Measures of Social and Emotional Skills for Children and Young People: A Systematic Review
Neil Humphrey, Afroditi Kalambouka, Michael Wigelsworth, Ann Lendrum, Jessica Deighton and Miranda Wolpert
Educational and Psychological Measurement 2011 71: 617
DOI: 10.1177/0013164410382896

The online version of this article can be found at:
http://epm.sagepub.com/content/71/4/617

Published by:
http://www.sagepublications.com

Additional services and information for Educational and Psychological Measurement can be found at:

Email Alerts: http://epm.sagepub.com/cgi/alerts
Subscriptions: http://epm.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Citations: http://epm.sagepub.com/content/71/4/617.refs.html

>> Version of Record - Jun 29, 2011

What is This?
Measures of Social and Emotional Skills for Children and Young People: A Systematic Review

Neil Humphrey1, Afroditi Kalambouka1, Michael Wigelsworth1, Ann Lendrum1, Jessica Deighton2, and Miranda Wolpert2

Abstract
This study presents the findings of a systematic review of measures of social and emotional skills for children and young people. The growing attention to this area in recent years has resulted in the development of a large number of measures to aid in the assessment of children and young people. These measures vary on a number of variables relating to implementation characteristics and psychometric properties. The methodology of the review followed the general principles of systematic reviewing, such as systematic search of databases, the adoption of predetermined set of inclusion and exclusion criteria, and a multistage filtering process. The review process resulted in the retention of 12 measures, which are presented and discussed in relation to key issues in this area, including difficulties with the underlying theory and frameworks for social and emotional skills, inconsistent terminology, the scope and distinctiveness of available measures, and more practical issues such as the type of respondent, location, and purpose of measurement.

Keywords
social and emotional skills, measurement, children

1University of Manchester, Manchester, UK
2The Anna Freud Centre, London, UK

Corresponding Author:
Neil Humphrey, School of Education, University of Manchester, Oxford Road, Manchester, M13 9PL, UK
Email: neil.humphrey@manchester.ac.uk
Introduction

This article reports on the findings of a systematic review of measures of social and emotional skills for children and young people. Its intended contributions are (a) to provide an overview of the state of the field in this important area of psychological measurement, (b) to highlight key issues pertaining to the definition and measurement of social and emotional skills in children and young people, and (c) to outline and discuss the implementation characteristics and psychometric properties of the 12 measures of social and emotional skills that passed our systematic review screening process and are firmly established in the academic literature in this area.

The article is the first (to our best knowledge) truly systematic and rigorous review of measures of social and emotional skills for children and young people, and as such it may provide a unique contribution to knowledge in this field. It is intended to be of use to researchers and practitioners with an interest in children’s social and emotional skills around the world. To this end, we have included information pertaining to the availability of norms and translated versions of the measures identified during the review process. This is of particular importance given the increasing interest in children’s social and emotional skills in countries other than the United States and the United Kingdom (Fundacion Marcelino Botin, 2008).

Defining Social and Emotional Skills

In recent years, educational policies in the United Kingdom, the United States, and elsewhere have reflected a growing interest and investment in the development of social and emotional skills as a means to promote children’s well-being, adjustment, and academic achievement. The issue of the mixed and inconclusive evidence on the effectiveness of these approaches is a matter of ongoing debate that has generated a great deal of academic and professional interest and will continue to do so for the foreseeable future (Humphrey, Curran, Morris, Farrell, & Woods, 2007; Humphrey et al., 2008; Zeidner, Roberts, & Matthews, 2002). However, more fundamental issues—such as inconsistencies in the definition, measurement, and utility of the concept of social and emotional skills—also need to be acknowledged and explored (Matthews, Zeidner, & Roberts, 2004; Wigelsworth, Humphrey, Kalambouka, & Lendrum, 2010). There seems to be little common consensus as regards to what is meant by social and emotional skills, and how they are best measured. As Humphrey et al. (2007) have argued, clear definition is a basic scientific requirement—yet one is still lacking in this area.

“Social and emotional skills” is referred to in this article as it is the currently preferred term in English government policy documentation (e.g., Department for Children, Schools and Families [DCSF], 2007a, 2007b). We use it to incorporate other terms, such as “social and emotional intelligence,” “emotional literacy,” and “social and emotional competence,” among others. Although some authors (e.g., Weare & Gray, 2003) have argued for the differentiation of these terms, reviews of the literature in this area (e.g., Humphrey et al., 2007; Wigelsworth et al., 2010) have led us to the
conclusion that they do not describe qualitatively different concepts. In defining social and emotional skills, we refer readers to the framework of social–emotional competence proposed by Denham (2005a). Drawing on the work of Rose-Krasnor (1997) and Payton et al. (2000), Denham proposes a delineation of relational/prosocial skills and emotional competence skills as given in Table 1.

This model benefits from reflecting the domains outlined in major theoretical models (e.g., Bar-On, 1997; Goleman, 1995; Salovey & Mayer, 1990) in this area, without suffering from the nebulous, imprecise nature of the more wide-ranging definitions (e.g., Petrides & Furnham, 2001) that have attracted criticism from skeptics (see Matthews, Zeidner, et al., 2004).

A number of key issues in reviewing measures of social and emotional skills are suggested in the literature. In brief, these include (a) the distinction between capturing typical and maximal behavior (Wilhelm, 2005), (b) the extent to which measures of social and emotional skills provide information that is distinct or unique among existing constructs (e.g., personality), (c) the scope and specificity of measures (e.g., single, unidimensional vs. complex, multidimensional), and (d) who provides information (e.g., children, teacher, parent, peers; see Wigelsworth et al., 2010).

**Rationale and Context for the Current Study**

The accurate assessment of social and emotional skills in children and young people has crucial implications for public health because of the associations with mental health, academic performance, and other key outcomes (Denham, Wyatt, Bassett, Echeverria, & Knox, 2009). In late 2008, the authors were commissioned by the DCSF to conduct a systematic review of measures of social and emotional skills for children and young people (Humphrey et al., 2009) as companion to a systematic review of measures of child mental health in children (Wolpert et al., 2009). The aim of the review was to identify a range of measures suitable for assessing social and emotional skills in children and young people for possible future inclusion in a population survey relating to progress toward the “2020 vision” in the Children’s Plan (DCSF, 2007b). This policy document

**Table 1. Denham’s (2005a) Framework for Social–Emotional Competence**

<table>
<thead>
<tr>
<th>Emotional competence skills</th>
<th>Self-awareness</th>
<th>Understanding self emotions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Self-management</td>
<td>Emotional and behavioral regulation</td>
</tr>
<tr>
<td></td>
<td>Social awareness</td>
<td>Understanding emotions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Empathy/sympathy</td>
</tr>
<tr>
<td>Relational/prosocial skills</td>
<td>Social problem solving</td>
<td>Cooperation</td>
</tr>
<tr>
<td></td>
<td>Relationship skills</td>
<td>Listening skills</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Turn-taking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Seeking help</td>
</tr>
</tbody>
</table>

by guest on January 1, 2012epm.sagepub.com

Downloaded from epm.sagepub.com by guest January 1, 2012
sets out an ambitious set of goals for children in England to be achieved by the year 2020, including a focus on improving social and emotional skills. Beyond this, we felt that a review of this kind was extremely timely. Although there are a handful of discussion papers (e.g., Wigelsworth et al., 2010) and reviews (e.g., Denham, 2005a; Stewart-Brown & Edmunds, 2003) in the literature, they are all selective and none have applied a rigorous and systematic approach. Furthermore, given the emphasis on monitoring and evaluation in currently popular initiatives such as the Social and Emotional Aspects of Learning (SEAL) program, a review of this kind should also be useful to practitioners. Although the DCSF have provided a guidance document on this topic (see DCSF, 2007a), it contains no substantive references to issues such as reliability and validity, and only one of the recommended measures has undergone any kind of appropriate validation process. Finally, given the high level of academic interest and debate around the concept of emotional intelligence, a review of this kind is crucial: “Determining whether EI [emotional intelligence] is a measurable quality is central to its development as a scientific construct” (Matthews, Roberts, & Zeidner, 2004, p. 182).

**Review Method**

The present review followed the general paradigm of a parallel systematic review of mental health measures for children and young people (Wolpert et al., 2009). The review followed a systematic process across six consecutive stages, summarized in the flow diagram given in Figure 1 and described below.

**Stages 1: Identification of Measures**

*Setting the Parameters of the Review and Search of Key Databases.* Following the parallel child mental health review (Wolpert et al., 2009), “outcome measures” were defined as any assessment or evaluation tools (e.g., questionnaires, checklists, or scales) that aim to measure the social and emotional skills of children and young people. The terms of the “measures of social and emotional skills” review were divided into three categories (see below) and their related terms (see Table 2):

- a. terms relating to social and emotional skills,
- b. terms relating to measurement, and
- c. terms relating to the population under study (children and young people)

Initial database searches focused on the main psychology, sociology, education, and medical databases. However, searches in the main medical databases produced limited relevant hits and a lot of irrelevant hits (e.g., those that were within the child mental health review remit), so it was decided that the medical databases would be dropped in favor of more extensive searches in the following educational, social and psychology databases: PsychInfo, Education Resources Information Centre (ERIC), British Education Index (BEI), Australian Education Index (AEI), and Applied Social Sciences Indexes by guest on January 1, 2012epm.sagepub.com Downloaded from epm.sagepub.com by guest on January 1, 2012
and Abstracts (ASSIA). All related terms for each of the three categories of interest (see Table 1) were then entered in the thesaurus facility within each database in order to ensure that the key words under which articles were catalogued would be identified through using the appropriate terms within each specific database. The thesaurus searches produced additional relevant key words and/or headings, which were then incorporated into subsequent searches. For example, the terms “emotional adjustment” and “emotional development” were identified within the ERIC thesaurus facility search and were later used to capture relevant articles. Any additional terms produced through the thesaurus/heading search that were too broad and produced too many irrelevant hits were later
discarded. This stage generated a list of final search terms, which were then used to run searches across each of the databases (see Table 3).

A total of 2,870 hits were transferred and saved into bibliographic applications (EndNote 9.0.1 for the ASSIA and PsychInfo databases; Reference Manager v11 for the ERIC, BEI, and AEI databases). Basic filtering was then applied through reading the title (and/or abstract in cases where the title was not sufficient) to discard the obviously irrelevant hits (i.e., papers not related to the measurement of children’s social and emotional skills). Basic filtering resulted in a total of 542 articles. The details of these articles (title, abstract, and publication details) were transferred into Excel spreadsheets, where further filtering took place involving the application of the following exclusion criteria: (a) no social and/or emotional skills measure mentioned in the abstract, (b) the measure mentioned was too narrow to provide a broad assessment of social and emotional skills (e.g., focused exclusively on just one specific aspect of social and emotional skills such as empathy), (c) the article referred to a measure that was not used with child or adolescent populations, (d) the article was not in English, and/or (e) the article was a duplicate of a previous hit within the database. All remaining hits after this filtering procedure were organized according to the measures they referred to, resulting to a total of 137 measures.

**Stage 2: Secondary Searches and Collective Knowledge Tranche**

The list of 137 measures was complemented by a number of measures from a “collective knowledge tranche.” These were measures that were known to members of the research team through previous research. Additional measures were identified through reviewing papers and compendiums (e.g., Denham, 2005a) and/or bibliographic references. The authors also contacted a number of relevant groups and organizations (e.g., Collaborative for Academic, Social and Emotional Learning; Antidote) via email, but

<table>
<thead>
<tr>
<th>Category</th>
<th>Related Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Social and emotional skills</td>
<td>Social/emotional&lt;br&gt;Social, emotional, socioemotional, socio-emotional, effect, affective, interpersonal&lt;br&gt;Skills&lt;br&gt;Skills, intelligence, competence, literacy, learning, awareness, ability, regulation</td>
</tr>
<tr>
<td>2. Measurement</td>
<td>Measure, measurement, questionnaire, survey, checklist, check list, tool, rating scale, scale, inventory, assessment, evaluation, screening, instrument, test</td>
</tr>
<tr>
<td>3. Population of interest</td>
<td>Children, adolescents, pupils, students</td>
</tr>
</tbody>
</table>

**Table 2. Categories of Interests and Related Terms for the Systematic Review**
Table 3. Final Search Terms for the Systematic Review

<table>
<thead>
<tr>
<th>Database</th>
<th>Search Terms for Subject Heading</th>
<th>Initial Hits</th>
<th>Hits After Basic Filtering</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASSIA</td>
<td>DE = (emotion* or social) and DE = (skills or intelligence or competence or learning or awareness or regulation or behavior or interaction) and DE = (measure* or questionnaire* or assessment or test* or survey* or scale* or evaluation or instrument*) and DE = (child* or (young people) or adolescent* or pupil* or student*)</td>
<td>172</td>
<td>80</td>
</tr>
<tr>
<td>ERIC, BEI, and AEI</td>
<td>(emotional skills OR socioemotional skills OR social intelligence OR social competence OR emotional literacy OR emotional development OR social development OR interpersonal competence OR social cognition OR affective behavior OR social behavior OR social skills OR emotional intelligence OR emotional competence OR socioemotional competence OR social literacy OR socioemotional literacy OR social learning OR emotional learning OR social awareness OR emotional awareness OR emotional adjustment OR emotional response OR social ability OR emotional ability OR socioemotional ability) DE AND (Measurement OR Questionnaire OR Rating-Scale OR surveys OR checklist OR psychometric OR assessment OR screening) DE AND (children OR adolescent OR pupil OR student) DE</td>
<td>1820</td>
<td>390</td>
</tr>
<tr>
<td>PsychInfo</td>
<td>(emotion* or social* or socio*) AND (measure* or question* or survey or check* or tool or rating or scale* or inventor* or instrument* or test* or screening or repository or assessment* or evaluation*) AND (child* or adolescent* or youth* or young* or pupil* or student*)</td>
<td>788</td>
<td>72</td>
</tr>
</tbody>
</table>

Note. ASSIA = Applied Social Sciences Indexes and Abstracts; ERIC = Education Resources Information Centre; BEI = British Education Index; AEI = Australian Education Index. “Hits” refers to the number of articles identified in each database.

This did not result in the procurement of any measures not already identified elsewhere. A total of 52 measures identified in this stage were added to the existing list for further sorting, increasing the total number of measures to 189.
Stage 3: Sifting of Measures

Filtering of Measures According to Inclusion and Exclusion Criteria. This stage involved reading of the relevant abstract and/or article and applying a list of inclusion/exclusion criteria. A measure was included if it met all of the inclusion criteria and excluded if it met one or more of the exclusion criteria:

Inclusion criteria. To include a measure of social and emotional skills for children and adolescents, the measure should

1. seek to provide measurement of generic social and emotional skills in children and young people (up to 18);
2. be either multidimensional or unidimensional;
3. refer to a measure that can be completed by child or parents/carers with the possible addition of professionals such as teachers (but not by teachers only);
4. have been validated in a child or adolescent context, even if not originally developed for this purpose;
5. be available in English language; and
6. be able to be used with a reasonably wide age range (e.g., not just for preschoolers).

Exclusion criteria. A measure would be excluded if it:

1. is not available in English;
2. does not measure social and/or emotional skills;
3. does not cover a broad range of social and/or emotional skills (e.g., focused exclusively on just one specific aspect of social and emotional skills such as empathy—see Table 1);
4. is not used with children and/or adolescents;
5. is based on professional report only (e.g., teacher, educational psychologist, or other practitioner);
6. takes more than 30 minutes to complete;
7. provides open-ended responses that have to be manually coded;
8. the age range is too narrow (e.g., preschool version of the measure only);
9. has not been used with a variety of populations (e.g., only used with specialist groups such as children with autism).

Following application of the above criteria, the list of 187 measures was reduced to 52 measures (for a list of the 52 measures see Humphrey et al., 2009). Measures for which information from the abstract/article did not provide adequate evidence to apply the above inclusion/exclusion criteria at this stage were still retained until further evidence was available to enable a decision on retention.
Stage 4: Collecting More Information on Retained Measures

Secondary Searches and Requests to Measure Developers for More Information. The above list of the 52 potentially suitable measures was organized in an Excel database, where information was entered on the following variables: a brief description of the measure, different versions available (e.g., child, parent, teacher), number of items and completion time, scales and subscales are included in the measure, response format (e.g., 4-point Likert-type scale), information on reliability (internal consistency, test–retest) and validity (e.g., construct validity), sample items, cost, and key references.

To provide as complete information as possible on these variables, several procedures were followed: (a) an attempt was made to obtain as many articles as possible on each measure and (b) measure developers were contacted through email/telephone to fill in gaps in our database and to request manuals and articles. As we found more detailed information on each measure it became obvious that some did not meet the inclusion/exclusion criteria set out in the previous stage (for instance, if we found that a given measure could only be completed by teachers). By the end of this stage a total of 23 measures were retained.

Stage 5: Filtering According to Academic Presence: Number of Articles Using Measure

This stage involved a further filtering process in which only measures that had been used in four or more articles in peer-reviewed academic journals were retained. This was done in an effort to retain only measures with an established and sustained presence in the academic literature, whilst also ensuring a workable number of measures for detailed examination in the final stage. Four or more articles was considered to be a useful benchmark since it would typically signify that the measure has gone beyond basic development and validation and is being used by the academic community for program evaluation, theory development, and so on. This process resulted in 12 measures being retained.

Stage 6: Detailed Examination of Final Measures

Presentation and Description: Implementation Characteristics and Psychometric Properties of Final 12 Measures. The final 12 measures were reviewed in depth in relation to their implementation characteristics and psychometric properties. These are presented in Tables 4 and 5, respectively.

Discussion

The primary purpose of our systematic review was to identify a range of measures suitable for assessing social and emotional skills in children and young people for possible future inclusion in a population survey relating to progress toward the “2020
Table 4. Implementation Characteristics of the Final 12 Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Acronym</th>
<th>Maximal or Typical</th>
<th>Scope</th>
<th>Age (Years)</th>
<th>Versions</th>
<th>Length</th>
<th>Completion Time (Minutes)</th>
<th>Scales and Subscales</th>
<th>Response Scales</th>
<th>Other Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bar-On Emotional Quotient Inventory: Youth Version (Bar-On &amp; Parker, 2008)</td>
<td>EQI:YV, EQI:YV(S)</td>
<td>Typical</td>
<td>Social—emotional</td>
<td>7-18</td>
<td>Child</td>
<td>60, 30(S)</td>
<td>20-25, 10-15 (S)</td>
<td>Total emotional intelligence, interpersonal, intrapersonal, adaptability, stress management, general mood, positive impression, inconsistency index</td>
<td>Likert 1-4</td>
<td>Pedi (South Africa)</td>
</tr>
<tr>
<td>Child Assertive Behaviour Scale (Michelson &amp; Wood, 1982)</td>
<td>CABS, CABS-SR</td>
<td>Maximal (child), typical (other)</td>
<td>Social</td>
<td>8-12</td>
<td>Child, parent, teacher</td>
<td>27</td>
<td>10-15</td>
<td>Total assertiveness, passivity, and aggressiveness</td>
<td>Multiple choice 1-5</td>
<td>Spanish, Dutch</td>
</tr>
<tr>
<td>Social Competence and Behavior Evaluation Scale (LaFreniere &amp; Dumas, 1996)</td>
<td>SCBE</td>
<td>Typical</td>
<td>Social</td>
<td>2.5-6.5</td>
<td>Teacher, parent</td>
<td>80, 30(S)</td>
<td>15</td>
<td>Depressive—joyful, anxious—secure, angry—tolerant, isolated—integrated, aggressive—calm, egotistical—prosocial, oppositional—cooperative, dependent—autonomous and 4 summary scales: social competence, externalizing problems, internalizing problems, and general adaptation</td>
<td>Likert 1-6</td>
<td>French-Canadian, Spanish, Japanese, Russian, Chinese, Portuguese, Italian, Austrian-German</td>
</tr>
<tr>
<td>Preschool and Kindergarten Behaviour Scales-2 (Merrell, 1996)</td>
<td>PKBS-2</td>
<td>Typical</td>
<td>Social</td>
<td>3-6</td>
<td>Teacher, parent</td>
<td>76</td>
<td>12</td>
<td>Social cooperation, social interaction, social independence, externalizing problems and internalizing problems</td>
<td>Likert 1-4</td>
<td>Spanish</td>
</tr>
<tr>
<td>Prosocial Tendencies Measure–Revised (Carlo, Hausmann, Christiansen, &amp; Randall, 2003)</td>
<td>PTM-R</td>
<td>Typical</td>
<td>Social—emotional</td>
<td>11-18</td>
<td>Child</td>
<td>25</td>
<td>30</td>
<td>Public, anonymous, dire, emotional, compliant and altruistic prosocial behavior</td>
<td>Likert 1-5</td>
<td></td>
</tr>
<tr>
<td>Measure</td>
<td>Acronym</td>
<td>Maximal or Typical</td>
<td>Scope</td>
<td>Age (Years)</td>
<td>Versions</td>
<td>Length</td>
<td>Completion Time (Minutes)</td>
<td>Scales and Subscales</td>
<td>Completion Time (Minutes)</td>
<td>Response Scales</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>---------</td>
<td>--------------------</td>
<td>----------------------</td>
<td>-------------</td>
<td>----------</td>
<td>--------</td>
<td>----------------------------</td>
<td>----------------------</td>
<td>--------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Child Rating Scale</td>
<td>CRS</td>
<td>Typical</td>
<td>Social</td>
<td>5-13</td>
<td>Child, parent, teacher</td>
<td>24, 38 (teacher), 18 (parent)</td>
<td>10-20</td>
<td>Rule compliance/acting out, anxiety/withdrawal, interpersonal social skills, self-confidence (child); acting out, shy-anxious, learning, frustration tolerance, assertive social skills, task orientation, peer sociability (teacher); Acting out, frustration tolerance, shy-anxious and peer sociability (parent)</td>
<td>Likert 1-3</td>
<td></td>
</tr>
<tr>
<td>Social Skills Improvement System</td>
<td>SSIS, SSRS</td>
<td>Typical</td>
<td>Social–emotional</td>
<td>3-18</td>
<td>Child, parent, teacher</td>
<td>79 (parent), 75 (child), 83 (teacher)</td>
<td>10-25</td>
<td>Social skills (communication, cooperation, assertion, responsibility, empathy, engagement, self-control), competing problem behaviors (externalizing, bullying, hyperactivity/ inattention, internalizing, autism spectrum), and academic competence (reading achievement, math achievement, motivation to learn)</td>
<td>Likert 1-4</td>
<td>Spanish, Iranian</td>
</tr>
<tr>
<td>Assessment of Children’s Emotion Skills</td>
<td>ACES</td>
<td>Maximal</td>
<td>Emotional</td>
<td>4-8</td>
<td>Child</td>
<td>56</td>
<td>10-25</td>
<td>Emotional attribution accuracy, anger attribution tendencies, happiness attribution tendencies, sadness attribution tendencies</td>
<td>Multiple choice 1-5</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. (continued)
### Table 4. (continued)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Acronym</th>
<th>Maximal or Typical</th>
<th>Scope</th>
<th>Age (Years)</th>
<th>Versions</th>
<th>Length</th>
<th>Completion Time (Minutes)</th>
<th>Scales and Subscales</th>
<th>Response Scales</th>
<th>Other Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotion Regulation Checklist (Shields &amp; Cicchetti, 1997)</td>
<td>ERC</td>
<td>Typical</td>
<td>Emotional</td>
<td>6-12</td>
<td>Parent, teacher</td>
<td>24</td>
<td>10</td>
<td>Negativity, emotional regulation</td>
<td>Likert 1-4</td>
<td></td>
</tr>
<tr>
<td>Matson Evaluation of Social Skills with Youngsters (Matson, Rotari, &amp; Helsel, 1983)</td>
<td>MESSY</td>
<td>Typical</td>
<td>Social-emotional</td>
<td>4-18</td>
<td>Child, teacher</td>
<td>64 (teacher), 62 (child)</td>
<td>10-25</td>
<td>Appropriate social skills, inappropriate assertiveness, impulsiveness, overconfident behavior, jealousy</td>
<td>Likert 1-5</td>
<td>Japanese, Spanish, Turkish, Chinese, Dutch, Portuguese</td>
</tr>
<tr>
<td>Diagnostic Analysis of Nonverbal Accuracy (Nowicki &amp; Duke, 1989)</td>
<td>DANVA</td>
<td>Maximal</td>
<td>Emotional</td>
<td>4-adult</td>
<td>Child</td>
<td>16-32</td>
<td>30</td>
<td>Interpretation of child and adult faces, paralanguage and posture</td>
<td>Multiple choice 1-4</td>
<td></td>
</tr>
<tr>
<td>Differential Emotions Scale (Izard, Dougherty, Bloxom, &amp; Kotsch, 1974)</td>
<td>DES-IV</td>
<td>Typical</td>
<td>Emotional</td>
<td>7-adult</td>
<td>Child</td>
<td>36</td>
<td>10-15</td>
<td>Positive emotionality (interest, joy, surprise), negative emotionality (shyness, sadness, anger, disgust, contempt, self-hostility, fear, guilt and shame)</td>
<td>Likert 1-5</td>
<td></td>
</tr>
</tbody>
</table>
Table 5. Psychometric Properties of the Final 12 Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>U.S. Norms</th>
<th>U.K. Norms</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Internal Test–Retest Other</td>
<td>Factorial Convergent Discriminative Predictive Other</td>
</tr>
<tr>
<td>EQi:YV</td>
<td>Yes</td>
<td>Yes</td>
<td>.67-.90 .77-.89</td>
<td>Yes</td>
</tr>
<tr>
<td>CABS</td>
<td>No</td>
<td>No</td>
<td>.79 (total scale) 0.66-0.78 (4 weeks)</td>
<td>Yes (for both single and dual structure)</td>
</tr>
<tr>
<td>SCBE</td>
<td>Yes</td>
<td>No</td>
<td>.80-.89 .74-.87 (2 weeks), .59-.70 (6 months)</td>
<td>Yes</td>
</tr>
<tr>
<td>PKBS-2</td>
<td>Yes</td>
<td>No</td>
<td>.84-.97 .58-.86 (3 weeks), .69-.78 (3 months)</td>
<td>Standard error of measurement available; interrater reliability: .36-.63</td>
</tr>
</tbody>
</table>

(continued)
### Table 5. (continued)

<table>
<thead>
<tr>
<th>Measure</th>
<th>U.S. Norms</th>
<th>U.K. Norms</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Internal</td>
<td>Test–Retest</td>
</tr>
<tr>
<td>PTM-R</td>
<td>No</td>
<td>No</td>
<td>.59-.86</td>
<td>54-82 (2 weeks)</td>
</tr>
<tr>
<td>CRS</td>
<td>?</td>
<td>No</td>
<td>.49-.80</td>
<td>.49-.79 (child), .61-.91 (teacher), .89-95 (teacher), .91 (parent)</td>
</tr>
<tr>
<td>SSIS</td>
<td>Yes</td>
<td>No</td>
<td>.72-.97</td>
<td>.72-.87 (parent), .68-.92 (teacher), .73-97 (teacher), .73-95 (child)</td>
</tr>
<tr>
<td>ACES</td>
<td>No</td>
<td>No</td>
<td>.46-.70</td>
<td></td>
</tr>
<tr>
<td>ERC</td>
<td>No</td>
<td>No</td>
<td>.84-.92</td>
<td></td>
</tr>
<tr>
<td>MESSY</td>
<td>Yes</td>
<td>No</td>
<td>Split-half ( \alpha = .78 ) (child), .87 (teacher)</td>
<td>.60-.90</td>
</tr>
</tbody>
</table>
Table 5. (continued)

<table>
<thead>
<tr>
<th>Measure</th>
<th>U.S. Norms</th>
<th>U.K. Norms</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Internal</td>
<td>Test–Re-test</td>
<td>Other</td>
<td>Factorial</td>
</tr>
<tr>
<td>DANVA</td>
<td>No</td>
<td>No</td>
<td>.69-.81</td>
<td>.66</td>
</tr>
<tr>
<td>DES-IV</td>
<td>No</td>
<td>No</td>
<td>.50-.85</td>
<td>.30-.75</td>
</tr>
</tbody>
</table>

Note. EQI;YV = Bar-On Emotional Quotient Inventory: Youth Version; CABS = Child Assertive Behavior Scale; SCBE = Social Competence and Behaviour Evaluation Scale; PKBS-2; Preschool and Kindergarten Behaviour Scales–2; PTM-R = Prosocial Tendencies Measure–Revised; CRS = Child Rating Scale; SSIS = Social Skills Improvement System; ACES = Assessment of Children’s Emotion Skills; ERC = Emotion Regulation Checklist; MESSY = Matson Evaluation of Social Skills with Youngsters; DANVA = Diagnostic Analysis of Nonverbal Accuracy; DES-IV = Differential Emotions Scale–IV; ADHD = attention deficit hyperactivity disorder.
vision” in the Children’s Plan (DCSF, 2007b). At a secondary level, we sought to present an overview of the current state of this field. The review process resulted in the retention of 12 measures that vary greatly in both their implementation characteristics and psychometric properties. In this final section of the current article we present some key issues that have arisen during the course of our review, providing illustrative examples from the measures identified.

In relation to the establishment and sustained presence of measures in the academic literature, very few measures appear to have a good “shelf life.” Although 12 measures have yielded 4 or more publications in peer-reviewed journals, only a very small handful (notably the Diagnostic Analysis of Nonverbal Accuracy [DANVA], Social Competence and Behaviour Evaluation Scale [SCBE], and Social Skills Improvement System [SSIS]) have been used with great frequency (e.g., more than 10 articles). Considering that 189 measures were originally identified, it is clear that many are short-lived. This is not uncommon though—Butler and Gasson (2005) identified a similar issue in their review of measures of self-concept and self-esteem. One must also consider that the upsurge in interest in social and emotional learning (and therefore measurement of social and emotional skills) among children and adolescents is relatively recent (although interest in social skills can be traced back to Thorndike, a standard watershed in the literature is Salovey and Mayer’s (1990) article on emotional intelligence). Furthermore, the rather protean nature of the construct in question provides a fundamental problem for measure developers. Indeed, it has led some critics to state that it is “bereft of any conceptual meaning” (Zeidner et al., 2002, p. 215).

A further observation we have made is the imbalance among the scope and type of the measures identified. There appear to be many more well-established measures that tap social skills, as opposed to emotional skills or both. Furthermore, there are relatively few maximal behavior measures. This has important implications in terms of recommending a truly “direct” measure, since there is the danger that “typical behavior” social and emotional skills measures may simply tap aspects of personality (Gannon & Ranzijn, 2005; Schulte, Ree, & Carretta, 2004; Van Rooy, Viswesvaran, & Pluta, 2005). However, this has to be balanced against the practical advantages (e.g., time) of typical behavior measures, and also the emergent pockets of evidence regarding their validity (e.g., Petrides et al., 2006, found that scores on the Trait Emotional Intelligence Questionnaire were associated with peer nominations of behavioral descriptions, such as “cooperative” and “disruptive”). There are also concerns relating to the practicalities of capturing certain domains—such as self-awareness—in a maximal behavior paradigm.

The majority of measures reviewed here have been developed and standardized with American populations. For some of these measures, evidence of research carried out with other populations has concluded that they can be translated and adopted in different cultures while still retaining acceptable reliability and validity. However, of the 12 final measures produced by this review, only one (the Bar-On Emotional Quotient Inventory: Youth Version [EQI:YV]) has U.K. norms; several have no norms at all, and none have norms for countries other than the United States or the United Kingdom.
This is a particular concern given that cultural transferability between the United States, the United Kingdom, and other countries cannot be automatically assumed. In terms of psychometrics, our retained measures largely demonstrated acceptable properties using standard techniques (e.g., internal consistency, test–retest reliability, factorial validity, construct validity, discriminative validity), but evidence pertaining to more advanced analysis (such as those based on item response theory—as recommended by Terwee et al., 2007) was lacking for most measures (notable exceptions include the Matson Evaluation of Social Skills with Youngsters [MESSY] and SSIS). This may again reflect the emergent state of the field.

Beyond the aforementioned issues, there appears to have been little analysis of the applicability of the measures to different groups of children—such as different ethnic groups—although the SSIS is an exception in this regard (Van Horn, Atkins-Burnnett, Karlin, & Snyder, 2007). This kind of work is of course essential to ensuring that a given measure is appropriate for use with diverse groups of children and young people (e.g., in the aforementioned population survey), but just as importantly it can lead to development and improvement of the instrument itself. For instance, in Van Horn et al.’s (2007) recent psychometric analysis of the SSIS, an alternative factor structure with a stronger data fit to that proposed in the manual for the measure emerged.

The range of possible respondents in our final 12 measures also varied greatly. Only three (Child Assertive Behavior Scale [CABS], Child Rating Scale [CRS], and SSIS) have child, parent, and teacher versions, and five (EQI:YV, Prosocial Tendencies Measure–Revised [PTM-R], Assessment of Children’s Emotion Skills [ACES], DANVA, and Differential Emotions Scale [DES]) can only be completed by the child/young person. Consideration of the range of respondents for a given measure is important because each has access to unique information. Through introspection, the child has access to the most detailed information about himself or herself of any of the possible respondents. Furthermore, recent policy and legislation has placed increasing emphasis on the importance of the child’s perspective (e.g., Every Child Matters—Department for Education and Skills, 2003). However, as self-awareness/perception follows a developmental trajectory, older children and adolescents are likely to be more accurate responders than younger children (Denham, 2005b).

The limited number of measures that incorporate child, teacher, and parent responses creates difficulties for researchers and/or practitioners attempting to create a triangulated profile of the social and emotional skills of a child or group of children. The advantages to triangulating measures of social and emotional skills are clear, especially in response to overcoming difficulties in self-report (see above). Of course, one must factor in the increased time and expense in collecting data from numerous perspectives, especially given typically low response rates of certain groups (such as parents). However, given the relatively low concurrence between teacher, parent, and child responses in this area (e.g., Humphrey et al., 2008, found correlations as low as .25, indicating only 6% shared variance), consideration needs to be given as to whose perspective is prioritized.
Conclusion

In concluding, it is also important to note the limitations imposed on our review process. Although we have conducted an entirely systematic review, the protean nature of the area under investigation caused some difficulties, and other authors’ interpretations of “social and emotional skills” may differ from our own (meaning that other authors may have used different definitions—leading to different results!). The fact that the precise keywords used and/or the precise question under investigation may determine the result of a systematic review has been acknowledged as one of the integral characteristics of systematic reviews (Hannes, Claes, & The Belgian Campbell Group, 2007; Nind, 2006). Furthermore, the systematic review process itself is not without criticism. For example, publication bias (the tendency for a greater proportion of studies with statistically significant positive results to be published—Torgerson, 2006) might have limited the published articles for a given measure and therefore the measure’s inclusion in the later stages of our review. While we acknowledge this, we felt that this criterion was an important way to assess the “academic presence” of the measures. Furthermore, all criteria were negotiated with users from the project steering group. Such close consultation with users not only contributes to the significance of the review and its fitness for purpose but also helps in addressing bias relating to the interpretation of findings (Harlen & Crick, 2004).

The systematic review reported in this article identified 12 measures with varying implementation characteristics and psychometric properties. We feel that this variability—particularly in relation to psychometric properties—reflects the emergent state of the field. Although there has been an interest in social and emotional learning over the last two decades, measure development has struggled to keep pace. This is undoubtedly reflected in the fact that many reports of social and emotional learning program evaluations have not contained any social and emotional skills outcome measures, instead using proxy indicators of success like reductions in mental health problems or increases in attendance (Humphrey et al., 2007).

Declaration of Conflicting Interests

The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Funding

The research reported in this article was funded by a grant from the Department for Children, Schools and Families (REF: 2008/056).

References

*References marked with an asterisk indicate key article/document for final 12 measures.


