
Development and Psychometric Properties of the Early Development Instrument (EDI): A Measure of Children's School Readiness*

MAGDALENA JANUS and DAVID R. OFFORD

Offord Centre for Child Studies, Department of Psychiatry and Behavioural
Neurosciences, McMaster University

Abstract

The Early Development Instrument (EDI), a teacher-completed measure of children's school readiness at entry to Grade 1, was designed to provide communities with an informative, inexpensive and psychometrically sound tool to assess outcomes of early development as reflected in children's school readiness. Its psychometric properties at individual level were evaluated in two studies. Five a priori domains – physical health and well-being, social competence, emotional maturity, language and communication, and cognitive development and general knowledge – were tested in a factor analysis of data on over 16,000 kindergarten children. The factor analyses upheld the first three domains, but revealed the need to develop two new ones, resulting in the final version of the EDI consisting of: physical health and well-being, social competence, emotional maturity, language and cognitive development, communication skills and general knowledge domains. These final domains showed good reliability levels, comparable with other instruments. A separate study ($N = 82$) demonstrated consistent agreements in parent-teacher, interrater reliabilities, concurrent validity, and convergent validity. These results establish the EDI as a psychometrically adequate indicator of child well-being at school entry.

Résumé

L'instrument de mesure du développement à la petite enfance a été conçu afin de fournir un outil approprié, peu coûteux et valide sur le plan psychométrique, permettant d'évaluer les capacités de l'enfant à entrer en première année primaire. Les propriétés psychométriques de l'instrument ont été évaluées au cours de deux études. Une analyse factorielle menée à partir des données recueillies auprès de 16 000 enfants de maternelle a permis de tester la présence de cinq facteurs théoriques : santé physique et bien-être, compétence sociale, maturité affective, langage et communication, développement co-

gnitif et connaissances générales. Les analyses factorielles ont pu confirmer la présence des trois premiers facteurs, mais ont dégagé la nécessité de modifier les deux derniers. Ceci a permis de formuler la version finale de l'instrument de mesure du développement à la petite enfance qui est composée des cinq facteurs suivants : santé physique et bien-être, compétence sociale, maturité affective, langage et développement cognitif, habiletés à la communication et connaissances générales. Ces cinq domaines offrent de bons indices de consistance interne, comparable à d'autres instruments. Une seconde étude ($N = 82$) a démontré la présence d'une bonne concordance entre l'évaluation de parents et des enseignants, la fidélité entre les évaluateurs, la validité concomitante et la validité convergente. Ces résultats confirment que l'instrument de mesure du développement à la petite enfance est un outil psychométrique adéquat pour évaluer les aptitudes de l'enfant lors de l'entrée à l'école primaire.

For many decades, the average percentage of children with impairing cognitive and behaviour problems in elementary school remained constant at about 25% (Achenbach, 1991; Offord & Lipman, 1996). This happened despite an increased awareness of the importance of the early years and more widely available intervention programs for young children. Although children's problems at school entry may generally occur at a level that would not necessarily require clinical treatment, there is evidence that as much as, or perhaps even more than, 25% of children experience some difficulties that prevent them from taking full advantage of the education offered by schools (e.g., Frank Porter Graham Child Development Center, 1999; Rimm-Kaufman, Pianta, & Cox, 2000). Differences in children's first years of school have long-term sequelae for their school career and later life (Alexander & Entwisle, 1988) since even minor differences in academic achievement at Grade 1

* Lorne Sulsky was the Action Editor for this manuscript
Rédaction de ce manuscrit : Lorne Sulsky

tend to intensify over the years rather than converge. A population-based model of health suggests that low-risk, small deficiencies in large populations contribute to the burden of ill health more than severe problems in a minority of patients (Rose, 1994). Seen in this context, children's school readiness is a health-relevant, measurable outcome that has long-term consequences for population health.

Because early child development is heavily influenced by the quality of stimulation, support, and nurturance in the environments where children grow up, school readiness can be broadly understood as an outcome of the early years. It is a useful construct because it acknowledges the importance of the early years for children's future development (Schonkoff & Phillips, 2000; Shore, 1997). The developmental outcomes, which could be operationalized as school achievement, behaviour, and cognitive outcomes, or school drop-out rates, depend on the combination of the individual and collective factors. While the individual variation routinely contributes a larger proportion of variance than neighbourhood factors (Boyle & Lipman, 2002), there is nevertheless a growing body of evidence suggesting that the neighbourhood and societal factors also matter, especially, though not only, within the context of poverty (Brooks-Gunn, Guo, & Furstenberg, 1993; Chase-Lansdale, Gordon, Brooks-Gunn, & Klebanov, 1997). Data from the National Longitudinal Study of Children and Youth in Canada (NLSCY) allow us to identify factors having an impact on children's outcomes beyond children's individual characteristics in three broad areas (C. Hertzman, personal communication, March 27, 2005): family (including income, education, parenting style); neighbourhood (including safety and cohesion, socio-demographic mix); and society (including support for parenting, e.g., access to high-quality care arrangements). Even small changes in any of these three areas can dramatically contribute to the social processes behind the well-being of all children, and change the distribution of risk at a given level (Offord et al., 1999). To paraphrase Rose (1994) and Offord, Kraemer, Kazdin, Jensen, and Harrington, (1998), a large number of children at a small risk for school failure may generate a much greater burden of suffering than a small number of children with a high risk. Yet, in the current climate, school or preschool interventions are implemented based on individual diagnostics usually only with serious, clinical cases, providing help to the few whose impairments are severe. Broad assessments of children's development in all relevant areas, such that could provide an overall evaluation of the range of developmental outcomes in a community

are rarely used. However, only such assessments can provide background to broadly cast interventions, called "universal" (Offord et al., 1998), which would have the advantage of helping all children, and thus raising the population level of school readiness.

In the last decade, the issue of children's readiness for school finally reached the forefront of interest not just among academics and educators, but also communities and even politicians. In Canada, the 1997 Speech from the Throne contained the commitment to "measure and report on the readiness to learn of Canadian children so that we can assess our progress in providing our children with the best possible start." This goal was picked up by a score of communities across the country, making its way into programs and coalitions (Janus & Offord, 2000; McCain & Mustard, 1999). More recently, one of the political parties in Canada embraced a set of principles summarized as "QUAD" (Liberal Party of Canada, 2005), to be used within the early child care and learning system, thus ensuring the relevance of developmental outcomes for social programs. QUAD stands for quality, universality, accessibility, and developmental outcomes and represents a promising opportunity to improve the early years' experiences for children.

A great deal of debate has been waged over the theoretical basis of school readiness and consequent methods of measurement: When should it be measured, who should be the informant, what should be included (Love, Aber, & Brooks-Gunn, 1994)? Readiness for school and its measurement have received their share of attention in the developmental and educational literature, and several reviews have been produced to highlight the difference in approaches over time (Meisels, 1998, 1999; Phillips & Love, 1995; Wenner, 1995).

In the first half of the 20th century, assessment of school readiness was virtually synonymous with decision-making for kindergarten entry or delay. The tests used focused on reading and writing, and were intended to identify children who should not start regular kindergarten classes. These trends can be traced to the history of the definition of school readiness. In the early formulations, it was an ability to perform indicated, usually cognitive, language or motor tasks on demand (e.g., Gesell test; Ilg, Ames, Haines, & Gillespie, 1978). Meisels (1998) classifies these types of definitions as "idealist/nativist" or "empiricist/environmentalist" perspectives. In the idealist/nativist view, readiness can be seen as a within-the-child phenomenon, whereby a child's readiness for school is achieved through a maturational process, with little or no impact from the environment (including parents, experiences, etc.). The

child's development proceeds through predictable stages and cannot be altered by external influences. Developmental tests were designed to measure this concept of readiness; however, by adhering too strictly to specific goals, they tended to misclassify too many children as not ready. The empiricist/environmentalist perspective claims that readiness is a set of particular behaviours, skills, and personality traits that are basic precursors to school achievements and are easily measured. Therefore, testing should focus on external evidence of what the child can do. This conceptualization of readiness provided a theoretical basis for a number of assessments, which tended to be curriculum-based or specific-tasks-oriented. Unfortunately, similarly to strict developmental tasks, such tests often resulted in inappropriate classification of many children.

Currently, kindergarten readiness or school readiness screening measures are often still utilized to provide a basis for decision-making on retention, tracking, and services (Meisels, 1998), or to be held as performance standards for schools' accountability (La Paro & Pianta, 2000). The measures could be skill-oriented, tapping into the degree of mastery of specific skills, or developmentally oriented, assessing the child's developmental age (Costenbader, Rohrer, & DiFonzo, 2000). Over and above those measures, school districts use locally constructed tests, or informal observations. New York State school districts, for example, use the four types of kindergarten screening in almost equal proportions (May & Kundert, 1992).

In view of the purposes for which they are commonly used, kindergarten and school readiness measures are usually reviewed and validated from the perspective of their accuracy in identifying children at risk for school failure (e.g., Costenbader et al., 2000), rather than their adequacy of reflecting the concept of school readiness (Meisels, 1998). Seven of the many well-known and widely used measures will be briefly described below. We will review their major domains, psychometric properties, and the training needs for assessment.

One of the earliest measures of school readiness is the Gesell School Readiness Test (GSRT), an assessment of skills that are purportedly achieved solely through a maturational process (Ilg et al., 1978). It is administered individually to children as an interview by a trained examiner, who needs to consider the content and manner of the child's response. The tasks include writing, drawing, visual and motor coordination, and the child's verbal expressions. Currently, the GSRT is described as an observational, qualitative tool, with results being interpreted clinically

(Lichtenstein, 1990). It has often been used to determine children's readiness for kindergarten, and followed up with placement decisions (Graue & Shepard, 1989). In Graue and Shepard's study, the developmental age measure on GSRT in kindergarten correlated with the Grade 1 report card only at 0.23. About 60% of children identified as not ready were misdiagnosed based on Grade 1 data. Similarly, no differences were detected between children classified as ready and unready by the GSRT before kindergarten entry in later measures of Grade 1 remedial placement, or academic scores in Grade two and three (Buntaine & Costenbader, 1997). Lichtenstein (1990) reports an interrater agreement of placement recommendation of 78%, based on 46 cases. Few other psychometric properties of the Gesell are available in literature.

Among some of the most frequently used skill-oriented measures are such readiness tests as the Developmental Indicators for the Assessment of Learning (DIAL-R) (Mardell-Czudnowski & Goldberg, 1998), and the Brigance Diagnostic Inventory of Early Development (Brigance, 1992; Glascoe, 1995). Both of these measures require a trained professional to administer the assessment to children. The assessments include motor, cognitive/conceptual, and language areas in 3 (DIAL-R) or up to 13 subtests (Brigance, 1992). Each of the two tests offers a parent-completed questionnaire to assess social skills and development. DIAL-R is reported to have high interrater and test-retest reliabilities (0.90 and 0.86, respectively), and both sensitivity and specificity around 85% (Mardell-Czudnowski & Goldenberg, 1998). A positive predictive value of only 0.53, demonstrated in one study (Jacob, Snider, & Wilson, 1988), suggests that if used for identifying children at risk for future academic difficulties, it carries a high "false-positive" rate. The Brigance is a criterion-referenced inventory of skills, with psychometric data similar to those reported for the DIAL-R. One study of 95 middle-class white 4-5-year-old children (Wenner, 1995) found that referrals to special problems and nonpromotion were correctly predicted with the Brigance scores for 67% of children in the sample.

Yet another school readiness assessment, the Lollipop Test (Chew & Lang, 1990), includes four subtests covering recognition and identification of shapes, colours, pictures, letters, and numbers, administered by trained examiners. Chew and Lang (1990) and Chew and Morris (1989) showed that Lollipop Test's domains mapped closely to those tested on DIAL-R and the Metropolitan Readiness Test (MRT; Swanson, Payne, & Jackson, 1981) yet required

a shorter testing time. Neither the Lollipop nor the MRT have specific “readiness levels” used to classify children as ready or not; their main purpose is to predict first grade academic success from a kindergarten testing. The ability of both MRT and the Lollipop to predict grades, and standardized achievement test results in Grades 1, 3, and 4 are similar and moderate to high in magnitude (Chew & Morris, 1989).

The Phelps Kindergarten Readiness Scale (Augustyniak, Cook-Cottone, & Calabrese, 2004; Duncan & Rafter, 2005), a newer addition to the spectrum of measures, was developed explicitly to measure “academic” readiness of children before entry to kindergarten. It contains six major domains: verbal processing, perceptual processing, and auditory processing, evaluating children’s language competence, ability to compare and reproduce shapes, and memory. Test-retest reliabilities vary from 0.61 to 0.87 for individual domains. Concurrent validity, established in the fall of the kindergarten year with the Woodcock-Johnson III Test of Achievement (Woodcock, McGrew, & Mather, 2001), is 0.59. Predictive validity values for the Phelps’ total readiness score are available for an eight-month period with the Woodcock-Johnson, and vary from 0.39 (reading subtest) to 0.53 (math subtest). In addition, a recent study demonstrated correlations of .47 and .51 between the Phelps’ kindergarten score and New York State fourth-grade assessments in language and mathematics tests, respectively (Augustyniak et al., 2005).

The Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn 1981), a test of receptive vocabulary, has also occasionally been mentioned as a measure of school readiness (e.g., Kohen, Brooks-Gunn, Leventhal, & Hertzman, 2002; Zill et al., 2001); however, it is rarely used as a sole screening method (Costenbader et al., 2000). Within its limited skill testing range, the PPVT has adequate psychometric properties for screening purposes (Dunn & Dunn, 1981), and is easy and quick to administer.

Of the seven assessments briefly reviewed above, only two offer an optional measure of children’s socio-emotional development, a parent-completed questionnaire (Brigance and DIAL-R). None of them allows for school-based evidence of children’s relationships with peers or social competence with adults other than parents. Most assessments provide some measure of children’s motor coordination, confined to fine motor skills (e.g., drawing, writing letters or numbers, copying shapes). None account for children’s gross motor skills (e.g., running, jumping) or physical independence. Few studies provided information on interrater reliability; however, it is proba-

bly implicit in the fact that the assessments have to be administered by professionals trained in the specific instrument. The need for an external examiner to administer the tool, rather than reliance on a report by an adult familiar with the child, explains why an examiner would not be informed well enough to rate the child’s social behaviour. From the implementation point of view, need for examiners trained in specific tools increases costs of assessments.

It is important to note that only three of the tools mentioned above, the DIAL-R, Brigance, and Phelps were purportedly validated specifically to screen children who were not ready. Nevertheless, Costenbader et al. (2000) and Duncan and Rafter (2005) suggest that even these three be used in conjunction with other, more detailed psychoeducational evaluation of readiness. Together with high implementation costs and lack of information on children’s social and emotional development, this indicates low cost-effectiveness of these measures.

Any of the assessment tools reviewed above could be used to provide information for groups of children, or even population-level data. Population-level community reporting theoretically can be achieved by aggregating any measurement available for all individuals in the community, or a representative sample, similarly to the way census reporting is carried out (Statistics Canada, 2005). However, most available school readiness assessments provide information only on the cognitive and language aspects of child development. Also, since all of them are implemented through a direct assessment with an individual child, it would be extremely costly to include all kindergarten children in such testing. At this point in time, direct cognitive assessments are rarely done for whole populations of young children; rather, schools have resources for assessments of children identified as at risk by teachers (Love et al., 1994).

There appears to be a consensus among educational and developmental experts that school readiness should be understood as not merely cognitive skills, but rather as a holistic concept involving several developmental areas such as cognitive, socio-emotional, and physical (Jimerson, Egeland, & Teo, 1999; Love et al., 1994; Meisels, 1999). Competence in all these areas will ensure that children are ready to benefit from educational activities offered in the school environment (Janus & Offord, 2000). Therefore, assessment of children’s cognitive status only is no longer adequate. Furthermore, making costly measures available for populations of children (as opposed to targeted subgroups) would require far greater investments than is currently feasible.

Meisels (1999) describes yet another perspective

on the measurement of school readiness, following Love et al. (1994): a “social constructivist” approach, where school readiness is defined with reference to how children’s behaviour and development are supported and what the children should be ready for. This approach requires a community-level measurement strategy, where assessment of children’s abilities is only one of the components, and has to be put in the context of the children’s past few years and the realities of where they will be educated. Moreover, this strategy explicitly involves the community’s willingness for action based on the results. By providing a strategy, and including a context, the social constructivist view is the most comprehensive approach to the measurement of children’s school readiness.

This article reports on the development of a new school readiness measurement tool. The holistic framework of children’s outcomes at school entry was adopted to provide communities with benchmarks useful for planning of intervention and prevention. By emphasizing the population-level of data interpretation, this tool overcomes the barrier of seeing the assessment of school readiness as an individual process labeling a child with a deficit. Because it comes at the cusp between early development and school entry, such an assessment has a potential to mobilize communities into providing opportunities accessible to all children.

The driving force behind the design of the current measure was the desire to provide communities with a feasible, acceptable and psychometrically reliable instrument that could be used for whole populations of children to monitor community efforts to improve early years’ outcomes over time (Janus & Offord, 2000). The Early Development Instrument (EDI) is a relatively short, easy-to-administer tool in the format of a teacher-completed checklist, whose results can be aggregated to various levels (e.g., groups like girls or boys, children living in a neighbourhood, children attending regular or immersion programs, as well as all children) and therefore easily lends itself to linkages with other population and community data (Janus, Walsh, Viveiros, & Offord, 2002). Within the theoretical framework of approaches to the measurement of school readiness, the EDI is positioned in the context of the “social constructivist” approach, by providing the “child” component necessary to complete the whole picture of community-based school readiness.

The focus of the current tool is on children’s readiness to enter grade one, rather than on their ability to start attending school at a kindergarten level. This follows conceptually the distinction made by Kagan

between the “readiness to learn” and “school readiness” (Kagan, 1992; Kagan & Neuman, 1997). The first refers broadly to the child’s neurosystem being ready from birth to process information it is being exposed to and develop accordingly; the second is a narrower view reflecting the specific domains of development relevant to school-based learning as children mature around the age of 4 to 5 years. Kindergarten attendance is still optional in many districts, yet it provides children with an undisputed advantage in first-grade outcomes (Entwisle & Alexander, 1999). Moreover, the structure of teaching in kindergarten classes is very different from grade one. Kindergarten provides the transition between the play-based preschool and home environment to the academically based environment of grade school, and ensures that children have the opportunity to consolidate skills relevant to grade-school learning. A school readiness measure taken at the beginning of the kindergarten year would fall back on the mistaken assumption of a common core of learning happening before school (Meisels, 1999). However, children who do poorly at school readiness measures taken prior to or at the beginning of kindergarten, often do well on similar measures of achievement by the end of the year (Meisels, 1987). Even comprehensive screening of children before school entry rarely provides highly reliable results (Pianta & McCoy, 1997). Readiness is a process occurring over time, and cannot simply be completed by the first day of kindergarten. As Meisels puts it, “... [since] readiness is a process and schools are by necessity a major contributor to this process, then a period of common schooling needs to occur in which this process can take place” (Meisels, 1999, p. 62). Therefore, an assessment of children’s school readiness for grade one should ideally be carried out well into the kindergarten year, yet with sufficient time before the end of the year to allow the use of the collected data for grade-one programming.

The EDI combines several areas that have been identified as relevant to children’s school readiness (Doherty, 1997; Kagan, 1992): physical health and well-being, social competence, approaches to learning, emotional maturity, language development, cognitive development, communication skills, and general knowledge. This paper describes the development, factorial structure, and initial psychometric properties of the Early Development Instrument (EDI): A Population-Based Measure for Communities.

Instrument Development

Conceptual Framework

Kagan (1992) and Doherty (1997) outlined the five areas of school readiness as pertaining to: physical

TABLE 1
Sample of Questions From the EDI

Physical Health and Well-Being	Social Competence	Emotional Maturity	Language and Cognitive Development	Communication Skills and General Knowledge
How often has the child arrived to school too tired to do school work?	How would you rate this child's ability to get along with peers?	Would you say that this child will try to help someone who has been hurt?	Would you say that this child is generally interested in books?	How would you rate this child's ability to communicate own needs in an understandable way?
How would you rate this child's ability to manipulate objects?	Would you say that this child accepts responsibility for actions?	Would you say that this child gets into physical fights?	Would you say that this child is showing awareness of rhyming words?	How would you rate this child's ability to tell a story?
How would you rate this child's level of energy throughout the school day?	Would you say that this child is able to solve day-to-day problems by himself?	Would you say that this child can't sit still, is restless?	Would you say that this child is interested in games involving numbers?	How would you rate this child's ability to take part in imaginative play?

well-being and appropriate motor development; emotional health and a positive approach to new experiences; age-appropriate social knowledge and competence; age-appropriate language skills; and age-appropriate general knowledge and cognitive skills. There is adequate evidence in literature to indicate that each area has an important impact on children's adjustment to school and short- or long-term school achievement (Doherty 1997; Jimerson et al., 1999; Love et al., 1992). This view was confirmed in a discussion held with educators and early childhood experts, who requested that each domain be represented in the new instrument to provide a comprehensive assessment of children's school readiness.

Item Selection

The items for the EDI were derived from existing instruments, key informant interviews, and focus groups, as suggested by Streiner and Norman (1995). A review of some commonly used teacher and parent-completed tools was carried out and items for the instrument were chosen to fit specific areas. An initial base of 128 questions was created, over 60% of which were modified from the items in the Canadian National Longitudinal Study of Children and Youth (NLSCY). The NLSCY is a federally funded study of a representative sample of Canadian children. The items relevant to child behaviour and language and cognitive areas in the NLSCY were based on a number of standardized instruments and consultations with experts (NLSC Project Team, 1995). Because it was apparent at the time that the NLSCY did not adequately cover all the areas relevant to school readiness (Morongiello, 1997), new questions were constructed by the authors for the missing areas, based on Doherty (1997), and field-tested with teachers and

researchers. The first draft of the EDI was reviewed by a group of educators, early years' professionals, and academics with expertise in the field. Changes were made to the draft, and subsequently four focus groups with kindergarten teachers were conducted. For several questions, wording was changed; others were dropped and some added, based on teachers' recommendations. Table 1 contains examples of questions in each domain. In addition, some answer/scoring options were modified in response to feedback from teachers. In particular, items referring to specific skills were provided with only yes/no options, rather than along the continuum. Teachers indicated to us that these were a better reflection of children's school readiness. Conversely, answer options to several questions on children's overall skills were expanded to five, as these were perceived to be more variable. An EDI guide, accompanying the instrument, was developed to provide brief explanations and anchors for the items.

Instrument Description

The first page of the instrument requests information on child demographic variables (gender, date of birth, language), as well as on selected variables related to the child's school-based designations (e.g., English as a second language, special needs, type of class), and the completion date.

Pages 2 to 7 of the EDI contain questions relevant to the five domains of school readiness: physical health and well-being; social competence; emotional maturity; language and cognitive development; and communication skills and general knowledge. Most of these are "core" questions, which means they directly contribute to one of the five domains. There are also questions related to children's special skills and spe-

cial problems. Finally, the last page of the instrument contains questions about children's prekindergarten experience (early intervention, child care, preschool, etc.). Only questions in the five core domains are used to score children's school readiness.

The questionnaire takes between 7 to 20 minutes to complete. It is recommended that it be completed in the second half of the kindergarten year, to give teachers the opportunity to get to know children in their class.

The descriptions below refer to the domains in the finalized instrument. The total number of core questions in the final version of the instrument is 103 (full version of the instrument is available from authors upon request, or at the website).

All core questions are scored from 0 (lowest score) to 10 (highest score). The domain score is calculated as a mean score of all the valid answers. Thus, scores for each domain have the same minimum and maximum values, even though there are different numbers of items. As the feedback from focus groups indicated, this way of scoring and presenting the results proved to be easier to communicate to audiences with little or no research background.

No more than 30% of missing answers are allowed per domain. If more than one domain is missing, the questionnaire is not considered complete and is discarded from analyses. On average, this occurs in no more than 3% of cases. There is no total score on the EDI.

Physical Health and Well-Being

This domain contains 13 items and refers to children's physical preparedness for the school day, fine and gross motor skills, energy level throughout the day, and physical independence (examples are in Table 1). Ten questions are answered on a 5-point scale (from never to always, or excellent to very poor), scored from 10 (best) to 0 (worst) in 2.5 point intervals: 10, 7.5, 5, 2.5, and 0. Three questions, about the child's washroom independence, hand preference, and level of coordination, are answered in a yes/no format. "Yes" is scored as 10 and "No" as 0.

Social Competence

This domain contains 26 items and covers the following areas: competence and cooperation in working together with others, ability to remember and follow rules, curiosity and eagerness, approaches to learning and problem-solving. (See Table 1 for exam-

ple questions.) All answers are scored on a 3-point scale: often or very true (10), sometimes or somewhat true (5), and never or not true (0).

Emotional Maturity

This domain contains 28 items¹ and covers prosocial behaviour, aggression, inattention and hyperactivity, and anxious behaviours. All answers are scored on a 3-point scale: often or very true (10), sometimes or somewhat true (5), and never or not true (0).

Language and Cognitive Development

This domain contains 26 items and refers to the child's ability to use language correctly and covers cognitive aspects of language and numeracy, in several areas: basic literacy and numeracy skills, interest and memory, and more complex literacy. All answers are scored on a 2-point scale: "yes" (10) if a child possesses a skill and "no" (0) if she/he does not.

Communication Skills and General Knowledge

This domain has eight questions and covers the child's ability to clearly communicate his/her own needs and thoughts in a way that is understandable to both adults and other children, the ability to understand others, to articulate clearly, as well as aspects of general knowledge. In contrast to the previous domain, this one is about effective communication regardless of the grammatical correctness. Seven answers are scored on a 5-point scale from very poor (0), to excellent (10), in 2.5 increments (0, 2.5, 5, 7.5, and 10). One answer is scored on a 3-point scale (often, 10, sometimes, 5, and never, 0).

Additional Questions

The three additional sections of the EDI cover children's special skills, special problems, and aspects of the prekindergarten history. Seven general areas in which young children could demonstrate special skills are listed: numeracy, literacy, arts, music, athletics/dance, problem-solving, and other. They are simply scored as "yes" (1) and "no" (0), and summed up, so for each child there is a total score indicating the number of special skills they demonstrate. Nine special problem areas are listed: physical, visual, hearing, speech, learning, emotional, behavioural, home environment, and other. These are scored in the same way as special skills. For the prekindergarten history, questions about the following areas are asked: child attendance at any early intervention program, preschool, language or religion classes, Junior Kindergarten level, and participation in non parental care. The prekindergarten history items are stand-alone questions.

1 Two additional items were added to this domain in the subsequent year, thus bringing the total number of items to 30. This addition did not change the factor structure of the instrument.

TABLE 2
Description of Sample in Each of the Studies

Study	Valid N	Age (in years)	% Girls	% with English as a second language (ESL)	% Mothers with more than high school education	% Living with married biological parents
2	82	5.6±0.36	50.0	6.9	79.3	83.3
1	Total:16,074				N/a	N/a
	JK 4,934	4.9±0.28	50.2	31.9		
	SK 11,140	5.9±0.30	48.4	27.7		

TABLE 3
Average Sociodemographic Characteristics for School Areas Participating in Study 1

Site	Average 1996 family income (\$)	Unemployment rate	% No high school diploma
1	75,291	8.5	16.0
2	70,331	9.4	16.4
3	66,343	10.9	19.7
4	37,901	14.1	26.6
5	45,348	13.2	31.3
6	24,193	15.4	43.7
Canada 1996	54,583	10.10	22.7

Study 1

Method

Participants. The EDI was implemented in six sites and completed for 16,583 students. Of those, 16,074 or 97% of questionnaires were complete (had no more than one domain and no more than 30 answers in total missing). The sites comprised three large urban ($N = 15,319$) and three smaller rural areas ($N = 755$). Thus, the rural sites contributed 5% of the sample, while the urban sites contributed 95%. Statistics Canada (2005) reports the distribution of the Canadian population to be 80% urban and 20% rural. All schools within the school boards were involved with an exception of one site where only about 25% of schools participated. As indicated in Table 2, there were approximately equal proportions of boys and girls in the sample, and for about 30% of children, English was not their first language. No other demographic data were available on the children.

Although information on the individual socioeconomic status of the families of the children in the sample was not available, the neighbourhood SES indicators (average income, unemployment rate, and high school education) were established for the enumeration areas in which participating schools were located, based on the 1996 Canadian census data, accessed through the DMTI Spatial Inc. Digital Data. Enumeration areas were the smallest geographical

areas for which census data were available. The mean SES indicators were computed for each site and are presented in Table 3, alongside Canadian national averages from the 1996 census. For three of the sites, the SES indicators were better than Canadian averages, and for the remaining three, they were lower.

Part of the data were collected in Ontario, Canada, where in many sites children can start kindergarten at a younger age level, called "Junior Kindergarten." These children turn four years old in the calendar year they enter school. The majority of children in Canada, however, start school at the 5-year-old level, called "Senior Kindergarten." Since the sample included children at both kindergarten levels, the reporting will be split, where appropriate, into the Junior (JK) and Senior levels (SK).

Analyses. The data were analyzed using several techniques to confirm the a priori domain/factor structure. A confirmatory factor analysis was computed on the full sample using principal axis factoring extraction method with promax rotation, allowing factors extracted to be correlated. Because of the natural clustering of the data by classroom, the within- and between-classroom factor structure was explored. A multilevel confirmatory factor analyses, developed by Muthen (1994), which involves a simultaneous analysis of both the within- and between-group factor

Development and Psychometric Properties of the Early Development Instrument (EDI) 9

TABLE 4
Summary of the Factor Analysis (Principal Axis Factoring)

Factor	Physical Health and Well-Being N of items, Mean loading (range)	Social Competence N of items, Mean loading (range)	Emotional Maturity N of items, Mean loading (range)	Language and Cognitive Development N of items, Mean loading (range)	Communication Skills and General Knowledge N of items, Mean loading (range)
1		17, 0.641 (0.411-0.938)			
2				16, 0.557 (0.204-0.961)	
3					8, 0.861 (0.426-1.061)
4			8, 0.836 (0.769-0.962)		
5	1, 0.119		8, 0.627 (0.448-0.876)		
6			7, 0.691 (0.500-0.833)		
7	5, 0.787 (0.704-0.858)				
8				6, 0.633 (0.310-0.924)	
9		4, 0.835 (0.515-1.050)		1, 0.211	
10	4, 0.611 (0.504-0.702)				
11			5, 0.511 (0.302-0.697)		
12		5, 0.570 (0.333-0.819)		1, 0.265	
13				2, 0.894 (0.876-0.913)	
14	3, 0.310 (0.271-0.387)				
Total	13, 0.572	26, 0.647	28, 0.682	26, 0.576	8, 0.861
Cumulative % variance explained	4.8	37.7	48.2	58.9	63.1

structure using the Mplus software (Muthen & Muthen, 2004), was employed to assess the factor structure for each domain. In order to assess the need for further multilevel analyses, the proportion of variance between teachers or the intraclass correlation coefficients (ICC) obtained in the above procedure were examined. Finally, the average teacher reliability (indicating consistency levels) for each domain was assessed using the unconditional multilevel models with the hierarchical linear modelling (HLM) methodology. Software used included SPSS and Mplus.

In addition, the internal consistency indicators (Cronbach's alpha) for the EDI domains were computed, and the convergent validity analyses on age

and gender relationship with the EDI scores were carried out.

Results

Factor structure. The principal axis factoring analysis revealed 14 factors, with eigenvalues greater than one. This was expected, since some of the broad domains covered more than one distinct factor, and forcing the distribution into only five would have been counterproductive (Gorsuch, 1983). The 14 factors were aggregated into the five domains based on the conceptual framework (Table 4). For all but three items, the highest loading belonged to a factor within the predicted domain. However, even for these three,

TABLE 5
Indices of Fit for Multilevel Models

Domain	Between-teacher			Within-teacher			Difference in CFI
	CFI	RMSEA	SRMR	CFI	RMSEA	SRMR	
Physical Health and Well-Being	0.699	0.248	0.190	0.739	0.140	0.085	0.105
Social Competence	0.414	0.224	0.202	0.502	0.151	0.141	0.065
Emotional Maturity	0.657	0.200	0.092	0.724	0.130	0.091	0.001
Language and Cognitive Development	0.626	0.157	0.091	0.739	0.098	0.073	0.018
Communication Skills and General Knowledge	0.955	0.170	0.015	0.972	0.101	0.019	0.004

Note. CFI = comparative fit index; RMSEA = root mean square error approximation; SRMR = standardized root mean residual.

TABLE 6
Results of the Multilevel Confirmatory Analyses

	Intraclass coefficients (ICC)	Teacher reliability estimates
Physical Health and Well-Being	0.305	0.843
Social Competence	0.196	0.759
Emotional Maturity	0.206	0.770
Language and Cognitive Development	0.245	0.804
Communication Skills and General Knowledge	0.258	0.813

the second highest loading belonged to the predicted domain. Seven items were retained despite loading less than 0.3 on a factor, due to perceived importance by teachers participating in the focus groups (three of those were the ones that did not separate as expected). These items were: independent in washroom, well-coordinated, sucks a finger, knows how to handle a book, interested in books, interested in reading, remembers things easily. All remaining items loaded 0.3 or higher on the factors.²

The 14-factor solution accounted for 63.1% of the variance. The factors contributed to the five domains in the following way: Physical Health and Well-Being, Factors 7, 10, and 14 (and one item from 5), 4.8% of variance; Social Competence, Factors 1, 9, and 12, 32.9% of variance; Emotional Maturity, Factors 4, 5, 6, and 11, 10.5% of variance; Language and Cognitive Development, Factors 2, 8, and 13 (and one items each from 9 and 12), 10.7% of variance; Communications Skills and General Knowledge, Factor 3, 4.2% of variance (Table 4).

The Muthen procedure for exploring between- and within-group factor variance confirmed the factor structure. Table 5 shows the fit indices for each

domain. The values for “between” and “within” comparisons are very close, regardless of the models employed (between or within). This indicates that the factor structure within classrooms is similar to the structure between classrooms.

Teacher reliability. Intraclass correlations for the five domains are in Table 6. The ICCs for all the items varied from the minimum of 0.017 to the maximum of 0.400, with 57% of items at 0.200 and less, indicating low levels of variability between classrooms or teachers. In the case of all items and domain scores, the majority of variance came from children (0.600 to 0.983).

Average teacher consistency in each domain, estimated with the HLM reliabilities, varied from 0.76 to 0.84 (Table 6).

Internal consistency. The internal consistency of the specified domains was explored using Cronbach’s alpha. All five domains showed satisfactory internal consistency levels: Physical Health and Well-Being 0.84; Social Competence 0.96; Emotional Maturity 0.92; Language and Cognitive Development 0.93; and Communications Skills and General Knowledge 0.95.

Relationship to age, gender, and English as a second language status. The EDI was intended to be an instrument based on the child’s developmental status and not achievement in relation to specific curriculum objectives. Therefore, it was imperative that it should be sensitive to the child’s age and gender. One-way analysis of variance (ANOVA) was used to compare the five domain scores for girls and boys. Table 7 shows the means and standard deviations, separately for the cohort of 4-year-olds and 5-year-olds. Girls were rated on average significantly higher than boys in all five domains.

Correlations of the EDI domains with age were

2 The actual loadings are available from the first author.

Development and Psychometric Properties of the Early Development Instrument (EDI) 11

TABLE 7
Gender and Age in the EDI Scores

		SK (5-year-olds)				JK (4-year-olds)				Correlation with age ^a
		N	Mean	Standard Deviation	F ratio ^a	N	Mean	Standard Deviation	F ratio ^a	
Physical health and well-being	Boys	5,628	8.39	1.15	172.00	2,448	7.87	1.25	107.12	.231
	Girls	5,267	8.66	1.03		2,470	8.23	1.16		
Social competence	Boys	5,278	7.80	1.94	438.17	2,210	7.23	2.09	207.13	.170
	Girls	4,965	8.54	1.62		2,220	8.07	1.78		
Emotional maturity	Boys	5,590	7.41	1.65	717.49	2,434	6.93	1.69	278.65	.162
	Girls	5,253	8.20	1.38		2,457	7.68	1.44		
Language and cognitive development	Boys	5,634	7.69	2.10	187.42	2,449	5.93	2.37	102.22	.377
	Girls	5,275	8.21	1.90		2,471	6.60	2.25		
Communication and general knowledge	Boys	5,609	6.85	2.20	125.63	2,449	6.20	2.35	52.29	.192
	Girls	5,247	7.31	2.12		2,470	6.69	2.32		

^aAll statistically significant at $p < .0001$.

TABLE 8
EDI Domains by ESL and JK/SK

EDI Domain	ESL Status	JK					SK				
		N	Mean	Standard Deviation	F-stat	p-value	N	Mean	Standard Deviation	F-stat	p-value
Physical	No ESL	3,096	8.18	1.19	111.92	< 0.001	7,625	8.56	1.11	48.45	< 0.001
	ESL	1,448	7.78	1.22			2,928	8.39	1.08		
Social	No ESL	2,814	7.86	1.93	107.26	< 0.001	7,228	8.21	1.83	22.59	< 0.001
	ESL	1,275	7.18	1.98			2,745	8.01	1.79		
Emotional	No ESL	3,077	7.45	1.64	107.17	< 0.001	7,593	7.82	1.61	20.59	< 0.001
	ESL	1,440	6.93	1.46			2,914	7.67	1.45		
Language/Cognitive	No ESL	3,098	6.68	2.21	359.03	< 0.001	7,637	8.10	1.94	165.09	< 0.001
	ESL	1,448	5.33	2.30			2,930	7.54	2.20		
Communication/General Knowledge	No ESL	3,097	7.47	1.80	3,279.51	< 0.001	7,587	7.76	1.83	4,189.20	< 0.001
	ESL	1,448	4.19	1.79			2,927	5.17	1.87		

also all statistically significant (Table 7; both cohorts analyzed together), and fairly low, with exception of the Language and Cognitive domain, where the correlation reached a moderate range, as was expected. These results demonstrate the EDI's expected sensitivity to age and gender.

Differences between children with and without ESL status were also explored with an ANOVA (the gender distributions did not differ significantly between the groups either for JK or SK level). As expected, the EDI scores were lower for children for whom English was a second language, both at the 4-year and 5-year-old level. The largest discrepancies between the two groups were in the Communication Skills and General Knowledge domain (Table 8).

Discussion

The factor solution replicated the domains of school readiness found in literature (e.g., Phillips &

Love, 1995) and accounted for 63% of variance. However, two domains, covering the language, communication, and cognitive abilities, did not emerge as the a priori hypothesized categories (language with communication, and cognitive development separately). Considering the range of abilities that are supposed to contribute to each domain in the theoretical models, it was to be expected that the factor analyses would reveal more than five factors. This multifactorial structure of the domains needs to be explored further. It was crucial, however, that most, if not all, items showed clear contributions to the set of factors belonging to a particular domain. In fact, all but three of the items loaded on the expected factors. Factor analysis experts suggest removing lowest-loading items (Gorsuch, 1983); however, we made the decision to keep the seven items that loaded the lowest on the finalized version of the EDI. This decision was dictated by the need to preserve the relevance of

12 Janus and Offord

TABLE 9
Parent Interview Questions and Their Relevance to the EDI Domains

Question	Responses	Association with EDI domains
How many times has your family consulted or visited a health professional within the past year?	Number	Physical Health and Well-Being
In general, would you say that your child's health is:	1-poor, 2-fair, 3-good, 4-very good, 5-excellent	Physical Health and Well-Being
In your opinion, how physically active is your child compared to other children the same age and sex?	1-not active at all, 2-a bit active, 3-average, 4-active, 5-very active	Physical Health and Well-Being
Would you describe your child as being usually:	1-happy and interested in life, 2-some-what happy, 3-somewhat unhappy, 4-unhappy with little interest in life, 5-so unhappy that life is not worthwhile	Physical Health and Well-Being Social Competence Emotional Maturity
How would you describe his/her usual ability to think and solve day-to-day problems? Is she/he:	1-able to think clearly and solve problems, 2-having a little difficulty, 3-having some difficulty, 4-having a great deal of difficulty, 5-unable to think or solve problems	Social Competence Emotional Maturity
Apart from school, about how many days a week does she/he do things with other children?	1-never, 2-1 day/wk, 3-2 to 3 days/wk, 4-4 to 5 days/wk, 5-6 to 7 days/wk,	Social Competence Emotional Maturity
During the past six months, how well has your child been getting along with other children such as friends or classmates (excluding siblings)?	1-very well, no problems, 2-quite well, hardly any problems, 3-pretty well, occasional problems, 4-not well at all, constant problems	
Since starting school in the fall, how well has she/he been getting along with her/his teachers?	As above	Social Competence Emotional Maturity
During the past six months, how well has she/he been getting along with her/her parents?	As above	Social Competence Emotional Maturity
With regard to how she/he feels about school, how often does she/he look forward to going to school?	1-almost never, 2-rarely, 3-sometimes, 4-often, 5-almost always	Social Competence Emotional Maturity
How often does your child look at books, comics, magazines, etc. on his/her own?	1-rarely, 2-less than once a month, 3-once a month, 4-a few times a months, 5-once a week, 6-a few times a week, 7-daily, 8-many times each day	Social Competence Emotional Maturity Language and Cognitive Development Communication Skills and General Knowledge
How often does she/he play with markers or pencils doing real or pretend writing?	As above	Language and Cognitive Development Communication Skills and General Knowledge
Currently, how often do you or another adult read aloud to him/her or listen to him/her read or attempt to read aloud?	As above	Language and Cognitive Development Communication Skills and General Knowledge
How old was she/he when you started (to the nearest month of age)?	Age in months	Language and Cognitive Development Communication Skills and General Knowledge

the questionnaire and its coverage to the community of teachers and educators.

If the data are correlated and clustered, as they are in the present study, the factor analyses of between-classroom data and within-classroom data could show different results. While the factor analyses methodology used accounted for the correlated factors, it did not account for clustering. The four-step Muthen procedure (Muthen, 1994) enables us to detect differences between the two factor structures, if there are any. The finding that there was very little difference between the fit coefficients in the two models allows us to say that for each domain, the EDI factor structure between classrooms is similar to the factor structure within classrooms. Because the clustering within classroom is an unavoidable natural phenomenon that is replicated when the EDI is used in the communities, it is important to assess the differences between levels of analyses. Muthen's multilevel confirmatory factor analyses methodology is suggested as the most adequate to test whether the structure of a construct differs across levels of analyses (Dyer, Hanges, & Hall, 2005). We chose this method of testing the impact of clustering, rather than a random selection of a student per classroom, or averaging results per classroom, because 1) neither of the other two would account for the variability occurring within students, 2) there are arguments in literature suggesting that factor-analysis of means can produce misleading results (Dyer et al., 2005).

The consistency of teachers' ratings was explored with the ICCs. The low ICCs indicated that the majority of variance among the item and domain scores was due to the variability of children within classrooms rather than between classrooms. The high average teacher reliability for each domain indicated that despite the fact that one teacher contributed the scores for children in the class, their ratings for individual children were sufficiently different to warrant the claim that the data were reliable at the individual level.

The internal consistency of finalized scales was acceptable. Convergent validity, as shown by associations of EDI scores with age and gender, was acceptable, though it requires further investigation with another sample to allow for inclusion of socioeconomic variables.

The magnitude of differences between boys and girls was especially large in the social and emotional domains, where 5-year-old boys (SK group) scored on average lower than 4-year-old girls (JK group). This appears to be a consistent difference between boys and girls, also found in other populations (e.g., Zill,

1999). This gap has been shown to persist into later years of school (Herbert & Stipek, 2005; Sheehan, Cryan, Wiechel, & Bandy, 1991). Clearly, it is an important educational issue, and as such it is receiving attention of practitioners (e.g., Spence, 2005). Age also has an impact on the EDI scores: In four domains, for a year increase in age, the scores increased on average by 0.5 points. In Language and Cognitive Development, the scores increased by almost two points between 4- and 5-year-olds. Children with an ESL status had lower scores than children for whom English was the first language. As the school's instruction language is English, it is not surprising that children with worse command of English have difficulties (Schwartz & Stiefel, 2006). ESL learners routinely struggle with acquiring the competence in the language of instruction (Roessingh & Kover, 2003). Combined with gender and age differences, these suggest that the composition of kindergarten classes is an important factor to be considered in planning educational activities.

Study 2

Method

Participants. Teachers in 10 schools in two large urban settings sent a letter describing the study to all parents of Senior Kindergarten children (that is, children who have their fifth birthday in the year of entry to school). Of the 117 letters sent out, 100 were returned (85%) with parental agreement to participate. Unfortunately, due to circumstances beyond the control of the research team, only 85 of the 100 could be contacted. For 82 families, complete data were collected from both parent and teacher. Fifty-three children in seven schools attended kindergarten at school (half-time), and a kindergarten-age program at a child-care centre (half-time).

Measures. The EDI was completed by school teachers and parents for all 82 children, and by child-care teachers for 53 children. Children's receptive vocabulary was directly assessed with the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981). The PPVT was administered to children within less than two weeks from the teacher completing the EDI.

Parents were interviewed to provide family background information, including parent education and marital status. They also answered additional questions about the child's health and behaviour. These were used to establish the external validity of the teacher-completed EDI. The questions, answer options, and coding, and their relevance to EDI domains are listed in Table 9.

TABLE 10
Reliability and Validity Data on the EDI (Pearson Correlations)

	Physical Health & Well-Being	Social Competence	Emotional Maturity	Language and Cognitive Development	Communication and General Knowledge
School-child care teacher correlations ($N = 53$)	.69	.80	.77	.72	.53
p	.00	.00	.00	.00	.00
Parent-teacher correlations ($N = 82$)	.36	.50	.36	.64	.41
p	.00	.00	.00	.00	.00
Correlations with PPVT language assessment ($N = 82$)	.05	.01	.13	.31	.47
p	NS	NS	NS	.00	.00

TABLE 11
Associations Between Teacher EDI Ratings on Social Competence and Emotional Maturity and Parent Ratings Related to Child Socio-Emotional Competence ($N = 82$)

		Happiness	Thinking/ Problem solving	Frequency of playing with children	Getting along with children at school	Getting along with teacher at school	Getting along with parents	Looking forward to school
Social competence	Pearson r	.36	.31	.39	.11	.48	.33	.31
	p	.00	.01	.00	.35	.00	.00	.01
Emotional maturity	Pearson r	.25	.30	.27	.14	.42	.21	.24
	p	.03	.01	.01	.20	.00	.06	.03

TABLE 12
Associations Between Teacher EDI Ratings on Language and Cognitive Development and Communication Skills and General Knowledge and Parent Ratings Related to Child Cognitive Competence ($N = 82$)

		Interest in books	Interest in writing	Frequency of reading with child	Age started being read to
Language and cognitive development	Pearson r	.24	.22	.25	.20
	p	.03	.05	.02	.08
Communication and general knowledge	Pearson r	.16	.21	.16	.26
	p	.15	.07	.14	.02

Results

Interrater reliability. In order to investigate the level of agreement between two independent observers completing the EDI, the EDI ratings were compared between school kindergarten teachers and early childhood educators (ECE), and between school teachers and parents.

The correlations between teachers and ECE ranged from 0.53 to 0.8 (Table 10), and all were statistically significant. Correlations between parent and teacher ratings ranged from 0.36 to 0.64 (average of 0.45) and

all were statistically significant (Table 10). The lowest agreement between parents and teachers occurred in Physical Health and Well-Being and Emotional Maturity; the highest (0.64) in the domain of Language and Cognitive Development.

Concurrent test-criterion relationship. Correlations of the EDI language-related domains, Language and Cognitive Development scale and the Communication Skills, with PPVT scores were statistically significant, though low to moderate (0.31 and

0.47, respectively, Table 10). These associations provide the evidence for test-criterion validity for the two domains (Joint Committee on Standards for Educational and Psychological Testing, 1999), that is, that these two different measures, purportedly measuring the same concept, indeed do so. PPVT scores were not correlated with the remaining three EDI domains.

Association with parent interviews. Parent-rated aspects of child health and behaviour (listed in Table 9) were correlated with teacher ratings on relevant EDI domains.

Of the four parent-based variables relevant to Physical Health and Well-Being (Items 1-4 in Table 9), only the correlation of the parent rating of the child's overall health was statistically significant ($r = 0.34$, $p < 0.05$).

Six of the seven parent-based variables relevant to Social Competence and Emotional Maturity domains were statistically significantly correlated with teacher ratings on the EDI (Table 11).

Teacher ratings of the child in Language and Communication domains of the EDI were significantly correlated with three out of four parent-based items – interest in books, writing, and frequency of reading with adult – while the Communications Skills score was significantly correlated with one out of four – the age at which the child was first being read to (Table 12). All correlations, however, were in the expected direction.

It is also important to note that there were only three statistically significant correlations between a parent-based variable and EDI domain not directly relevant to this variable. These were: Language and Cognitive Development with frequency of seeing other children, $r = 0.26$, $p = 0.002$, and liking school, $r = 0.30$, $p = 0.009$, and Communication Skills and General Knowledge with the ability to think and solve problems, $r = 0.33$, $p = 0.003$. None of the parent-based items not relevant to either Physical Health and Well-Being, or Social Competence, or Emotional Maturity were correlated with teacher ratings in these domains.

Discussion

Interrater agreements on the EDI domains were moderate to high for the two teacher ratings, and low to moderate for parent-teacher ratings. Agreement between multiple respondents on children's behaviour is notoriously low (e.g., Boyle et al., 1996; Verhulst & Akkerhuis, 1989; Culp, Howell, Culp, & Blankemeyer, 2001; Winsler & Wallace, 2002). In particular, teachers and parents appear to have low

agreement rates, although there is a fairly high rate of agreement between parents (Grietens et al., 2004). It has been argued that respondents hold differing thresholds and standards (Grietens et al., 2004), resulting in low agreement. Low concordance could also be attributed to unique variance contributing to the ratings (Dishion, French, & Patterson, 1995): Schools may elicit different behaviour patterns in children than do home settings. Moreover, some behaviours, especially problem behaviours, have low frequency or low visibility (Campbell, 2002; Deng, Liu, & Roosa, 2004), which makes them hard to notice reliably. All of these possibilities are likely reflected in the interrater agreements on particular domains of the EDI. First, agreements between the two teachers are higher than between the teacher and parent. This suggests that 1) children behave similarly in educational settings, but differently at home, in particular in terms of their emotional expressions, and 2) school teachers and teachers in early childhood educational settings have similar perspective in assessing children's behaviour. This second finding is especially important since the results of the EDI are frequently aggregated across different teachers and this basic trust in teachers' reliability is crucial. Also, these similarities indicate that the concepts captured by the EDI are clear and easily assessed by trained educators. Second, low parent-teacher agreement ($r = 0.36$) in the Emotional Maturity domain may well reflect the low – and variable – frequency of problem behaviours, especially internalizing ones like anxiety (Cambell, 2002), which are part of that domain. Similar results were found in research on the reliability of the Strength and Difficulties Questionnaire (SDQ; Goodman, 2001). Two scales of the SDQ, Emotional Symptoms and Prosocial Behaviour, which are conceptually close to the Emotional Maturity domain of the EDI, had the lowest parent-teacher agreement rates (0.27 and 0.25, respectively) in a community sample of over 7,000 children.

A fairly high parent-teacher agreement was achieved for the Language and Cognitive Development domain, which includes letter knowledge, number knowledge, memory, and basic reading and writing skills. This agreement is higher than expected based on the evidence from largely behaviour-based scales (see above). However, parent ratings have not been commonly used to evaluate children's cognitive ability, in particular for school-age children. There is some evidence that parents tend to overestimate their children's development (Deimann, 2005; Glascoe & Sandler, 1995). Maternal predictions of their 4-year-olds' performance on 96 test items

were highly correlated with the children's actual performance, yet the "errors" in judgment were mostly overestimates (Hunt & Paraskevopoulos, 1980). A study of kindergarten-age children with developmental disabilities demonstrated that parent and teacher ratings of children's language development were positively and significantly correlated (Sigafoos & Pennell, 1995), in particular in the area of expressive language. Moreover, maternal education contributes to the accuracy in assessment of their children's abilities (Hunt & Paraskevopoulos, 1980). Almost 80% of mothers in our study were well educated, which most likely contributed to their knowledge about their children's cognitive abilities.

The language-related EDI domains were significantly associated with directly tested children's receptive vocabulary. Since receptive vocabulary is a part of the larger assessment of children's IQ and correlates well with composite IQ measures (Dunn & Dunn, 1981), PPVT scores are often taken as a proxy of a child's intelligence. Significant correlations with the appropriate EDI domains indicate good criterion validity on these domains. Nevertheless, further evidence is needed to ascertain that other areas of child cognitive development (number concepts, problem-solving, expression, memory) are accurately reflected in EDI scores. At the same time, the lack of correlations between PPVT and the three remaining EDI domains clearly provides evidence of the discriminant validity of the EDI domains.

The patterns of correlations between parent-based variables and relevant teacher-reported EDI domains further indicate that the domains discriminate among the aspects of school readiness. Although the magnitude of correlations was, at best, moderate (0.24-0.48), they were all in the expected direction. Unlike the correlations between parent ratings on the EDI and teacher ratings on the EDI, where the same questions were asked of different observers, these parent variables are based on interview questions in general areas relevant to the specific EDI domains. The correlations suggest that there is a certain small (0.06-0.23) amount of shared variance among the variables. Because most parent interview variables were based on much narrower concepts than the EDI domains, the low-level associations are not surprising.

A case in the above point is provided by parents' judgment of their children getting along with school friends, which was the only explored aspect of the socio-emotional skills not significantly correlated with teacher EDI ($r_s = 0.11$ and 0.14). Social Competence and Emotional Maturity domains each cover a spectrum of related concepts, not only the "getting along" or "prosocial" behaviours. It is possi-

ble that the lack of power was caused by associating a single aspect of a spectrum with an EDI domain combining many components. Moreover, as argued by Dishion et al. (1995), school context influences child behaviours, which may differ from those observed at home. Parents rarely have a chance to observe their child in an environment with 20 peers, rather than just one or two, and therefore their perception may not be the same as that of the teacher.

General Discussion

The Early Development Instrument was designed to fill the gap in the population-level measurement of children's school readiness with a tool that is feasible and quick to complete, informative, and psychometrically adequate, while at the same time lending itself well to aggregation for social reporting. The analyses in this paper suggest that the EDI's psychometric properties are acceptable and comparable with other instruments measuring children's behaviour (e.g., CBCL; Achenbach, 1991) and academic skills (e.g., PKRT; Duncan & Rafter, 2005).

Internal consistency of the EDI scales ranged from 0.84 to 0.96; the 14-factor solution replicated the domains of school readiness suggested in literature (Phillips & Love, 1995), and accounted for 63% of variance. The interrater reliability correlations were moderate (0.53) to high (0.80). While not reported here (Duku & Janus, 2004), the test-retest correlations were also high (0.82-0.94). Validity investigations encompassed several analyses. Parent-teacher agreements on the EDI were moderate (0.36-0.64). Concurrent test-criterion validity of the EDI, as explored in comparisons with direct language test and parent interview about children's behaviour demonstrated low to moderate, yet consistent, relationships.

The age and gender difference patterns demonstrated in other large samples of kindergarten children were also replicated by the EDI results. Zill (1999) found that boys and children with birthdays late in the year were more likely to have problems in kindergarten; male gender, and younger age at school entry significantly contributed to "school unreadiness" in Farkas and Hibbel's (2005) analysis of the ECLS-K data in the U.S. Interestingly, among kindergarten children in the sample analyzed by Farkas and Hibbel, boys were significantly older at entry than girls, a finding interpreted by the authors as a possible "strategizing" effort by parents. In jurisdictions where rules about the age of school entry are less uniformly observed, the EDI scores need to be grouped by actual age intervals rather than by "grade level." This procedure is currently being used in

Australia (Goldfeld et al., 2006). The EDI scores were consistently lower for children with the ESL status. Lack of proficiency in the language of instruction frequently contributes to children's lower achievement in school (Fontaine, Torre, & Grafwallner, 2006). There is evidence, however, that foreign-born ESL children have better academic achievement than native-born children (Schwartz & Stiefel, 2006). Moreover, as Bialystok points out, the quality of home environment and its promotion of reading and learning will have an impact on the school achievement of children with the ESL status (Bialystok, 2001). Unfortunately, in our study we were not able to control for either of these factors, and therefore this issue has to be explored further.

Associations with various other measures were usually only statistically significant where there was a strong theoretical basis for them to be so. For example, direct language tests were not significantly correlated with the noncognitive EDI domains; parent ratings of child getting along with friends *at school* was not significantly correlated to social and emotional competence rated by the teacher on the EDI, and the cognitively oriented EDI domains were not, as a rule, correlated significantly with parent ratings of children's social competence. These findings emphasize the discriminatory character of the instrument, and underline our view of reporting on each domain separately, rather than producing a composite total score, which could obscure real differences.

Unlike the many existing assessments of school readiness, the EDI has not been validated for screening at an individual or diagnostic level. In contrast to an instrument like the Child Behaviour Checklist (CBCL; Achenbach, 1991), for example, which has set thresholds indicating clinical diagnoses, an EDI score in a certain range cannot be taken as indicative of a clinical problem. However, even the CBCL author warns of equating the CBCL scores with particular disorders, and instead recommends integrating the CBCL "descriptions of the child" with other types of data on the child and family in order to arrive at a diagnosis (Achenbach, 1991). Measurement experts suggest that a test used for decision-making at an individual level needs to be more reliable than one used for group-level analysis and research (Streiner & Norman, 1995). Establishing a diagnostic use for the EDI would considerably increase its costs, and thus the availability for population-level use. With the exception of clinical identification, the EDI psychometric properties described here are at similar levels as those of other teacher questionnaires used for assessment of behaviour of preschool and early school-age children (e.g., Bulotsky-Shearer &

Fantuzzo, 2004; Goodman, 2001; Lutz, Fantuzzo, & McDermott, 2002). Together with a moderate predictive validity of the EDI from kindergarten to third grade (Gaskin, Duku, & Janus, 2005), also comparable with other measures (LaParo & Pianta, 2000), these properties suggest that the EDI could be a useful addition to the spectrum of measures available to students of children's behaviour and school adjustment in the preschool and early school years.

The major advantage of the EDI is its combination of several domains of child development into one comprehensive instrument, which sets it apart from the other available measures of school readiness. Questions are based on behaviours and skills easily observable in a school setting, and responses are rated based on observed frequency of behaviours or presence of skills, rather than on the child's performance in relation to a specific group (e.g., "top half of the class"). These properties make teachers experts in providing the information on children without the necessity of additional training. On the other hand, teacher ratings could be subject to individual bias, due to characteristics of teacher, child, school, or interactions of all three (Pianta, Steinberg, & Rollins, 1995). Although it is impossible to fully address the question of teacher bias with the data from studies reported in this paper, two findings raise the confidence in teachers' fairly uniform standards of answers: interrater reliabilities, with both other teachers and parents, and the high teacher consistencies. Elsewhere, teacher ratings of overall summary skills were reported to have only moderate association with later outcomes (Mashburn & Henry, 2004; Meisels, Bickel, Nicholson, Xue, & Atkins-Burnett, 2001). Nevertheless, a recent study suggests that the population context should be taken into account in assessing the appropriateness of the methodology used: Teacher ratings, while not specific enough to warrant early identification, are valid enough to suggest intervention models (Crooks & Peters, 2005). The teacher measures used in the cited studies all contained less than 20 items of varying generality. It appears that although the EDI is longer, it may be a compromise between multiple, costly, standardized assessments and brief rating scales, as it provides anchored teacher ratings of detailed competencies.

Several limitations have to be noted here. One of the studies had a small and moderately variable sample. In particular, very few of the children were non-English speakers. Although parent country of birth does not have a significant impact on children EDI scores (Janus & Duku, 2006; Janus, Offord, & Walsh, 2001), children for whom language of instruction is a second language face obvious disadvantages enter-

ing the school system. Second, only limited data on families' socioeconomic status were available. Unfortunately, in this respect the validation of the EDI in this study is very similar to validation of other instruments, which often have been criticized for small samples. These limitations are addressed in the next study (Janus & Duku, 2006). Moreover, as the EDI is currently used in many communities, local researchers are encouraged to include some validation components in their projects.

Throughout the process of the EDI development the engagement of representatives of the communities of stakeholders – teachers, early childhood educators, child health professionals, and parents – was highly valued. Their engagement helped to achieve an instrument relevant to the community. On the other hand, the emphasis that we put on ensuring this participation may have somewhat impeded the achievement of psychometric elegance, for example, in keeping some items due to their highly perceived importance by teachers despite their low factor loading, or having an inconsistent number of answer options in domains. Although it may be questioned, we believed that this approach was crucial for communities' perceived ownership of the measure. Developing the EDI with the stakeholders' input appears to have been a sound decision, as interest in measurement and improvement of children's school readiness in many communities in Canada has continued to be widespread.

In conclusion, while clearly some of the analyses have to be repeated with larger or different samples, the EDI's psychometric properties have proven to be acceptable. The EDI's simplicity, ease of use, and low cost all lend themselves easily to community-wide implementation. Data collected for whole populations of children have the advantage of giving the community the true picture and, especially in conjunction with other locally relevant data, allow for making useful recommendations (Janus et al., 2002) and provide a baseline for future assessments of progress. It is, however, important to point out that while the EDI already has a demonstrated ability to distinguish among communities (Janus et al., 2002), and is related to neighbourhood collective efficacy (Seward & Loomis, 2005), its sensitivity to changes in communities due to reform still remains to be investigated.

The only types of individual measurements that have been commonly aggregated to represent population-level results are school-based achievement tests (e.g., Willms & Jacobsen, 1990). This is usually done for later grades, and rarely, if ever, to provide communities with a useful overview for implementing action. Yet, the population-level analyses allow

identification of inequalities related to characteristics that may be remediable by appropriate policies (Starfield, 2002). The EDI is the first tool for which community aggregation is feasible due to its low cost, relevance in covering the five developmental domains, and proven psychometric properties.

The authors would like to acknowledge The Founder's Network, Invest in Kids Foundation, and Human Resources Development Canada for their financial support. We would like to thank teachers and parents in Toronto and Calgary for volunteering their time, and Eric Duku for his invaluable statistical expertise.

Correspondence concerning this article should be addressed to Dr. Magdalena Janus, Offord Centre for Child Studies, McMaster University, 1200 Main St. W., Patterson Bldg. Rm. 215, Hamilton, Ontario, Canada L8N 3Z5.

References

- Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont.
- Alexander, K. L., & Entwisle, D. R. (1988). Achievement in the first 2 years of school: Patterns and processes. *Monographs of the Society for Research in Child Development, 53*(2), 1-157.
- Augustyniak, K. M., Cook-Cottone, C. P., & Calabrese, N. (2004). The predictive validity of the Phelps Kindergarten Readiness Scale. *Psychology in the Schools, 41*, 509-516.
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York: Cambridge University Press.
- Boyle, M. H., & Lipman, E. L. (2002). Do places matter? Socioeconomic disadvantage and behavioral problems of children in Canada. *Journal of Consulting and Clinical Psychology, 70*(2), 378-389.
- Boyle, M. H., Offord, D. R., Racine, Y. A., Szatmari, P., Sanford, M., & Fleming, J. E. (1996). Interviews versus checklists: Adequacy for classifying childhood psychiatric disorder based on adolescent reports. *International Journal of Methods in Psychiatric Research, 6*, 309-319.
- Brigance, A. (1992). *Brigance Diagnostic Inventory of Early Development-Revised*. North Billerica, MA: Curriculum Associates.
- Brooks-Gunn, G. J., Guo, G., & Furstenberg, F. F. (1993). Who drops out and who continues beyond high school? A twenty-year follow-up of black urban youth. *Journal of Adolescence, 3*, 271-294.
- Bulotsky-Shearer, R., & Fantuzzo, J. (2004). Adjustment scales for preschool intervention: Extending validity and relevance across multiple perspectives. *Psychology*

- in the Schools*, 4(7), 725-735.
- Butaine, R., & Costenbader, V. (1997). The effectiveness of a transitional prekindergarten program on later academic achievement. *Psychology in the Schools*, 34(1), 41-50.
- Campbell, S. B. (2002). *Behavior problems in preschool children: Clinical and developmental issues*. New York: Guilford Press.
- Chase-Lansdale, P. L., Gordon, R. A., Brooks-Gunn, J., & Klebanov, P. K. (1997). Neighborhood and family influences on the intellectual and behavioral competence of pre-school and early school-age children. In G. J. Brooks-Gunn & J. L. Duncan (Eds.), *Neighborhood poverty* (pp. 79-118). New York: Russell Sage Foundation.
- Chew, A. L., & Lang, W. S. (1990). Predicting academic achievement in kindergarten and first grade from prekindergarten scores on the Lollipop Test and Dial. *Educational and Psychological Measurement*, 50, 431-437.
- Chew, A. L., & Morris, J. D. (1989). Predicting later academic achievement from kindergarten scores on the Metropolitan Readiness Tests and the Lollipop Test. *Educational and Psychological Measurement*, 49, 461-465.
- Costenbader, V., Rohrer, A. M., & DiFonzo, N. (2000). Kindergarten screening: A survey of current practice. *Psychology in the Schools*, 37(4), 323-332.
- Crooks, C. V., & Peters, R. D. (2005). Predicting academic difficulties: Does a complex, multidimensional model outperform a unidimensional teacher rating scale? *Canadian Journal of Behavioural Science*, 37(3), 170-180.
- Culp, R. E., Howell, C. S., Culp, A. M., & Blankemeyer, M. (2001). Maltreated children's emotional and behavioral problems: Do teachers and parents see the same things? *Journal of Child and Family Studies*, 10(1), 39-50.
- Deimann, P. (2005). Mothers' evaluation of their kindergarten children's developmental status. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37, 122-134.
- Deng, S., Liu, X., & Roosa, M. W. (2004). Agreement between parent and teacher reports on behavioral problems among Chinese children. *Developmental and Behavioral Pediatrics*, 25(6), 407-413.
- Dishion, T. J., French, D. C., & Patterson, G. R. (1995). The development and ecology of antisocial behavior. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (pp. 421-471). New York: Wiley.
- Doherty, G. (1997). *Zero to six: The basis for school readiness*. Ottawa, ON: Human Resources Development Canada. R-97-8E.
- Duku, E., & Janus, M. (2004, April). *Stability and reliability of the early development instrument: A population-based measure for communities (EDI)*. Poster presentation, 16th Annual Research Day, Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON.
- Duncan, J., & Rafter, E. M. (2005). Concurrent and predictive validity of the Phelps Kindergarten Readiness Scale-II. *Psychology in the Schools*, 42, 355-359.
- Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised. Manual for Forms L and M*. Circle Pines, MN: American Guidance Service.
- Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly*, 16, 149-167.
- Entwisle, D. R., & Alexander, K. L. (1999). Early schooling and social stratification. In R. C. Pianta & M. J. Cox (Eds.), *The transition to kindergarten* (pp. 13-39). Baltimore, MD: Paul H. Brookes.
- Farkas, G., & Hibel, J. (2005, October). *Being unready for school: Factors affecting risk and resilience*. Paper presented at the National Symposium on Family Issues, Pennsylvania State University.
- Fontaine, N. S., Torre, D. L., & Grafwallner, R. (2006). Effects of quality early care on school readiness skills of children at risk. *Early Child Development and Care*, 176(1), 99-109.
- Gaskin, A., Janus, M., & Duku, E. (2005, April). *Concurrent and predictive validity of the Early Development Instrument (EDI)*. Poster presentation, 17th Annual Research Day, Department of Psychiatry & Behavioural Neurosciences, McMaster University, Hamilton, ON.
- Glascoe, F. P., & Sandler, H. (1995). Value of parents' estimates of children's developmental ages. *The Journal of Pediatrics*, 127, 831-835.
- Goldfeld, S., Brinkman, S., Sayers, M., Liddelow, T., Daws, A., & Coutts, M., et al. (2006, April). *A national approach to measuring child development in Australia: Process challenges and policy implications*. Paper presented at the Measuring Early Child Development Conference, Vaudreuil, QC.
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(11), 1337-1345.
- Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, NJ: Erlbaum.
- Graue, M. E., & Shepard, L. A. (1989). Predictive validity of the Gesell School Readiness Tests. *Early Childhood Research Quarterly*, 4, 303-315.
- Grietens, H., Onghena, P., Prinzie, P., Gadeyne, E., VanAssche, V., Ghesquiere, P., et al. (2004). Comparison of mothers', fathers', and teachers' reports on problem behavior in 5- to 6-year-old children. *Journal of Psychopathology and Behavioral Assessment*, 26(2), 137-145.
- Herbert, J., & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic

- competence. *Journal of Applied Developmental Psychology*, 26(3), 276-295.
- Hunt, J. M., & Paraskevopoulos, J. (1980). Children's psychological development as a function of the inaccuracy of their mothers' knowledge of their abilities. *Journal of Genetic Psychology*, 136(2), 285-298.
- Ilg, F., Ames, L., Haines, J., & Gillespie, C. (1978). *Gesell School Readiness Test*. New York: Harper & Row.
- Jacob, S., Snider, K., & Wilson, J. (1988). Validity of the DIAL-R for identifying children with special education needs and predicting early reading achievement. *Journal of Psychoeducational Assessment*, 6, 289-297.
- Janus, M., & Duku, E. (2006). *The school entry gap: Socioeconomic, family, and health factors associated with children's school readiness to learn*. Unpublished manuscript.
- Janus, M., & Offord, D. (2000). Readiness to learn at school. *ISUMA*, 1(2), 71-75.
- Janus, M., Offord, D., & Walsh, C. L. (2001, April). *Population-level assessment of readiness to learn at school in 5-year-olds: Relationship to child and parent measures*. Proceedings of the 2001 Biennial Meeting, Society for Research in Child Development, Minneapolis, MN.
- Janus, M., Walsh, C., Viveiros, H., & Offord, D. R. (2002, June). *Community, neighbourhood and 5-year-olds' readiness to learn at school*. Poster session presented at the Head Start 6th National Conference, Washington, DC.
- Jimerson, S., Egeland, B., & Teo, A. (1999). A longitudinal study of achievement trajectories: Factors associated with change. *Journal of Educational Psychology*, 91(1), 116-126.
- Joint Committee on Standards for Educational and Psychological Testing. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Kagan, S. L. (1992). Readiness past, present, and future: Shaping the agenda. *Young Children*, 48(1), 48-53.
- Kagan, S. L., & Neuman, M. J. (1997). Defining and implementing school readiness: Challenges for families, early care and education, and schools. In R. P. Weissberg, T. P. Gullotta, R. L. Hampton, B. A. Ryan, & G. R. Adams (Eds.), *Establishing preventive services* (pp. 61-96). Thousand Oaks, CA: Sage Publications.
- Kohen, D. E., Brooks-Gunn, J., Leventhal, T., & Hertzman, C. (2002). Neighborhood income and physical and social disorder in Canada: Associations with young children's competencies. *Child Development*, 73(6), 1844-1860.
- La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research*, 70, 443-484.
- Liberal Party of Canada. (2005, March) *Supporting Canadian Families, Liberal Party of Canada Biennial Convention Paper*, Ottawa, ON: Liberal Party of Canada.
- Lichtenstein, R. (1990). Psychometric characteristics and appropriate use of the Gesell School Readiness Screening Test. *Early Childhood Research Quarterly*, 5, 359-378.
- Love, J. M., Aber, J. L., & Brooks-Gunn, J. (1994). *Strategies for assessing community progress toward achieving the first national educational goal*. Princeton, NJ: Mathematica Policy Research, Inc.
- Lutz, M. N., Fantuzzo, J., & McDermott, P. (2002). Multidimensional assessment of emotional and behavioral adjustment problems of low-income preschool children: Development and initial validation. *Early Childhood Research Quarterly*, 17(3), 338-355.
- Mardell-Czudnowski, C., & Goldberg, D. S. (1998). *Developmental indicators for assessment of learning (DIAL-3)* (3rd ed.). Circle Pines, MN: American Guidance Service.
- Mashburn, A. J., & Henry, G. T. (2004). Assessing school readiness: Validity and bias in preschool and kindergarten teachers' ratings. *Educational Measurement: Issues and Practice*, 23(4), 16-30.
- May, D. C., & Kundert, D. K. (1992). Kindergarten screening in New York State: Tests, purposes and recommendations. *Psychology in the Schools*, 29, 35-41.
- McCain, M. N., & Mustard, J. F. (1999). *Report on the early years study*. Toronto, ON: Province of Ontario, Ontario Children's Secretariat. ISBN 0-7778-8583-2.
- Meisels, S. J. (1987). Uses and abuses of developmental screening and school readiness testing. *Young Children*, 42(4), 6-68.
- Meisels, S. J. (1988). *Assessing readiness*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement. Ciera Report #3-002.
- Meisels, S. J. (1999). Assessing readiness. In R. C. Pianta & M. J. Cox (Eds.), *The transition to kindergarten* (pp. 39-66). Baltimore, MD: Paul H. Brookes Publishing Co.
- Meisels, S. J., Bickel, D. D., Nicholson, J., Xue, Y., & Atkins-Burnett, S. (2001). Trusting teachers' judgments: A validity study of a curriculum-embedded performance assessment in kindergarten to grade 3. *American Educational Research Journal*, 38(1), 73-95.
- Morongiello, B. (1997). *Tapping school readiness in the NLSCY: Measurement issues and solutions*. Ottawa, ON: Applied Research Branch, Strategic Policy, Human Resources Development Canada. T-98-1E.
- Muthen, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods and Research*, 22, 376-398.
- Muthen, B. O., & Muthen, L. K. (2004). *Mplus version 3.0*. Los Angeles, CA: Muthen & Muthen.
- NLSC Project Team. (1995). *National longitudinal survey of children: Overview of survey instruments for 1994-95 data collection cycle 1* (pp. 95-102). Ottawa, ON: Statistics

- Canada Human Resources Development Canada.
- Offord, D. R., Kraemer, H. C., Kazdin, A. E., Jensen, P. S., & Harrington, R. (1998). Lowering the burden of suffering from child psychiatric disorder: Trade-offs among clinical, targeted, and universal interventions. *Journal of American Academy of Child and Adolescent Psychiatry, 37*, 686-94.
- Offord, D. R., Kraemer, H. C., Kazdin, A. E., Jensen, P. S., Harrington, R., & Gardner, J. S. (1999). Lowering the burden of suffering: Monitoring the benefits of clinical, targeted, and universal approaches. In D. Keating & C. Hertzman (Eds.), *Developmental health and the wealth of nations*. (pp. 293-310). New York: The Guilford Press.
- Offord, D. R., & Lipman, E. L. (1996). Emotional and behavioural problems. In Anonymous, *Growing up in Canada: National longitudinal survey of children and youth* (pp. 119-126). Ottawa, ON: Human Resources Development Canada.
- Phillips, D., & Love, J. (1995, May). *Indicators for school readiness, schooling, and child care in early to middle childhood*. Washington, DC: Institute for Research on Poverty. Special Report No. 60b.
- Pianta, R. C., & McCoy, S. J. (1997). The first day of school: The predictive validity of early school screening. *Journal of Applied Developmental Psychology, 18*, 1-22.
- Pianta, R. C., Steinberg, M. S., & Rollins, K. B. (1995). The first two years of school: Teacher-child relationships and deflections in children's classroom adjustment. *Development and Psychopathology, 7*, 295-312.
- Rimm-Kaufman, S. E., Pianta, R. C., & Cox, M. J. (2000). Teachers' judgments of problems in the transition to kindergarten. *Early Childhood Research Quarterly, 15*, 147-165.
- Roessingh, H., Kover, P. (2003). Variability of learners' acquisition of cognitive academic language proficiency: What can we learn from achievement measures? *TESL Canada Journal, 21*(1), 1-21.
- Rose, G. (1994). *The strategy of preventive medicine*. Oxford: Oxford University Press.
- Schonkoff, J. P., & Phillips, D. A. (2000). *From neurons to neighborhoods*. Washington, DC: National Academy Press.
- Schwartz, A. E., & Stiefel, L. (2006). Is there a nativity gap? New evidence on the academic performance of immigrant students. *Education Finance and Policy, 1*(1), 17-49.
- Seward, C. L., & Loomis, C. (2005, April) *Rural neighbourhood collective efficacy influences children's readiness to learn*. Poster session, Biennial Meeting, Society for Research in Child Development Meeting, Atlanta, GA.
- Sheehan, R., Cryan, J. R., Wiechel, J., & Bandy, I. G. (1991). Factors contributing to success in elementary schools: Research findings for early childhood educators. *Journal of Research in Childhood Education, 6*(1), 66-74.
- Shore, R. (1997). *Rethinking the brain: New insights into early development*. New York: Families and Work Institute.
- Sigafoos, J., & Pennell, D. (1995). Parent and teacher assessment of receptive and expressive language in preschool children with developmental disabilities. *Education and Training in Mental Retardation and Developmental Disabilities, 30*(4), 329-335.
- Spence, C. M. (2005). *Creating a literacy environment for boys: Ideas for teachers, administrators, and parents*. Toronto, ON: Thomson Nelson.
- Starfield, B. (2002). Forward to: Improving children's health: How population-based research can inform policy – The Manitoba Experience. *Canadian Journal of Public Health, 93*, S7.
- Statistics Canada (2005). *Canada at a glance 2005*. Catalogue number [12-581-XPE]. Ottawa, ON: Statistics Canada, Communications Division.
- Streiner, D. L., & Norman, G. R. (1995). *Health measurement scales: A practical guide to their development and use* (2nd ed.). New York: Oxford University Press.
- Swanson, B., Payne, D., & Jackson, B. (1981). A predictive validity study of the Metropolitan Readiness Tests and Meeting Street School Test against first grade Metropolitan Achievement Test scores. *Educational and Psychological Measurement, 41*, 575-578.
- Verhulst, F. C., & Akkerhuis, G. W. (1989). Agreement between parents' and teachers' ratings of behavioral/emotional problems of children aged 4-12. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 30*(1), 123-136.
- Wenner, G. (1995). Kindergarten screens as tools for the early identification of children at risk for remediation or grade retention. *Psychology in the Schools, 32*, 249-254.
- Willms, J. D., & Jacobsen, S. (1990). Growth in mathematics skills during the intermediate years: Sex differences and school effects. *International Journal of Educational Research, 14*, 157-174.
- Winsler, A., & Wallace, G. L. (2002). Behavior problems and social skills in preschool children: Parent-teacher agreement and relations with classroom observations. *Early Education and Development, 13*(1), 41-58.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.
- Zill, N. (1999). Promoting educational equity and excellence in kindergarten. In R. C. Pianta and M. J. Cox (Eds.), *The transition to kindergarten*. Baltimore, MD: Paul H. Brookes.
- Zill, N., Resnick, G., Kim, K., Hubbel McKey, R., Clark, C., Pai-Samant, S., et al. (2001). *Head Start FACES*:

Longitudinal finding on program performance, third progress report. Washington, DC: Research, Demonstration, and Evaluation Branch & Head Start Bureau, Administration on Children, Youth and Families, U.S. Department of Health and Human Service.

*Received June 2, 2005
Revised February 6, 2006
Accepted July 31, 2006*