Routledge
Taylor & Francis Group

Early Child Development and Care

ISSN: 0300-4430 (Print) 1476-8275 (Online) Journal homepage: www.tandfonline.com/journals/gecd20

Measuring interactional quality in pre-school settings: introduction and validation of the Sustained Shared Thinking and Emotional Wellbeing (SSTEW) scale

Steven J. Howard, Iram Siraj, Edward C. Melhuish, Denise Kingston, Cathrine Neilsen-Hewett, Marc de Rosnay, Elisabeth Duursma & Betty Luu

To cite this article: Steven J. Howard, Iram Siraj, Edward C. Melhuish, Denise Kingston, Cathrine Neilsen-Hewett, Marc de Rosnay, Elisabeth Duursma & Betty Luu (2020) Measuring interactional quality in pre-school settings: introduction and validation of the Sustained Shared Thinking and Emotional Wellbeing (SSTEW) scale, Early Child Development and Care, 190:7, 1017-1030, DOI: 10.1080/03004430.2018.1511549

To link to this article: https://doi.org/10.1080/03004430.2018.1511549

Published online: 17 Aug 2018.

Submit your article to this journal

Article views: 5755

View related articles

View Crossmark data

Citing articles: 22 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=gecd20

EARLY CHILD DEVELOPMENT AND CARE
2020, VOL. 190, NO. 7, 1017–1030
https://doi.org/10.1080/03004430.2018.1511549

Routledge
Taylor & Francis Group

Check for updates

Measuring interactional quality in pre-school settings: introduction and validation of the Sustained Shared Thinking and Emotional Wellbeing (SSTEW) scale

Steven J. Howard [a,c], Iram Siraj [a,b], Edward C. Melhuish[b,c], Denise Kingston[d], Cathrine Neilsen-Hewett[a], Marc de Rosnay [c], Elisabeth Duursma[a] and Betty Luu [a]

[a]Early Start and School of Education, University of Wollongong, NSW, Australia; [b]Department of Education, University of Oxford, Oxford, UK; [c]Early Start and School of Psychology, University of Wollongong, NSW, Australia; [d]Department of Education, University of Sussex, Brighton, UK

ABSTRACT
Research increasingly acknowledges the importance of high quality interactions that support and extend children's thinking. Few measurement tools currently exist, however, to capture this specific aspect of process quality. The Sustained Shared Thinking and Emotional Wellbeing (SSTEW) scale was developed to assess interactional quality in early childhood education and care, and it includes dimensions of process quality based on developmental theories and practice in effective settings. This study compared ratings on the SSTEW and Early Childhood Environment Rating Scale – Extension (ECERS-E) to consider the impact of varying levels of curricular and interactional quality on child development in 45 Australian pre-school centres; namely the language, numeracy and socio-behavioural development of 669 children at the end of their pre-school year. Results indicated a level of predictive validity for interactional quality ratings as measured by SSTEW which, while related to curricular quality ratings on ECERS-E, differed in associations across domains of child development.

ARTICLE HISTORY
Received 15 June 2018
Accepted 9 August 2018

KEYWORDS
Preschool; sustained shared thinking; environment rating; process quality; child development; predictive validity

Introduction

The notion of 'quality' within Early Childhood Education and Care (ECEC) remains a contested term (Siraj-Blatchford & Wong, 1999). The main objection to the use of the term 'quality' lies within its subjectivity. It is well documented that different people view ECEC quality in different ways and value different aspects of provision (Penn, 2011). For example, parents/carers commonly value the proximity of an ECEC setting to their homes and perceived happiness of children (Plantenga, 2011). Others, such as policymakers, often prioritize the setting's ability to support parents/carers back into the workforce (West, Roberts, & Noden, 2010).

There is a large body of research in which quality has been defined more objectively to avoid such differences in emphasis (e.g. Melhuish et al., 2015; Pianta et al., 2005; Sylva et al., 2014). For instance, quality can be defined in relation to children's learning and development, such that a high quality ECEC setting is one that supports and enhances children's outcomes (Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2004). While this objective definition adds clarity, it raises some questions: for example, which aspects or characteristics of ECEC settings are associated with enhancing children's learning outcomes; how can quality be measured reliably; and do ratings on quality measures

CONTACT Steven Howard stevenh@uow.edu.au School of Education, University of Wollongong, New South Wales, 2522, Australia

© 2018 Informa UK Limited, trading as Taylor & Francis Group

predict outcomes for the children who attend? This study aims to answer these questions in relation to a new quality measurement tool, the Sustained Shared Thinking and Emotional Wellbeing (SSTEW) scale (Siraj, Kingston, & Melhuish, 2015), and its ability to index interactional quality in ECEC and predict impact on children's outcomes.

Background

In research that investigates associations of ECEC quality with children's learning and development (Siraj-Blatchford et al., 2003; Sylva et al., 2004), quality is measured by considering process and structural aspects of quality (Donabedian, 1980). Process aspects seek to capture the child's everyday experiences – what it feels like to be a child in the environment (Katz, 2008). This includes the social, emotional, physical and instructional aspects of the setting's activities, opportunities, and interactions available to the children, which are mediated by the educators, other children, and the materials available and accessible within the setting (Howes et al., 2008; Slot, Leseman, Verhagen, & Mulder, 2015). Structural aspects include the more general, and often regulated, aspects in ECEC – including child/adult ratios, educators' qualifications, and group sizes (Howes et al., 2008; Slot et al., 2015). Yet, Fukkink and Lont (2007), in a meta-review of the impact of educator qualifications and professional development on quality, noted that it was educators' ability to create high quality pedagogic environments and interactions that made the difference for children's development, rather than the structural aspects of their qualifications. What appears to be critical is the educator's sensitivity and responsiveness, the quality of their interactions, and their ability to extend and scaffold children's learning and thinking (Organisation for Economic Co-operation and Development, 2012).

In order to measure structural and process aspects, researchers often use Environment Rating Scales (ERS). These increase the objectivity of observations and allow comparisons across studies internationally. As a result, ERS have developed an international reputation for: (i) measuring important aspects of ECEC quality associated with children's outcomes; (ii) the standardization processes they have undergone; and (iii) well-established psychometric properties (e.g. validity, reliability). ERS support the view that there are indeed complex interactions between the two aspects of quality, and they also point to structural quality as a possible pre-condition for high process quality. Nevertheless, the international research base increasingly demonstrates that the process aspects of adult-child and child-child interactions are the most powerful predictors of children's outcomes (Melhuish et al., 2015; Siraj & Kingston, 2015). Structural quality is then seen as important because of the ECEC characteristics it identifies (e.g. adult-child ratios, training, qualifications) impact on process quality. This suggestion is supported by research showing that ERS emphasizing curricula (i.e. Early Childhood Environment Rating Scale-Extension; ECERS-E) are more predictive of children's academic outcomes than scales with a greater emphasis on structural quality (i.e. Early Childhood Environment Rating Scale-Revised; ECERS-R) (Sylva et al., 2004).

Some key studies in this evidence base have emphasized the importance of high quality interactions that support and extend children's thinking (termed 'sustained shared thinking': Siraj et al., 2017; Siraj-Blatchford et al., 2003). Sustained shared thinking (SST) refers to 'an interaction where two or more individuals 'work together' in an intellectual way to solve a problem, clarify a concept, evaluate an activity, or extend a narrative. Both parties must contribute to the thinking, and it must develop and extend the understanding' (Siraj-Blatchford, Sylva, Muttock, Gilden, & Bell, 2002, p. 8). SST occurs when an interaction supports and extends a child or children's thinking and includes successful support of children's communication, language, thinking, and learning. International research has endorsed the view that such interactions are key to practice if children's learning and development are to be enhanced (Katz, 2008; Pianta, 2012; Sylva et al., 2014). As a result, SST is acknowledged in many early years frameworks across the world. For instance, SST's influence is reflected in the development of the English Early Years Foundation Stage (Department for Education, 2012) and the Australian Early Years Learning Framework (Department of Education, Employment and Workplace Relations, 2009). Despite this, practices associated with SST are relatively poor and

hard to find (Siraj-Blatchford et al., 2002; Sylva et al., 2004). The need to clarify and support understanding, recognition and practices related to SST is now a priority.

These discrepancies between research and practice, and newer understandings about the importance of process quality, have led to the development of tools designed to specify and capture pedagogies and practices associated with SST. The SSTEW scale (Siraj et al., 2015), for example, was designed to support, increase, and improve the practice and identification of episodes of SST within ECEC settings (with children aged 2–5). It was developed from the evidence base generated by the Researching Effective Pedagogy in the Early Years (REPEY) study (Siraj-Blatchford et al., 2002), and from international evidence that, while not using the term SST, also pointed to the importance of high-quality interactions (Bodrova & Leong, 2007; Downer et al., 2012; Mashburn et al., 2008). While most existing ERS deal with staff-child interactions, they do not take account of important characteristics of interactions in the way SSTEW intends. That is, SSTEW was designed to extend and add to the existing ERS; to support the recognition and differentiation of more- and less-effective interactions within ECEC settings, where effectiveness is defined by its influence on young children's developmental outcomes (Siraj-Blatchford et al., 2002).

SSTEW is underpinned by principles and practices that reflect the evidence base around effective practice. It is grounded in current thinking and developmental theories (see, for example, Mayer & Beckh, 2016; Moreno, Shwayder, & Friedman, 2017) and considers: relational pedagogy and practices that encourage educators to engage in responsive, warm, and affectionate interactions with children (see Ainsworth, 1979; Bowlby, 1969; Deci & Ryan, 1985; Hamre & Pianta, 2001; Ryan & Deci, 2000); intentional pedagogy that supports and extends critical thinking, reasoning, and problem solving (e.g. Bodrova & Leong, 2007; Vygotsky, 1978); and concept development and encouragement for learning (e.g. Downer et al., 2012; LoCasale-Crouch et al., 2007; Mashburn et al., 2008).

SSTEW recognizes the importance of child-centered and developmentally appropriate practice that shows progression in learning in order to support children in becoming self-regulated, autonomous learners. It focuses on relational and intentional pedagogy, and links successful interactions and SST to a deep knowledge and understanding of effective ECEC pedagogies and practices. Higher scores are awarded in SSTEW when: staff show that they know individual children well, including their interests, beliefs, cultures, and achievements; there is a culture in the setting that supports children's curiosity, thinking, and questioning; children are seen to engage in appropriate, cognitively challenging activities and discussions with the educators and each other; educators support confidence, risk taking, and autonomy in the children's learning; each child is supported according to their needs; and educators show a range of different teaching and learning strategies together with a comprehensive and relevant content knowledge that they apply flexibly with contextual, individual, and socio-cultural sensitivity. Finally, SSTEW recognizes the power of intergenerational components of educators' work, as educators work with parents and carers to support quality interactions in the Home Learning Environment (Kingston & Siraj, 2017).

Like other ERS, the SSTEW scale can be used for research, audit, and inspection or by educators themselves as a self-assessment tool and part of their own quality improvement processes. Thus, the SSTEW scale is a tool developed to bridge the gap between research and practice, and support understanding, integrity, and sustainability in the area of high quality interactions. It was designed to support educators, external quality assurance personnel, and researchers in assessing early interactions (including aspects of pedagogy and practice which support or hinder these), as well as promoting improvement and professional development through identification of strengths and areas requiring attention. However, given that the SSTEW scale has only recently been created, it is important to determine its associations with children's outcomes (in alignment with conceptions of quality in relation to child outcomes; Sylva et al., 2004).

The current study represents the first attempted validation of the SSTEW scale, and as such it seeks to introduce the scale and evaluate its predictive validity within key domains of child development. Specifically, this study evaluates the SSTEW scale's reliability (i.e. internal consistency), its concurrent validity with another established measure of process quality (i.e. ECERS-E), and its predictive validity

in relation to key child development outcomes nearly a year later. While current evidence suggests the strength of curricular quality (as measured by ECERS-E) over more structural aspects of quality for predicting children's developmental progress (Sylva et al., 2004), it was expected that SSTEW would show at least an equivalent predictive strength in accounting for children's subsequent developmental progress. This was based on the background, development and links to the effectiveness evidence-base shared by both scales. If found, this would reinforce and substantiate the importance of interactional quality, in addition to and independent of curricular quality, for enhancing aspects of children's early development.

Materials and method

Centre and participant sample

Data from 45 ECEC centres were drawn from the control group of a larger cluster RCT study–the Fostering Effective Early Learning (FEEL) study (Melhuish et al., 2016)–for the purposes of the current investigation and analysis. These centres were recruited to ensure diversity of representation across: centre quality (as indexed by Australia's statutory National Quality Standard (NQS) ratings, whereby assessors for the regulatory authority identify centres as Working Towards, Meeting or Exceeding these standards), location (i.e. metro, regional), centre type (i.e. long-day care, preschool) and socioeconomic areas (Deciles 1–8, based on the Australian Bureau of Statistics' area-level Socio-Economic Indexes for Areas; SEIFA). Table 1 summarizes the centres' baseline characteristics. These 45 centres yielded a sample of 54 pre-school rooms and 669 children, which corresponds to 50.5% of eligible children having parent consent to participate in the study and 100% of consented children participating in the study.

The children ranged in age from 3.10 to 5.69 years, with a mean age of 4.59. Girls were modestly underrepresented in this sample ($n = 297$ girls; 44.4%), while the proportion of children who were identified as of Aboriginal or Torres Strait Islander origin (4.1%) aligned closely with Australian census data for this age (Australian Institute of Health and Welfare, 2012). Most children spoke English at home (91.4%), with the remaining children speaking one of 12 different languages. Most children's parents were born in Australia (88.0%); the next most frequent being the United Kingdom (1.9%), New Zealand (1.6%), and India (1.4%).

Measures

Environment quality ratings

Environmental quality ratings were conducted by highly trained observers throughout a one-day observation of each pre-school room in all the participating centres. Observers took a fly-on-the-wall approach to observation, so as to observe (but not influence) typical practice. Further, given a one-day observation in isolation limits the ability to index typical practices over time, observation days also involved an in-depth review of programming, records, and other relevant materials

Table 1. Pre-School Centre Characteristics.

Centre Hubs	22 from southern regional hub (within proximity, and inclusive of major metropolitan areas) 14 from a northern regional hub (within proximity, and inclusive of major metropolitan areas) 9 from a north-eastern regional hub
ECERS-E	T1: $M = 3.09$, $SD = 0.94$; T2: $M = 3.19$, $SD = 1.12$
SSTEW	T1: $M = 3.96$, $SD = 1.25$; T2: $M = 3.83$, $SD = 1.28$
Geographic Location	18 regional, 27 metro
Service Type	31 LDC, 14 preschool
NQS Rating	12 WT, 14 M, 18 EX (1 not yet rated)
SEIFA Decile	$M = 3.89$ (49% Deciles 1–3)

Note. LDC = long-day care. NQS = National Quality Standard. WT = Working Towards, M = Meeting, and EX = Exceeding the NQS.

(e.g. portfolios, day books) to inform quality ratings. A 30–60 min interview with centre leadership (i.e. the Centre Director and/or Room Leader) also informed the final quality ratings. All observers were trained intensively for five days, including in-field practice ratings with a highly experienced trainer/observer, followed by rigorous inter-rater reliability checks. Before entering the field, observers had to meet a rigorous inter-rater reliability standard against a highly experienced trainer/observer: (1) an intra-class correlation exceeding .70 ($M = .86$); (2) a correlation exceeding .70 ($M = .86$); (3) a mean difference in ratings less than 0.75 ($M = 0.43$); and (4) an agreement of ratings (within 1 point) of at least 80% ($M = 93\%$). Observers who were certified as reliable against these standards administered two ERS in the field: ECERS-E (Sylva, Siraj-Blatchford, & Taggart, 2010) and SSTEW (Siraj et al., 2015).

Early Childhood Environment Rating Scale – Extended (ECERS-E). ECERS-E contains 18 elements which combine to create four subscales: Literacy, Mathematics, Science/Environment, and Diversity. These subscales were derived from Curriculum Guidance for the Foundation Stage (Department for Education and Employment/Qualifications and Curriculum Authority, 2000) - the English curriculum document available at the time of the scale development. ECERS-E was tailored specifically to tap dimensions of quality defined both by England's curriculum and by notions of emergent academic skills at that time. Using on-balance judgements derived from a one-day room observation, each element is rated from 1 (inadequate practice) to 7 (excellent practice) based on patterns of the presence or absence of that element's constituent indicators. ECERS-E has been shown to have good reliability and predictive validity of child development progress at school entry (Sylva et al., 2006). The elements within each subscale are averaged to create four subscale scores; and these four are then averaged to generate an overall scale score.

Sustained Shared Thinking and Emotional Wellbeing (SSTEW) scale. SSTEW brings together different dimensions of the ECEC environment to consider practice that supports children aged 2–5 to develop skills in sustained shared thinking and emotional wellbeing. SSTEW contains 14 elements, which combine as five sub-scales related to two domains of development. While these domains cannot be separated completely, the focus of the first domain, social-emotional development, is found predominantly in two sub-scales: (1) building trust, confidence, and independence; and (2) social and emotional wellbeing. The focus for the second domain, cognitive development, is found mostly in three sub-scales: (3) supporting and extending language and communication; (4) supporting learning and critical thinking; and (5) assessing learning and language. Like ECERS-E, each element is rated from 1 (inadequate practice) to 7 (excellent practice) based on the pattern of presence/absence of indicators. The elements are averaged to yield five subscale scores and these five are then averaged to generate an overall scale score.

Child assessments

Highly trained assessors, who had undertaken one full day of training and extensive in-field administration practice with feedback, conducted child assessments. Assessments were selected within domains established as important for child outcomes, namely: language, early numeracy, self-regulation, and social-behavioural development. Assessors who administered child assessments were different from environmental raters. In all cases, two fieldworkers visited each centre to complete the child assessments.

Differential Ability Scales (DAS-II). Two DAS-II subscales were administered to capture aspects of language and number concept development: verbal comprehension and early number concepts. The verbal comprehension subtest, consisting of 42 items, requires children to identify and manipulate objects in response to verbal instructions. Early number concepts, comprising 33 items, requires children to count, identify digits and quantities, perform mathematical operations, and demonstrate knowledge of basic numerical concepts. In both cases, administration continues until the earlier of completion or non-satisfaction of a performance threshold at identified stop rule junctures. DAS is appropriate for using with children aged 2.5–17 years, and has shown good reliability and validity in children within and outside of typical development ranges (Elliott, 2007).

Preschool Early Numeracy Scale (PENS). Four PENS subscales were administered to capture elements of early numeracy not assessed by DAS-II, specifically: one-to-one counting, counting subsets, number order, and set-to-numerals. Twenty-one items were administered in total. PENS was designed for use with children from 3 years of age, with good reliability and predictive validity (Purpura & Lonigan, 2015).

Early Years Toolbox (EYT) Expressive Vocabulary. EYT Vocab is a 54-item measure of a child's expressive vocabulary development that requires children to produce verbally the correct label for each stimulus depicted on an iPad (Howard & Melhuish, 2017). Participants respond verbally and a data collector records this response within the app. In cases of an incorrect label, the data collector prompts participants by asking 'what else might this be called' until either a correct production or an indication that the child is unable to produce the required word. The measure ceases at the earlier of completion or six consecutive incorrect responses. This assessment has been used successfully with children aged 2.5–6 years, with good internal consistency and concurrent validity in a large and demographically diverse sample (Howard & Melhuish, 2017). It is also unique in its ability to leverage affordances of technology (e.g. standardized administration and automated scoring via iPad, to reduce inter-assessor variability) and brevity (i.e. ~5 minutes to complete).

Strengths & Difficulties Questionnaire (SDQ). The 25-item educator-report version of the SDQ was implemented, which yields subscales of prosocial behaviour (e.g. 'considerate of other people's feelings'), hyperactivity (e.g. 'restless, overactive'), peer problems (e.g. 'has at least one good friend'), conduct problems (e.g. 'often fights with other children') and emotional problems (e.g. 'many fears, easily scared') (Goodman, Meltzer, & Bailey, 1998). Educators rate items according to the frequency with which a child engages in the identified behaviour, ranging from 0 (Not True) to 2 (Certainly True). SDQ has strong reliability and validity in diverse international samples (Downs, Strand, Heinrichs, & Cerna, 2012; Sharp, Croudace, Goodyer, & Amtmann, 2005).

Child Self-Regulation and Behaviour Questionnaire (CSBQ). This 33-item educator-report questionnaire yields subscales of cognitive self-regulation (e.g. 'likes to work things out for self'), behavioral self-regulation (e.g. 'waits their turn in activities') and emotional self-regulation (e.g. 'gets over being upset quickly'), as well as other social-behavioral outcomes. Each item asks the educator to evaluate the frequency of an identified behavior on a scale from 1 (not true) to 5 (certainly true). CSBQ has shown good internal consistency, as well as structural and concurrent validity, in a diverse sample (Howard & Melhuish, 2017).

Child demographic data. Child demographic data, as reported by the child's primary caregiver, included age at commencement of the prior-to-school year, child's sex, maternal education, identification as Aboriginal or Torres Strait Islander, language spoken at home, and family income.

Procedures

Observation of each pre-school room was conducted initially at the end of 2016, just before the participating children entered the room. This is a consequence of the FEEL study design; however, it was not expected that there would be dramatic changes in process quality between these end-of-year ratings and those experienced by participating children upon entry to the room some two months later. Follow-up environmental ratings were taken at the end of the following year. Baseline environmental quality ratings were used to evaluate concurrent validity and internal consistency. Child assessments were conducted in 2017, at the start and end of the pre-school year. To maximize attention and prevent fatigue, assessments were conducted across two sessions on the same day, each lasting no more than 30 min. The assessments were conducted in the child's ECEC centre and were administered in the same fixed order, as follows: early number concepts and expressive vocabulary (session 1); verbal comprehension and PENS (session 2). The educator who was most familiar with each child (most commonly the ECEC room leader) completed the social-behavioural questionnaires. Child assessments were not conducted on observation days and observations always preceded child assessments.

Plan for analysis

Internal consistency was evaluated through Cronbach alphas within SSTEW subscales. Concurrent validity was evaluated through correlations of SSTEW with ECERS-E and its subscales. For predictive validity, multi-level regression modelling (i.e. hierarchical linear modelling) was conducted, adjusting for the clustering of children within rooms, for each of the child outcome variables measured at the end of the pre-school year. Each model of a child outcome was constructed as a two-level model, with child at the first level and room at the second level. The environmental quality ratings were the only variables that applied at the second level (room). For these analyses, ERS ratings (SSTEW, ECERS-E) were averaged across Time 1 (end of year prior to entry) and Time 2 (end of pre-school year) to best index quality that the children experienced over their pre-school year. These scores, and child development outcome variables, were then standardized into standard deviation units (z-scores) to ensure comparability of multi-level estimates (beta weights) and standard errors. Additional covariates were included in the model at the first (child) level to allow for their associations with children's developmental progress. These were the child's age and sex, scores at baseline (so models evaluated change in scores), identification as Aboriginal or Torres Strait Islander, maternal education, family income, and language spoken at home.

Results

Concurrent validity and internal consistency of the SSTEW scale

Descriptive statistics are presented in Table 2 for centre-level variables and Table 3 for child-level variables. In evaluating concurrent validity of the SSTEW scale, there was a high and significant association between ECERS-E and SSTEW overall scores ($r = .88$). This strong correlation indicates that both scales were capturing aspects of environment process quality (i.e. ECEC centres that are higher in process quality will typically show both higher curricular quality and higher-quality interactions), despite their very different items and foci. Notwithstanding this strong overall correlation, the conclusion that these two scales capture distinct aspects of quality was supported by subscale correlations: subscale correlations within each scale ranged from $r = .44$ to .71 (for ECERS-E) and $r = .43$ to .88 (for SSTEW); subscale correlations between scales ranged from $r = .53$ to .84. Internal consistency within SSTEW subscales ranged from acceptable to excellent (alphas ranging from .76 to .89).

Predictive validity of the SSTEW scale

Multi-level estimates for SSTEW (and, for comparison and replication, ECERS-E) predicting each of the child-level outcomes are presented in Table 4. Covariates that were significant predictors of child-level outcomes were: child assessment results at baseline (all child outcomes), age (for numeracy

Table 2. Descriptive statistics for environmental rating scales (mean of baseline and follow-up).

	n	*Mean*	*SD*	*Range*
ECERS-E	*54*	*3.14*	*0.87*	*1.75–5.13*
1. Literacy	54	3.80	0.97	2.33–6.08
2. Mathematics	54	3.04	1.12	1.33–5.83
3. Science and Envir.	54	3.14	1.00	1.33–5.17
4. Diversity	54	2.60	0.82	1.17–4.50
SSTEW	*54*	*3.89*	*1.06*	*2.13–6.13*
1. Build. Trust, Conf.	54	4.68	1.02	2.33–6.67
2. Social & Emotional	54	4.07	1.30	1.50–6.50
3. Lang. & Comm.	54	4.30	1.18	2.13–6.75
4. Learning & Critical	54	3.00	1.11	1.38–6.00
5. Assess Learn/Lang.	54	3.40	1.26	1.25–6.25

Note. Environmental rating scale scores are averages of baseline and follow-up ratings, as a means to best capture the quality levels experienced over the course of the year.

Table 3. Descriptive statistics for child outcome variables.

	n	*Mean*	*SD*
Number Concepts	600	22.26	4.66
Early Numeracy	597	0.68	0.21
Verbal Comprehension	599	21.16	4.72
Expressive Vocabulary	600	30.93	6.46
CSBQ			
Sociability	608	3.91	0.73
Behavioural Self-Reg	608	3.97	0.96
Cognitive Self-Reg	608	3.82	0.91
Emotional Self-Reg	608	3.86	0.81
SDQ			
Externalising	608	1.42	0.43
Internalising	607	1.32	0.33
Prosocial	603	2.49	0.47
Hyperactivity	605	1.61	0.60
Conduct Problems	605	1.24	0.37
Emotional Problems	602	1.32	0.42
Peer Problems	604	1.32	0.37

Note. Number Concepts = Differential Ability Scales' early number concepts subscale. Early Numeracy = Preschool Early Numeracy Scale. Verbal comprehension = Differential Ability Scales' verbal comprehension subscale. Expressive Vocabulary = Early Years Toolbox's Expressive Vocabulary measure. CSBQ = Child Self-Regulation and Behaviour Questionnaire. SDQ = Strengths & Difficulties Questionnaire.

measures and verbal comprehension), maternal education (for number concepts), and income (for number concepts). After taking into account these socio-demographic and centre-level covariates, SSTEW and ECERS-E overall scores were independent and significant predictors of children's progress in numeracy from baseline to follow-up (typically a 7-month gap). In all models, the β statistic represents the change in standard deviation units in the outcome for a one standard deviation change in predictor.

The multi-level models for interactional quality (the SSTEW scale) were significant for both numeracy measures, DAS: $F(1, 48.96) = 4.29$, $p = .044$; PENS: $F(1, 39.80) = 5.09$, $p = .030$, such that SSTEW overall scores were predicting changes in scores on both measures, DAS: $\beta = .08$, $SE\ \beta = .04$, 95% CI [.00, .17]; PENS: $\beta = .08$, $SE\ \beta = .04$, 95% CI [.01, .15]. These SSTEW results appeared to be driven largely by multi-level estimates for two SSTEW subscales: extending language and communication, DAS: $\beta = .09$, $SE\ \beta = .04$; PENS: $\beta = .07$, $SE\ \beta = .04$; and supporting learning and critical thinking, DAS: $\beta = .10$, $SE\ \beta = .04$; PENS: $\beta = .09$, $SE\ \beta = .03$. Interactional quality, as measured by SSTEW, did not significantly predict verbal comprehension or expressive vocabulary performance after controlling for socio-demographic factors, centre-level covariates, and initial language and literacy scores. Models evaluating the association of SSTEW ratings with changes in children's socio-behavioural outcomes indicated that, after controlling for included covariates, SSTEW scores were not significant predictors of SDQ or CSBQ subscale scores.

The multi-level model for curricular quality, as measured by ECERS-E, was significant only for early numeracy, $F(1, 34.72) = 13.35$, $p = .001$, predicting change in children's early numeracy, $\beta = .12$, $SE\ \beta = .03$, 95% CI [.05, .18]. Each of the ECERS-E subscales predicted change in early numeracy scores, over and above the included covariates (see Table 4). In addition, the science and environment subscale of ECERS-E significantly predicted number concepts, $\beta = .09$, $SE\ \beta = .04$. Like SSTEW, ECERS-E ratings were not significant predictors of language and literacy scores over and above the included covariates. Similarly, the models evaluating the associations of ECERS-E with change in child socio-behavioural development indicated that, after controlling for covariates, ECERS-E scores were also non-significant predictors of SDQ or CSBQ scores.

In summary, overall, children's numeracy development was more consistently related to SSTEW scores than ECERS-E scores, suggesting that the quality of educators' sustained shared thinking and emotional well-being practices (as measured by SSTEW) were slightly more predictive of children's numeracy than were aspects of the educational curriculum (as measured by ECERS-E).

Table 4. Associations of ECEC quality with children's cognitive and socio-behavioural development: Multi-level estimates (and standard errors) from multi-level models.

	Cognitive				Socio-behavioural (SDQ)				
	Number Concepts	Early Numeracy	Verbal Comp.	Express. Vocab	Prosocial	Hyperact.	Conduct Problems	Peer Problems	Emotional Problems
SSTEW									
Overall average	.08 (.04)*	.08 (.04)*	−.06 (.04)	.04 (.03)	−.02 (.05)	−.03 (.05)	−.07 (.06)	.08 (.05)	.04 (.05)
1. Build. Trust, Conf.	.07 (.04)	.06 (.04)	−.02 (.04)	.05 (.03)	.02 (.05)	−.04 (.05)	−.06 (.05)	.07 (.06)	.04 (.06)
2. Social & Emotional	.08 (.04)#	.04 (.04)	−.08 (.04)#	.05 (.03)	−.02 (.05)	−.07 (.05)	−.06 (.05)	.04 (.06)	.01 (.06)
3. Lang. & Comm.	.09 (.04)*	.07 (.04)*	−.05 (.04)	.02 (.03)	.01 (.05)	−.05 (.05)	−.09 (.06)	.05 (.06)	.04 (.06)
4. Learning & Critical	.10 (.04)*	.09 (.03)*	−.04 (.04)	.03 (.03)	−.05 (.05)	−.01 (.05)	−.06 (.06)	.09 (.06)	.05 (.06)
5. Assess Learn/Lang.	.05 (.04)	.09 (.03)*	−.07 (.04)	.03 (.03)	−.06 (.06)	.03 (.05)	−.04 (.05)	.11 (.06)#	.06 (.06)
ECERS-E									
Overall average	.08 (.04)#	.12 (.03)*	−.03 (.04)	.04 (.03)	.00 (.06)	−.02 (.05)	−.07 (.06)	.06 (.06)	.04 (.06)
1. Literacy	.08 (.04)#	.10 (.04)*	−.05 (.04)	.02 (.03)	−.04 (.06)	−.02 (.05)	−.05 (.06)	.09 (.06)	.03 (.06)
2. Mathematics	.07 (.11)	.13 (.03)*	−.02 (.04)	.04 (.03)	.01 (.06)	−.01 (.05)	−.03 (.06)	.06 (.06)	.02 (.06)
3. Science and Envir.	.09 (.04)*	.10 (.03)*	−.01 (.04)	.06 (.03)	−.02 (.06)	−.02 (.05)	−.08 (.06)	.08 (.06)	.08 (.06)
4. Diversity	.04 (.04)	.09 (.03)*	−.04 (.04)	.03 (.03)	.04 (.06)	−.05 (.05)	−.10 (.06)	−.03 (.06)	.00 (.06)

Note. ERS scores (SSTEW, ECERS-E) at Time 1 and Time 2 were averaged to best index the quality experienced by the children in those rooms. ERS scores and outcome variables were then transformed into standard deviation units (z-scores) for comparability across multi-level estimates (beta weights) and standard errors. Overall average = average of all subscale scores. For verbal comprehension all statistics reported are for models that includes random intercepts and/or slopes for ERS rating, given significant improvement in model with this inclusion. Estimates for all CSBQ subscales were similar to SDQ, and non-significant, for all subscales. *$p < .05$. #$p < .08$.

Discussion

The current study serves to introduce and provide initial evaluation of a new measure of interactional quality in ECEC that is distinct in its nature, depth, and scope of focus on high-quality educator-child and child-child interactions. Specifically, this study demonstrated the SSTEW scale's concurrent validity with an established measure of process quality (ECERS-E). Whereas ECERS-E is an established scale focusing on quality of educational curricula, SSTEW focuses on interactional quality and has started to be used only recently in research contexts. This study is also the first to demonstrate that ratings of interactional quality using the SSTEW scale predicted change across a range of child outcome measures over a narrow window of exposure to those settings. Further, children's early numeracy emerged as being particularly sensitive to the process quality of ECEC settings.

Indeed, both SSTEW and ECERS-E showed some prediction of numeracy outcomes. SSTEW was predictive of both numeracy outcomes while ECERS-E was predictive of only one numeracy outcome. The fact that both process quality ratings more consistently predicted early numeracy development is consistent with a UK longitudinal study that showed ECEC quality effects are associated more strongly with numeracy outcomes than either language or social-behavioural outcomes (Sammons et al., 2002, 2003). Further, another UK longitudinal study found that non-verbal outcomes were more sensitive to differences in ECEC experience than verbal outcomes (Barnes & Melhuish, 2016). The results from the current study conform to this pattern, as both SSTEW and ECERS-E were similarly predictive of numeracy progress. It has been suggested (e.g. Melhuish et al., 2012) that this pattern may be a consequence of the ubiquity of language learning experiences at home and formal pre-school services, yet lesser focus in number and breadth of numeracy learning experiences.

Similar results for SSTEW and ECERS-E might be expected given their high level of correlation, indicating that ECEC settings high in one aspect of process quality tend also to be strong in other aspects of process quality. This strong correlation is predictable given that both ERS focus on process aspects of quality and the educator's role in supporting learning, including: recognition of the importance of intentional pedagogy; a child-centered approach; and appropriate concept development and content knowledge that should be applied flexibly and co-constructed with the children. There are, however, also notable differences between the scales, which may account for the additional discrepant results.

The most notable difference between the two scales are the domains of learning and development, and the content/subject knowledge, covered within their subscales. While it may be impossible to separate practice into individual domains of learning – as learning is inevitably inter-related and its domains are reciprocally determined – differences in the focus of the subscales are evident. The ECERS-E requires educators to have content knowledge of, and provide support for: emergent literacy; mathematics; science and exploration; as well as understanding how to support diversity and promote equality with young children. It provides a 'snapshot' of curricula knowledge and practice in these areas. The subject areas covered are typically considered to sit within the cognitive and academic domains of development. The SSTEW scale, in contrast, has a specific focus on quality interactions and its subscales tend to straddle both the cognitive and social-emotional domains of development.

There are two of SSTEW's fourteen items which directly overlap with ECERS-E: one considering literacy (Item 10: Encouraging sustained shared thinking through story telling, sharing books, singing, and rhymes); and another, around mathematics and/or science and exploration (Item 11: Encouraging sustained shared thinking in investigation and exploration). During the SSTEW's development this duplication was acknowledged, however, these two items were nevertheless included in the subscale *Supporting Learning and Critical Thinking* due to their high potential for leading to SST. That is, engaging with the scientific process is highly likely to support thinking and metacognitive development, as children explore, hypothesize, and evaluate. Engaging with mathematics – taking a more contemporary view of mathematics as problem solving and supporting creativity, rather than just

numbers and counting – is likely to also enhance children's thinking. Literacy activities such as dialogic reading, can support children in participating, anticipating, predicting, evaluating, taking another's perspective, and recognizing different views. The remaining items in the subscale *Supporting Learning and Critical Thinking,* as well as the other subscales in SSTEW, expand this particular and unique focus on supporting interaction. That is, the pedagogy and practice described within the SSTEW scale can occur at any time during the day and whilst engaged in any activity. The SSTEW scale thus considers the child and their learning as a whole, within and outside of specific domains of learning, resources, and experiences.

A final difference between the two scales relates to their structure. In ECERS-E, all intentional pedagogy is found in indicators at the higher end of each item. As such, given the tradition amongst many ERS of considering higher-quality pedagogies and practices only if lower-level indicators are also met, interactional quality only influences an ECERS-E quality rating amongst higher-quality centres (and are disregarded even if they are occurring within comparatively lower-quality centres). With the SSTEW scale's focus strongly on interaction, it has entire subscales dedicated to intentional pedagogy around interactions that may lead to SST. This includes the subscale *Supporting Learning and Critical Thinking*, which requires intentionality for scores of 2 and above. This difference may enhance the SSTEW scale's particular sensitivity to concepts and domains that are reliant on sustained and high quality interactions, perhaps explaining why it associates with children's performance on two measures of numeracy.

It is noteworthy that, while some associations were approaching significance, neither scale predicted change in any of the language or social-behavioural outcomes. This contrasts previous research, in which scores on ECERS-E were predictive of literacy, numeracy, and non-verbal reasoning, after allowing for children's prior attainment, family characteristics, parental care, and home learning environment (Sylva et al., 2006). One possible explanation for this discrepancy, and limitation to this study, is the time available for ECEC to influence child outcomes across these studies. Whereas in Sylva et al.'s (2006) study the pre-test was at age 3 and the post-test at age 4–5 (around 18–24 months of ECEC exposure), in the current study the duration of ECEC exposure from pre- to post-test was an average of 7 months (a consequence of the design of the FEEL study). Hence, the time available for the ECEC environment to influence child outcomes and the age range of the children differed in the two studies. This suggests that the results of the current study may indeed under-estimate (yet still demonstrate their ability to detect) the influence of differences in ECEC quality, as measured by SSTEW and ECERS-E, upon child development. Another factor potentially influencing the lack of prediction of changes in social-behavioural development is the adult-report nature of the tools adopted. That is, raters normally anchor their judgement of a child to children of a similar age, which renders scores as relative rankings within an age group and context. These rankings typically remain stable, even when real-world change has occurred (Howard & Melhuish, 2017). Such an effect is particularly relevant over shorter periods of time. Further research using objective measures of socio-behavioural outcomes are needed to circumvent this issue.

Finally, building on the opportunities provided by the SSTEW scale to inform practice, an interesting area of further research would be evaluating the impact of use of the scale by ECEC practitioners. While this would likely reduce the objectivity inherent in independent observation, highly rigorous standards of reliability are less essential to the aim of practice change. The SSTEW scale provides a unique opportunity in this regard.

Conclusion

In summary, this study introduces and demonstrates concurrent validity of the SSTEW scale through comparison with an established measure of ECEC quality and provides initial evidence of the SSTEW scale's ability to capture environmental differences that are likely to be important for, at a minimum, children's numeracy development. Indeed, that interactional aspects of quality (such as SST) were more consistently associated with children's numeracy development than even the quality of

curricula around literacy and mathematics, suggests the unique and independent importance of these previously overlooked interactional aspects of process quality. Further, the stringent nature of the longitudinal analyses in the current study (i.e. brief duration of exposure to ECEC) may well underestimate SSTEW scale's potential to capture environmental aspects relevant for fostering children's development more broadly. The additional potential of the SSTEW scale for informing practice and in-service learning is an area of opportunity that requires further investigation. In the context of regulatory authorities continuing to increase their focus on improving quality in ECEC, the current study suggests that aspects of interactional quality might be particularly indicative of quality, where quality is referenced to its influence of children's development and outcomes.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work derives from data collected in the Fostering Effective Early Learning (FEEL) study, funded by NSW Department of Education and Training [grant number DECEAR-15-35].

Notes on contributors

Steven J. Howard's research investigates early child development, as well as the educational implications of these changing capacities. This includes research to identify effective approaches to supporting children's development that are low-cost and flexible, as well as identifying ways to monitor children's developmental progress that are accessible, acceptable, and actionable by those who spend their time working with young children.

Iram Siraj's research focuses on understanding the relationship between quality, pedagogy, and early assessment on young children's developmental outcomes. She works in transdisciplinary teams on longitudinal studies to assess the long-term impact of the quality of pre-school and home environments on children from differing backgrounds.

Edward C. Melhuish is an international expert in early child development and the impact of home learning and pre-school experiences on children's later outcomes. He also has vast experience linking child development with social policy, and is a consultant to the OECD, WHO, and EU.

Denise Kingston teaches across undergraduate and graduate courses within the Department of Education and Social Work. She is a qualified teacher and Educational Psychologist and her current research interests focus on supporting and improving the quality of early childhood education and care through evidence-based professional development.

Cathrine Neilsen-Hewett is Director of the Early Years at University of Wollongong, and has particular expertise in development in early childhood, early childhood pedagogy and practice, and staff training and childhood socialization.

Marc de Rosnay is Academic Director at Early Start, University of Wollongong, and has specific interest and expertise in social, emotional, and communicative development in the early years.

Elisabeth Duursma is interested in parenting and, in particular, how fathers engage with their young children and the impact this has on children's early language and literacy development. Her research focuses specifically on vulnerable populations such as low-SES and linguistically and ethnically diverse families.

Betty Luu's research interest is in understanding how early environments can best support children's development. With a background in developmental psychology, she has contributed to child-related research areas, including: children's psychosocial and cognitive development, the impact of early childhood teaching practices on children's outcomes, and open adoption for children and young people in out-of-home care.

ORCID

Steven Howard http://orcid.org/0000-0002-1258-3210
Iram Siraj http://orcid.org/0000-0001-6752-881X
Marc de Rosnay http://orcid.org/0000-0001-8333-9541
Betty Luu http://orcid.org/0000-0003-4461-500X

References

Ainsworth, M. D. S. (1979). Infant-mother attachment. *American Psychologist*, *34*, 932–937.

Australian Institute of Health and Welfare (AIHW). (2012). *A picture of Australia's children* 2012. Retrieved from http://www.aihw.gov.au

Barnes, J., & Melhuish, E. C. (2016). Amount and timing of group-based childcare from birth and cognitive development at 51 months: A UK study [online first]. *International Journal of Behavioral Development*, *41*, 361–370.

Bodrova, E., & Leong, D. J. (2007). *Tools of the mind: The Vygotskian approach to early childhood education* (2nd ed.). New York: Pearson.

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

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum Publishing Co.

Department for Education and Employment/Qualifications and Curriculum Authority (DfEE/QCA). (2000). *Curriculum guidance for the Foundation Stage*. Retrieved from http://webarchive.nationalarchives.gov.uk/20040117041426/http://www.dfes.gov.uk/foundationstage/download.shtml

Department for Education (DfE). (2012). *Statutory framework for the early years foundation stage*. Manchester, UK: Department for Education.

Department of Education, Employment and Workplace Relations. (DEEWR). (2009). *The Early Years Learning Framework for Australia: Belonging, being, becoming*. Retrieved from https://www.education.gov.au/early-years-learning-framework-0

Donabedian, A. (1980). *The definition of quality and approaches to its assessment*. Ann Arbor: Health Administration Press.

Downer, J. T., Lopez, M. L., Grimm, K. J., Hamagami, A., Pianta, R. C., & Howes, C. (2012). Observations of teacher–child interactions in classrooms serving Latinos and dual language learners: Applicability of the classroom assessment scoring system in diverse settings. *Early Childhood Research Quarterly*, *27*, 21–32.

Downs, A., Strand, P. S., Heinrichs, N., & Cerna, S. (2012). Use of the teacher version of the strengths and difficulties questionnaire with German and American preschoolers. *Early Education and Development*, *23*(4), 493–516.

Elliott, C. D. (2007). *Differential ability scales-II [measurement instrument]*. San Antonio, TX: Pearson.

Fukkink, R., & Lont, A. (2007). Does training matter? A meta-analysis and review of caregiver training studies. *Early Childhood Research Quarterly*, *22*, 294–311.

Goodman, R., Meltzer, H., & Bailey, V. (1998). The strengths and difficulties questionnaire: A pilot study on the validity of the self-report version. *European Child & Adolescent Psychiatry*, *7*(3), 125–30.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, *72*(2), 625–638.

Howard, S. J., & Melhuish, E. C. (2017). An Early Years Toolbox (EYT) for assessing early executive function, language, self-regulation, and social development: Validity, reliability, and preliminary norms. Manuscript accepted for publication in *Journal of Psychoeducational Assessment*.

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

Katz, L. (2008). *Current Perspectives on the Early Childhood Curriculum*, Paper presented at OPEN EYE Conference, London.

Kingston, D., & Siraj, I. (2017). Supporting the implementation of the foundation phase through effective professional development. *Wales Journal of Education. Special Edition: Professional Learning*, *19*(1), 39–68.

LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., & Barbarin, O. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly*, *22*, 3–17.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., … Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*(3), 732–749.

Mayer, D., & Beckh, K. (2016). Examining the validity of ECERS-R: Results from the German national study of child care in early childhood. *Early Childhood Research Quarterly*, *36*, 415–426.

Melhuish, E., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K., … Leseman, P. (2015). A review of research on the effects of Early Childhood Education and Care (ECEC) upon child development. Retrieved from http://ecec-care.org/resources/publications/

Melhuish, E. C., Howard, S. J., Siraj, I., Neilsen-Hewett, C., Kingston, D., de Rosnay, M., … Luu, B. (2016). Fostering effective early learning (FEEL) through a professional development program for early childhood educators to improve professional practice and child outcomes in the year before formal schooling: Study protocol for a cluster randomized controlled trial. *Trials*, *17*(1), 118.

Melhuish, E., Quinn, L., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2012). Preschool affects longer term literacy and numeracy: Results from a general population longitudinal study in Northern Ireland. *School Effectiveness and School Improvement*, *24*(2), 234–250.

Moreno, A. J., Shwayder, I., & Friedman, I. D. (2017). The function of executive function: Everyday manifestations of regulated thinking in preschool settings. *Early Childhood Education Journal*, *45*, 143–153.

Organisation for Economic Co-operation and Development (OECD). (2012). *Starting Strong III: A quality toolbox for early childhood education and care*. Retrieved from http://www.oecd.org/edu/school/startingstrongiiiaqualitytoolboxforearlychildhoodeducationandcare

Penn, H. (2011). *Quality in early childhood services. An international perspective*. Maidenhead: OUP.

Pianta, R. C. (ed.). (2012). *Handbook of early childhood education*. London: Guildford Press.

Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, R. M., & Early, D. (2005). Features of pre-kindergarten programs, classrooms and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*, 144–159.

Plantenga, J. (2011). Local providers and loyal parents: Competition and consumer choice in the Dutch childcare market. In E. Lloyd & H. Penn (Eds.), *Childcare markets: Can they deliver an equitable service?* (pp. 63–78). Bristol, UK: The Policy Press.

Purpura, D. J., & Lonigan, C. J. (2015). Early numeracy assessment: The development of the preschool numeracy scales. *Early Education & Development*, *26*(2), 286–313.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68–78.

Sammons, P., Sylva, K., Melhuish, E. C., Siraj-Blatchford, I., Taggart, B., & Elliot, K. (2002). *The effective provision of Pre-school education project (EPPE), technical paper 8a: Measuring the impact of pre-school on children's cognitive progress over the pre-school period*. London: DfES/Institute of Education, University of London.

Sammons, P., Sylva, K., Melhuish, E. C., Siraj-Blatchford, I., Taggart, B., & Elliot, K. (2003). *The effective provision of Pre-school education project (EPPE), technical paper 8b: Measuring the impact of pre-school on children's socio-behavioural development over the pre-school period*. London: DfES/Institute of Education, University of London.

Sharp, C., Croudace, T. J., Goodyer, I. M., & Amtmann, D. (2005). The strengths and difficulties questionnaire: Predictive validity of parent and teacher ratings for help-seeking behaviour over one year. *Educational & Child Psychology*, *22* (3), 28–44.

Siraj, I., & Kingston, D. (2015). *An independent review of the Scottish Early Learning and Childcare (ELC) workforce and Out of School Care (OSC) workforce*. Retrieved from http://www.gov.scot/Publications/2015/06/5902/downloads

Siraj, I., Kingston, D., & Melhuish, E. (2015). *Assessing quality in early childhood education and care. Sustained shared thinking and emotional wellbeing (SSTEW)sScale for 2–5 year olds provision*. London: UCL and IOE Press.

Siraj, I., Kingston, K., Neilsen-Hewett, C., Howard, S. J., Melhuish, E., de Rosnay, M., … Luu, B. (2017). *A review of the current international evidence considering quality in early childhood education and care programmes – in delivery, pedagogy and child outcomes*. Sydney: NSW Department of Education.

Siraj-Blatchford, I., Sylva, K., Muttock, S., Gilden, R., & Bell, D. (2002). *Researching effective pedagogy in the early years (REPEY): DfES research report 356*. London: DfES, HMSO.

Siraj-Blatchford, I., Sylva, K., Taggart, B., Sammons, P., Melhuish, E., & Elliot, K. (2003). *The Effective Provision of Pre-School Education [EPPE] Project A Longitudinal Study funded by the DfES (1997–2003) Technical Paper 10 Intensive Case Studies of Practice across the Foundation Stage*. Retrieved from http://www.ioe.ac.uk/EPPETechnicalPaper_10_2003.pdf

Siraj-Blatchford, I., & Wong, Y. (1999). Defining and evaluating 'quality' early childhood education in an international context: Dilemmas and possibilities. *Early Years*, *20*(1), 7–18.

Slot, P. L., Leseman, P., Verhagen, J., & Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early education and care settings. *Early Childhood Research Quarterly*, *33*, 64–76.

Sylva, K., Melhuish, E., Sammons, P., Siraj, I., Taggart, B., Smees, R., … Welcomme, W. (2014). *Students' Educational and Developmental Outcomes at age 16. Effective Pre-school, Primary and Secondary Education EPPSE 3–16 project*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/351496/RR354Students educational_and_developmental_outcomes_at_age_16.pdf

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2004). *The effective provision of pre- school education (EPPE) project: Final report*. Nottingham: DfES Publications.

Sylva, K., Siraj-Blatchford, I., & Taggart, B. (2010). *Early childhood environment rating scale-extension (ECERS-E)*. Stoke-on-Trent: Trentham Books.

Sylva, K., Siraj-Blatchford, I., Taggart, B., Melhuish, E. C., & Sammons, P. (2006). Capturing quality in early childhood through environmental rating scales. *Early Childhood Research Quarterly*, *21*, 76–92.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

West, A., Roberts, J., & Noden, P. (2010). Funding early years education and care: Can a mixed economy of providers deliver universal high quality provision? *British Journal of Educational Studies*, *58*, 155–179.