council, Ottawa, Ontario.

Council of Chief State School Officers (CCSSO) (2011), "Moving Forward With Kindergarten Readiness Assessment Efforts: A Position Paper of the Early Childhood Education State Collaboration on Assessment and Student Standards", Washington, DC.

Cubey, P. and C. Dalli (1996), *Quality Evaluation of Early Childhood Education Programmes*, Occasional Paper No. 1, Institute for Early Childhood Studies, Wellington.

Dahlberg, G. and P. Moss (2005), *Ethics and Politics in Early Childhood Education*, RoutledgeFalmer, London and New York.

Danmarks Evalueringsinstitut (2012), *Læreplaner i praksis. Daginstitutionernes arbejde med pædagogiske læreplaner*, Danmarks Evalueringsinstitut, Denmark.

Domitrovich, C.E., S. D. Gest, D. Jones, S. Gill and R. M. Sanford DeRousie (2010), "Implementation quality: Lessons learned in the context of the Head Start REDI trial", *Early Childhood Research Quarterly*, Vol. 28, No. 3, pp 284-298.

Drange, N. and K. Telle (2011), *Gratis kjernetid i barnehage: Effekt på skolekarakterene til barn fra innvandrerfamilier*, Statistics Norway.

Drouin, C., N. Bigras, C. Fournier, H. Desrosiers and S. Bernard (2004), *Grandir en qualité 2003: Enquête québécoise sur la qualité des services de garde éducatifs*, Institut de la Statistique du Québec, Québec.

Early Years Institute website: http://www.eyi.org/ accessed March 7, 2013.

Early Years Institute (2012), *The Early Development Instrument (EDI). A Population-Based Measure of Children's Early Development*, PowerPoint Presentation on approach of EDI and its outcomes.

Education Review Office (ERO) (2007), *The Quality of Assessment in Early Childhood Education*, ERO, Wellington, New Zealand.

Edwards, C.P., S.M. Sheridan and L. Knoche (2008), *Parent engagement and school readiness: parent-child relationships in early learning*, Nebraska Center for Research on Children, Youth, Families and Schools, University of Nebraska, Nebraska.

Elicker, J., K.M. Ruprecht, C. Langill, J. Lewsader, T. Anderson and M. Brizzi (2013), "Indiana Paths to QUALITY™: Collaborative Evaluation of a New Child Care Quality Rating and Improvement System", *Early Education and Development*, Vol. 24, No. 1, pp. 42-62.Espinosa, L.M. and M.L. López (2007), *Assessment Considerations for Young English Language Learners Across Different Levels of Accountability*, Paper prepared for The National Early Childhood Accountability Task Force, Los Angeles, CA.

Faubert, V. (2009), *School Evaluation: Current Practices in OECD Countries and a Literature Review*, OECD Education Working Paper No. 42, Paris.

Firestone, W. (2002), "Educational Accountability", *Encyclopedia of Education*.

Forum for Early Child Development Monitoring. (n.d.). *Population Measures*. Retrieved 05/06/2014 from http://www.childdevelopmentmonitoring.net/population-measures

Frede, E., W. S. Barnett, K. Jung, C. E. Lamy and A. Figueras (2007), *The Abbott Preschool Program Longitudinal Effects Study (APPLES)*, Interim Report, National Institute for Early Education Research, New Brunswick, NJ.

Frede, E., K. Jung, W. S. Barnett and A. Figueras (2009), *The APPLES Blossom: Abbott Preschool Program Longitudinal Effects Study (APPLES) – Preliminary Results through 2nd Grade*, Interim Report, National Institute for Early Education Research, New Brunswick, NJ.

Frede, E. C., W. S. Gilliam and L. J. Schweinhart (2011), "Assessing Accountability and Ensuring Continuous Program Improvement: Why, How and Who" in E. Zigler, W. Gilliam and W. S. Barnett (eds.), *The pre-K debates: Current Controversies & Issues*. Baltimore, MD: Brookes Publishing.

Fukkink, R. (2011), "Prettiger in een goed pedagogisch klimaat", *Management Kinderopvang*, Vol. 11. No. 4, pp. 12-14.

Gatt, S., M. Ojala and M. Soler (2011), "Promoting social inclusion counting with everyone: Learning Communities and INCLUD-ED", *International Studies in Sociology of Education*, Vol. 21, No. 1, pp. 33-47.

Goe, L. (2007), *The link between teacher quality and student outcomes: A research synthesis*, National Comprehensive Center for Teacher Quality, Washington DC.

Golan, S., D. Petersen, and D. Spiker (2008), *Kindergarten Assessment Process Planning Report*, Washington State Department of Early Learning , Lacey, Washington.

Grisham-Brown, J. (2008), "Best practices in implementing standards in early childhood education", in A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V*, National Association of School Psychologists, Washington DC.

Gulbrandsen, L. and E. Eliassen (2013), *Kvalitet i barnehager: rapport fra en undersøkelse av strukturell kvalitet høsten 2012 [Quality in the kindergarten: report of a survey of structural quality in autumn 2012]*, NOVA [Norwegian Social Research], Oslo, Norway.

Hallam, R., J. Grisham-Brown, X. Gao and R. Brookshire (2007), "Effects of Outcomes-Driven Authentic Assessment on Classroom Quality", *Early Childhood Research & Practice*, Vol. 9, No. 2.

Harlen, W. (2007), "Criteria for Evaluating Systems for Student Assessment", *Studies in Educational Evaluation*, Vol. 33, No. 1, pp. 15-28.

Hauser-Cram, P., M.E. Warfield, C.C. Upshur and T.S. Weisner (2000), "An expanded view of program evaluation in early childhood intervention", in J. P Shonkoff and S. J. Meisels (Eds.), *Handbook of early childhood intervention*, Cambridge University Press, New York.

Helmerhorst , K.O.W., J. M. Riksen-Walraven, H. J. Vermeer, R.G. Fukkink and L.W.C. Tavecchio (2014) "Measuring the Interactive Skills of Caregivers in Child Care Centers: Development and Validation of the Caregiver Interaction Profile Scales", *Early Education and Development*, Vol. 25, No. 5, pp. 770-790.

Hidalgo, N.M., J.K. Epstein and S. Siu (2002), "Research on families, schools, and communities. A multicultural perspective", in Banks J. A. and C.A. Banks (Eds.), *Handbook of multicultural education*, Macmillan, New York.

Horton, C. and B.T. Bowman (2002), *Child assessment at the pre-primary level: Expert opinion and state trends*, Erikson Institute for Advanced Study in Child Development, Chicago, IL.

Hoyt, W. T. and M.D. Kems (1999), "Magnitude and moderators of bias in observer ratings: A meta-analysis", *Psychological Methods*, Vol. 4, pp. 403–424.

Isoré, M. (2009), *Teacher Evaluation: Current Practices in OECD Countries and a Literature Review*, OECD Education Working Paper No. 23, Paris.

Japel, C., R.E. Tremblay and S. Côté (2005). "Quality Counts! Assessing the Quality of Daycare Services Based on the Québec Longitudinal Study of Child Development", *IRPP Choices*, Vol. 11, No. 5.

Kamerman, S. B. (Ed.). (2001), *Early childhood education and care: International perspectives*, Columbia University, ICFP, New York.

Katz, L. G. and S.C. Chard, (1996), *The contribution of documentation to the quality of early childhood education*, ERIC Clearinghouse on Elementary and Early Childhood Education, Champaign, Illinois.

Kohnstamm Instituut (2012), *Pre-COOL cohortonderzoek. Technisch rapport tweejarigencohort, eerste meting 2010 – 2011*, Kohnstamm Instituut, Amsterdam.

Lee, J.-H. and D. J. Walsh (2004), "Quality in Early Childhood Programs: Reflections from Program Evaluation Practices", *American Journal of Evaluation*, Vol. 2, pp. 351-373.

Levitt, R., B. Janta and K. Wegrich (2008), *Accountability of Teachers: A Literature Review*, Rand Cooperation, Santa Monica, CA.

Litjens, I. and M. Taguma (2010), *Revised Literature Overview for the 7th Meeting of the Network on Early Childhood Education and* Care, OECD, Paris.

Lockwood, J., T. Louis, T. and D. McCaffrey (2002), "Uncertainty in rank estimation: Implications for value-added modelling accountability systems", *Journal of Educational and Behavioral Statistics*, Vol. 27, No. 3, pp 255–270.

Margo J., M. Benton, K. Withers and S. Sodha (2008), *Those who can?*, Institute for Public Policy Research, London.

Matthews, P. and P. Sammons (2004), *Improvement Through Inspection: An evaluation of the impact of Ofsted's work*, Office for Standards in Education, London.

McDonnell, L.M. and C. Choisser (1997), "Testing and Teaching: Local Implementation of New State Assessments", *CSE Technical Report 442*, Center for Evaluation, Standards, and Student Testing, Los Angeles, CA.

Meisels, S. J. and S. Atkins-Burnett (2000), "The elements of early childhood assessment", in J. P. Shonkoff and S. J. Meisels (Eds.), *Handbook of early childhood intervention*, Cambridge University Press, New York.

Meisels, S. J., S. Atkins-Burnett, Y. Xue, D.D. Bickel and S. Son, (2003), "Creating a system of accountability: The impact of instructional assessment on elementary children's achievement test scores", *Education Policy Analysis Archives*, Vol. 11, No. 9, pp 1–18.

Meisels, S.J. (2007), "Accountability in Early Childhood: No Easy Answers", in R.C. Pianta, M.J. Cox, K.L.Snow (Eds.), *School Readiness and the Transition to Kindergarten in the Era of Accountability*, Paul H. Brookes Publishing Co., Baltimore, Maryland.

Ministry of Education of Chile (Ministerio de Educación) website, accessed May 7 2013 at: http://www.evaluacioninicia.cl/

Mitchell, L. and K. Brooking (2007), *First NZCER national survey of early childhood education services*, New Zealand Council for Educational Research, Wellington.

Mitchell, L. (2008), *Assessment practices and aspects of curriculum in early childhood education. Results of the 2007 NZCER national survey for ECE services*, New Zealand Council for Educational Research, Wellington, New Zealand.

Morris, P., C.M. Lloyd, M. Millenky, N. Leacock, C.C. Raver and M. Bangser (2013), *Using Classroom Management to Improve Preschoolers' Social and Emotional Skills. Final impact and implementation findings from the foundations of learning demonstration in Newark and Chicago*, MDRC, New York.

Morris, A. (2011), *Student Standardised Testing: Current Practices in OECD Countries and a Literature Review*, OECD Education Working Paper No. 65, Paris.

Munté, A., O. Serradell and T. Sordé (2011), "From research to policy: Roma participation through communicative organization", *Qualitative Inquiry*, Vol. 17, No. 3, pp. 256-266.

Munton, A. G., A. Mooney and L. Rowland (1997), "Quality in Group and Family Day Care Provision: Evaluating Self-Assessment as an Agent of Change", *European Early Childhood Education Research Journal*, Vol. 5, No. 1, pp. 59-75.

National Association for the Education of Young Children (NAEYC) (2010), *Quality Rating and Improvement Systems (QRIS) Toolkit*, NAEYC, Washington, DC.

National Association of School Psychologists (2005), *Position statement on early childhood assessment*, NASP, Bethesda, MD.

National Center for Education Statistics (NCES) (1997), *Measuring the Quality of Program Environments in Head Start and Other Early Childhood Programs: A Review and Recommendations for Future Research*, Working Paper No. 97-36, Washington DC.

National Institute of Child Health and Human Development (NICHD) (2002). *Early childhood education and school readiness: Conceptual models, constructs, and measures*, Workshop Summary, Washington, DC.

Nederlands Consortium Kinderopvang Onderzoek (NCKO) website, accessed May 10 2013 at: http://www.kinderopvangonderzoek.nl/drupal/

Neisworth, J. and S.J. Bagnato (2004), "The mismeasure of young children: The authentic assessment alternative", *Infants and Young Children*, No. 17, pp 198–212.

Niemeyer, J. and C. Scott-Little (2001), *Assessing Kindergarten children: A compendium of assessment instruments*, SERVE, Tallahassee, Florida.

Norris, D. J. and L. Dunn (2004), *"Reaching for the Stars" Family Child Care Home Validation Study Final Report*, Early Childhood Collaborative of Oklahoma.

Norris, D. J., L. Dunn and L. Eckert (2003), *"Reaching for the Stars" Center Validation Study Final Report*, Early Childhood Collaborative of Oklahoma.

Odom, S. L., K. Fleming, K. Diamond, J. Lieber, M. Hanson, G. Butera, E. Horn, S. Palmer and J. Marquis (2010), "Examining different forms of implementation and in early childhood curriculum research", *Early Childhood Research Quarterly*, Vol. 25, No. 3, pp. 314-328.

OECD (2001), *Starting Strong: Early Childhood Education and Care*, OECD Publishing. doi: 10.1787/9789264192829-en.

OECD (2005), *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD Publishing. doi: 10.1787/9789264007413-en.

OECD (2006), *Starting Strong II: Early Childhood Education and Care*, OECD Publishing. doi: 10.1787/9789264035461-en.

OECD (2012a), *Starting Strong III: A Quality Toolbox for Early Childhood Education and Care*, OECD Publishing. doi: 10.1787/9789264123564-en.

OECD (2012b), *Education at a Glance 2012: OECD Indicators*, OECD Publishing. doi: 10.1787/eag-2012-en.

OECD Network on Early Childhood Education and Care (2011), "Survey for the Quality Toolbox and ECEC Portal", OECD, Paris.

OECD Network on Early Childhood Education and Care (2012a), *Draft Position Paper of the Thematic Working Group on Workforce Quality*, Background document for the 12th ECEC Network Meeting, OECD, Paris.

OECD Network on Early Childhood Education and Care (2012b), *Draft Position Paper of the Thematic Working Group on Monitoring Quality*, Background document for the 12th ECEC Network Meeting, OECD, Paris.

Office of Child Development and Early Learning (2010), "Keystone Stars: Reaching Higher for Quality Early Education", *Program Report*, Department of Public Welfare, Pennsylvania.

Ofsted (Office for Standards in Education, Children's Services and Skills) (2011), *The Annual Report of Her Majesty's Chief Inspector of Education Children's Services and Skills 2010/11*, Ofsted, London.

Oliver, E., L. de Botton, M. Soler and B. Merril (2011), "Cultural intelligence to overcome educational exclusion", *Qualitative Inquiry*,Vol. 17, No. 3, pp. 267-276.

Østrem, S., H. Bjar, L. Føsker Rønning, H. Hogsnes Dehnæs, T. Jansen Thorsby, S. Nordtømme, S., and K. Tholin Rydjord (2009), *Alle teller mer. En evaluering av hvordan Rammeplan for barnehagens innhold og oppgaver blir innført, brukt og erfart [Everybody counts more. An evaluation of how the Framework Plan for the Content and tasks are introduced, used and experienced]*, Høgskolen i Vestfold 1/2009 [Vestfold University College], Tønsberg, Norway.

Peterson, K., C. Wahlquist, J. Esparza Brown and S. Mukhopadhyay (2003), "Parents Surveys for Teacher Evaluation", *Journal of Personnel Evaluation in Education*, Vol. 17, No. 4, pp 317-330.

Philips, D., D. Mekos, S. Scarr, K. McCartney and M. Abbott-Shim (2000), Within and Beyond the Classroom Door: Assessing Quality in Child Care Centers, *Early Childhood Research Quarterly*, Vol. 15, No. 4, pp. 475-496.

Pianta, R. C., A. J. Mashburn, J. R. Downer, B. K. Hamre and L. Justice (2008),"Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms", *Early Childhood Research Quarterly*, Vol. 23, pp 431-451.

Picchio, M., D. Giovannini, S.Mayer and T. Musatti (2012), "Documentation and analysis of children's experience: an ongoing collegial activity for early childhood professionals, *Early Years: An International Research Journal*, Vol. 32, No. 2, pp, 159-170.

Pramling, N. and I. Pramling Samuelsson (2011), *Educational encounters: Nordic studies in early childhood didactics*, Springer, Dordrecht, Netherlands.

Pramling Samuelsson, I. and M. Asplund Carlsson (2008), The playing learning child: Towards a pedagogy of early childhood, *Scandinavian Journal of Educational Research*, Vol. 52, No. 6, pp 623-641.

Puigvert, L., M. Christou and J. Holdford (2012), "Critical Communicative Methodology: including vulnerable voices in research through dialogue", *Cambridge Journal of Education*, Vol. 42, No. 4, pp. 513-526.

RAND (2008), "Assessing Child-Care Quality: How Well Does Colorado's Qualistar Quality Rating and Improvement System Work?", *Policy Brief*, RAND Corporation, Santa Monica, CA.

Raudenbush, Sd. W., A. Martinez, H. Bloom, P. Zhu, P. and F. Lin (2008), *An eightstep paradigm for studying the reliability of group-level measures*, Working Paper, University of Chicago.

Raver, C. (2002), "Emotions Matter: Making the case for the role of young children's emotional development for school readiness", *Society for Research in Child Development's Social Policy Report*, Vol. 16, No. 3, pp 3-19.

Rivas, S. and A. Sobrino (2011), "Determining quality of early childhood education programmes in Spain: a case study", *Revista de Educación*, No. 355, pp 257-283.

Rosenkvist, M. A. (2010), *Using Student Test Results for Accountability and Improvement: A Literature Review*, OECD Education Working Paper No.54, Paris.

Ruhm, C. and J. Waldfogel (2012), "Long-term effects on early childhood care and education", *Nordic Economic Policy Review*, No. 1, Nordic Council of Ministers.

Safer, N. and S. Fleishman (2005), "Research Matters: How Student Progress Monitoring Improves Instruction", *How Schools Improve*, Vol. 62, No. 5, pp 81-83.

Sattler, J. R. (1998), *Assessment of children (3rd ed.)*, J. R. Sattler Publishing, San Diego, CA.

Sheridan, S. (2011), "Pedagogical quality in preschool: A commentary", in N. Pramling and I. Pramling Samuelsson (Eds.), *Educational encounters: Nordic studies in early childhood didactics*, Springer, Dordrecht, The Netherlands.

Sheridan, S., I. Pramling Samuelsson and E. Johanssson. (Eds.) (2009), *Barns tidiga lärande. En tvärsnittstudie av förskolan som miljö för barns lärande* [Children's early learning. A current study of preschool as an environment for children's learning], Acta Universitatis Gothoburgensis, Göteborg, Sweden.

Sheridan, S. (2009), "Discerning pedagogical quality in preschool", *Scandinavian Journal of Educational Research*, Vol. 53, No. 3, pp. 245-261.

Sheridan, S. (2001), "Quality Evaluation and Quality Enhancement in Preschool - A Model of Competence Development", *Early Child Development and Care*, No. 166, pp. 7-27.

Sheridan, S. (2001), "Quality Evaluation and Quality Enhancement in Preschool - A Model of Competence Development", *Early Child Development and Care*, No. 166, pp. 7-27.

Sheridan, S. (2001), *Pedagogical quality in preschool: An issue of perspectives*, Acta Universitatis Gothoburgensis, Göteborg, Sweden.

Sheridan, S. (2000), "A comparison of external and self-evaluations of quality in Early Childhood Education", *Early Child Development and Care*, No. 164, pp. 63-78.

Shore, R. (1998), *Ready schools*, National Education Goals Panel, Washington, DC.

Siraj-Blatchford, I. (2007), "Creativity, communication and collaboration: The identification of pedagogic progression in sustained shared thinking", *Asia- Pacific Journal of Research in Early Childhood Education*, No. 2, pp. 3-23.

Siraj-Blatchford, I., K. Sylva, S. Muttock, R. Gilden and D. Bell (2002), *Researching Effective Pedagogy on the Early Years*, Research report 356, Institute of Education, University of London, London.

Snow, K.L. (2007), "Integrative views of the domains of child function: Unifying school readiness", in R.C. Pianta, M.J. Cox and K.L. Snow (Eds.), *School Readiness & the Transition to Kindergarten in the Era of Accountability*, Paul H. Brookes Publishing Co., Baltimore, Maryland.

Stuart, D., H. Aiken, K. Gould and A. Meade (2008), *Evaluation of the Implementation of Kei Tua o te Pae Assessment for Learning: Early Childhood Exemplars: Impact evaluation of the Kei Tua o te Pae 2006 professional development*, Ministry of Education New Zealand, Wellington.

Sylva, K., E. Melhuish, P. Sammons, I. Siraj-Blatchford and B. Taggart (2010), *Early Childhood Matters. Evidence from the Effective Pre-school and Primary Education project*, Routledge, London.

Taguma, M., I. Litjens and K. Makowiecki (2012), *Quality Matters in Early Childhood Education and Care: New Zealand*, OECD, Paris.

Thornburg, K.,, W.A. Mayfield, J.S. Hawks and K.L. Fuger (2009). *The Missouri Quality Rating System School Readiness Study*. University of Missouri--Columbia. Center for Family Policy and Research.

Tout, K., M. Zaslow, T. Halle and N. Ferry (2009), "Issues for the Next Decade of Quality Rating and Improvement Systems", *Issue Brief No. 3*, Office of Planning, Research and Education, U.S. Department of Health and Human Services, Washington, DC.

Visscher, A., S. Karsten, T. de Jong, and R. Bosker (2000), "Evidence on the intended and unintended effects of publishing school performance indicators", *Evaluation and Research in Education*, No. 14, pp 254-267.

Waterman, C., P.A. McDermott, J.W. Fantuzzo and J.L. Gadsden (2012), "The matter of assessor variance in early childhood education - Or whose score is it anyway?", *Early Childhood Research Quarterly* Vol. , No. 1, pp 46– 54.

Weiss, H., M. Caspe and M. E. Lopez. (2008), "Family Involvement Promotes Success for Young Children: A Review of Recent Research", in M. M. Cornish (ed.), *Promising Practices for Partnering with Families in the Early Years*, Information Age Publishing, Plymouth.

Winswold, A. and L. Gulbrandsen (2009), *Kvalitet og kvantitet. Kvalitet i en barnehagesektor i sterk vekst [Quality and quantity. Quality in a nursery sector growing strongly]*, NOVA [Norwegian Social Research], Oslo, Norway.

Woodhead, M. (1998), "Quality in early childhood programmes: a contextually appropriate approach", *International Journal of Early Years Education*, Vol. 6, No. 1, pp 5-17.

Zaslow, M., K. Tout, T. Halle and N. Forry (2009), "Multiple Purposes for Measuring Quality in Early Childhood Settings: Implications for collecting and communicating information on quality", *Issue Brief No. 2*, Office of Planning, Research and Evaluation, U.S. Department of Health and Human Services, Washington, DC.

Zaslow, M., J. Calkins and T. Halle (2000), *Background for community-level work on school readiness: A review of definitions, assessments, and investment strategies. Part I: Defining and assessing school readiness – building on the foundation of NEGP Work*, Child Trends Inc., Washington DC.

Zellman, G. L., M. Perlman, V.-N. Le and C. M. Setodji (2008), *Assessing the Validity of the Qualistar Early Learning Quality Rating and Improvement System as a Tool for Improving Child-Care Quality*, RAND Corporation, Santa Monica, CA.

Zellman, G. and M. Perlman (2008), *Child care Quality Rating and Improvement Systems in five pioneer states: implementation issues and lessons learned*, RAND Corporation, Santa Monica, CA.

Zucker, S. (2004), Administration *Practices for Standardized Assessments*, Pearson Assessment Report.