The Early Development Instrument: An Examination of Convergent and Discriminant Validity

Shelley Hymel · Lucy LeMare · William McKee

Abstract   The convergent and discriminant validity of the Early Development Instrument (EDI), a teacher-rated assessment of children's "school readiness", was investigated in a multicultural sample of 267 kindergarteners (53% male). Teachers evaluations on the EDI, both overall and in five domains (physical health/well-being, social competence, emotional maturity, language/cognition, communication/general knowledge), were related to direct, child-based assessments of performance on two standardized measures of school readiness, and measures of phonological awareness and early social competence. Regression analysis indicated that together the four comparison measures accounted for 36% of variance in overall EDI scores, each making a significant and unique contribution. Results supported the convergent validity of overall EDI scores but not the discriminant validity of EDI domain scores. Moreover, correlations between EDI scores and comparison measures varied widely across teachers, suggesting considerable individual differences in teacher's ability to evaluate school readiness relative to direct, child-based assessments, and confirming that the EDI is more appropriate for deriving inferences at higher aggregated levels

S. Hymel · W. McKee
Department of Educational and Counselling Psychology and Special Education,
University of British Columbia, Vancouver, BC, Canada
e-mail: william.mckee@ubc.ca

L. LeMare
Faculty of Education, Simon Fraser University, Burnaby, BC, Canada
e-mail: lemare@sfu.ca

S. Hymel (✉)
Faculty of Education, University of British Columbia, 2125 Main Mall, Vancouver, BC V6T 1Z4, Canada
e-mail: shelley.hymel@ubc.ca

such as community or region. The validation of EDI domain scores remains an important challenge in future research.

Keywords   School readiness · Early Development Instrument · Validity

1 The Early Development Instrument: An Examination of Validity

The Early Development Instrument (EDI) is a relatively new, teacher-report assessment tool designed by Janus and Offord (2000, 2007) to measure school readiness in populations of children. Although the EDI is completed for individual children, results are intended to be aggregated at various levels (e.g., school, neighborhood, community, see Kershaw et al. 2005 as one example) to assess the school readiness of the group. Evidence in support of the measurement and psychometric quality of the EDI is growing, though some significant gaps remain (c.f. Keating 2007). These include examination of the EDI in relation to direct, child-based assessments (as opposed to other adult ratings) and other indicators of school readiness. The need for such data is highlighted by the current widespread use of the EDI in Canada (e.g., Guhn et al. 2007; Kershaw et al. 2007), Australia (e.g., Brinkman et al. 2007; Sayers et al. 2007) and around the world (see Janus 2006) and attendant questions concerning the validity of EDI scores at various levels (child, classroom, community, etc.). This paper describes research on the validity of the EDI aimed at addressing these issues.

1.1 School Readiness

Young children's readiness for school has long been of interest to academics and educators, and has increasingly garnered the attention of politicians and policy makers. Public sector interest in school readiness can be linked to increasing recognition that early child development affects well-being across the life-span (e.g., Keating and Hertzman 1999; National Research Council and Institute of Medicine 2000) and to consequent demands for universal programs that support the development of all children (e.g., see Irwin et al. 2007), as well as increasing emphases on program accountability and evidence-based practice.

Considerable debate has occurred over the conceptualization of school readiness and resultant approaches to its measurement (see Vernon-Feagans and Blair 2006), including what should be measured, who should be the informant, and whether assessment should be an individual or a group process. Traditional readiness measures (e.g., Brigance K-1 Screen—Revised [Brigance 1997; Glascoe 1997]; McCarthy Scales of Children's Abilities [McCarthy 1972; Sturner et al. 1984]; Metropolitan Readiness Test [Nurss and McGauvran 1995; Clancy and Pianta 1993]) reflect the assumptions that (1) cognitive development is paramount to school readiness, (2) standardized, direct, child-based assessment of children's skills provides the most valid indicator, and (3) assessment is an individual process to identify those who are and are not deficient.

More recent conceptualizations of school readiness extend beyond the contributions of maturation and cognitive abilities and reflect the broader notion that school readiness involves the child's ability to meet the task demands of school (e.g., Doherty and Stuart 1997; Konold and Cox 2005; Meisels 1999), including such things as being cooperative, attending to and understanding the teacher, getting along with peers, and handling materials. These and other task demands entail physical, social, emotional, and language competencies in addition to those in the cognitive domain. Concomitant with foci on

universal programming and accountability, there is also a need and desire to measure school readiness at the level of schools and communities (Connor and Brink 1999; Janus et al. 2002), owing in part to debates regarding whether "readiness" resides in the child or in the school (Graue 2006; Keating 2007). Given these trends, as well as concerns regarding the predictive validity of traditional school readiness measures (Janus and Offord 2007; Meisels 1987), these instruments may no longer be adequate. Moreover, direct, child-based assessments are time-consuming and costly, rendering them impractical for community or population-level reporting.

1.2 The Early Development Instrument

In response to the need for population-level assessments, the Early Development Instrument (EDI) was developed as a group-level index of school readiness assessing a broad range of competencies believed to contribute to children's preparedness for school learning (Janus and Offord 2000, 2007). Specifically, kindergarten teachers evaluate each student on more than 100 items that assess physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge. As a population measure, the obvious benefits of the EDI over traditional readiness assessments are that: (1) it addresses several key areas of development; (2) it is relatively short and easy to administer; and as such (3) it is relatively inexpensive to use in evaluating large numbers of children. As with any measure, however, it is incumbent on researchers to demonstrate the psychometric soundness of the EDI.

Considerable effort toward this end has been made by the developers of the EDI. For example, the internal structure and consistency of the EDI, its convergent and discriminant validity, and its inter-rater reliability have been addressed in different samples of varying sizes with mostly favorable results (see Janus and Offord 2007 for details). In addition, Duku and Janus (2004) found EDI domain scores to be quite stable over a 2–4 week interval across three small samples.

In a special issue of *Early Education and Development* (see Guhn et al. 2007) researchers other than the developers of the EDI addressed the instrument's psychometric properties. Forget-Dubois et al. (2007) explored the short-term predictive validity of the EDI, along with a battery of tests of school readiness and cognition, in the prediction of Grade 1 academic outcomes. EDI ratings explained 23% of the variance in Grade 1 achievement beyond the 13% explained by age, gender, and socioeconomic status (SES). Almost identical results were obtained in predicting Grade 1 achievement from the cognitive battery and school readiness (i.e., 23% of the variance after accounting for age, gender and SES). Two domains of the EDI, Physical Health/Well-being and Language/Cognitive Development, made a unique contribution to the variance in Grade 1 school achievement over and above the variance accounted for by the cognitive battery and measures of school readiness.

In a sample of 642 Australian, 4-5-year-olds, Brinkman et al. (2007) assessed the validity of the EDI by correlating EDI subscales with a variety of conceptually-similar measures of early learning and development collected contemporaneously, including four parent reports, two teacher reports, and two direct, child-based measures of receptive vocabulary (Peabody Picture Vocabulary, PPVT, Dunn and Dunn 1981) and writing/copying tasks. Overall, the EDI correlated reasonably well with other measures (r's ranged from .10 to .69), the strongest correlates being other teacher-reports of similar constructs. Correlations between the EDI and analogous parent-reports were small to non-existent.

Direct, child-based assessments of early language and writing skills correlated moderately (.34 and .49, respectively) with the EDI Language/Cognition domain.

Although the aforementioned studies provide support for the validity of the EDI, further work remains to be done. Establishing validity is an ongoing process (see Hubley and Zumbo 1996), requiring repeated and varied examination until a body of evidence is accumulated that provides a basis for conclusions about whether a measure is valid for a particular purpose. Keating (2007) notes that what is missing from the published literature on the psychometric properties of the EDI is a thorough examination of the validity of the instrument based on concurrent, direct, child-based assessments for individual children. The published studies in which validity has been addressed have relied primarily on teacher and parent reports as comparison measures; a practice that has a number of shortcomings. First, using teacher report measures capitalizes on shared method variance, which can yield inflated estimates of the association between variables.

A second (and related) problem concerns the potential biases inherent in parent and teacher reports. Whether consciously or not, teachers may hold assumptions about the populations of children they serve as well as individuals within those populations that will undoubtedly influence the judgments they make when completing readiness instruments. The same can be said about parents. Characteristics of the community or population being served (e.g., ethnic and gender composition, socioeconomic status, etc.) may be general-ized or may introduce halo effects that decrease sensitivity to individual differences (e.g., Chambers and Windschitl 2004). Keating (2007) reminds us that "this is one of the major reasons that 'objective' testing, either in standardized group tests or clinical assessment, was introduced in the first place, precisely as an attempt to safe-guard against such bias" (p. 565). The only way to determine the extent to which inherent bias is a problem for the EDI itself is to examine it in relation to direct, child-based measures of the same constructs the EDI purports to assess. Direct, child-based assessments have been addressed in the published validity work on the EDI, but only in the language domain and primarily with the PPVT (e.g., Brinkman et al. 2007). Given that the EDI is intended to assess five developmental domains, all believed to be important to school readiness, it behooves researchers to examine EDI scores in all domains in relation to analogous direct, child-based assessments to the degree that such assessments exist.

The aim of the present study was to extend research on the validity of the EDI by examining the convergent and discriminant validity of the EDI in a fair-sized, multicultural sample of kindergarten students by correlating EDI scores with other, established, indi-vidually-administered indices of school readiness across several domains of development. In doing so, we faced two major challenges. First, we needed measures that tapped each of the five domains assessed by the EDI and second, we needed measures that entailed the direct, child-based assessment of individual children rather than relying on adult reports. Given young children's limited cognitive and language skills and attention spans, it is not surprising that many researchers rely on adult ratings as an efficient way to evaluate child attributes. However, in order to complete a stringent test of convergent and discriminant validity that does not capitalize on shared method variance between measures, we chose not to correlate adult ratings with other adult ratings on conceptually-related items, but instead evaluated the validity of the EDI using current and well-established, direct, child-based measures that assess readiness across multiple domains at the individual level. When standardized measures were not available, research-based assessment instruments were identified. Hence, the present study adds to the body of evidence on the validity of the EDI by examining teacher-reported EDI data in relation to direct, child-based assessments of children's functioning across domains.

2 Method

2.1 Participants

Data were collected for 267 kindergarten children (142 boys, 125 girls, mean age = 5½ years, range: 4 years, 4 months to 6 years, 3 months at the completion of testing) evaluated on the EDI by 27 teachers in 16 schools (3 districts) serving neighborhoods in southwestern British Columbia that were socioeconomically and culturally/ethnically diverse. The sample was equally divided with respect to sex (53% female), was primarily non-Aboriginal (97%), English speaking (71%), non-ESL (70%) and included few children with designated special needs (1%). Teachers who completed the EDI were primarily female and Caucasian.

2.2 Procedure

Kindergarten students who received parent permission for participation were normally involved in two testing sessions of approximately 30 minutes each in their schools; for some students, testing was completed over several shorter sessions due to class schedules or limited child attention spans. Nine different testers (male and female university students) assessed the children on the comparison measures. Each received prior training in the administration of each comparison measure and achieved mastery on administration and scoring before undertaking data collection. Scoring for each measure was verified prior to data entry. The children's teachers provided EDI ratings within 3 months (either before or after) of the child assessments, with all testing completed in the spring of 2003.

2.3 Measures

The Early Development Instrument.[1] The EDI is an adult-report questionnaire on which kindergarten teachers evaluate children's development in five domains (Janus and Offord 2000, 2007). The EDI assessment of Physical Health and Well-being includes 12 questions answered on a 5-point scale that assess physical preparedness for the school day, fine and gross motor skills, energy level and physical independence. Three items about washroom independence, hand preference, and coordination are answered in a yes/no format. The 26 Social Competence items, all answered on a 3-point scale, evaluate competence and cooperation in working with others, ability to remember and follow rules, curiosity and eagerness, approaches to learning and problem solving. The 28 Emotional Maturity items, also answered on 3-point scales, address prosocial behaviour, aggression, inattention and hyperactivity, and anxious behaviors. The EDI Language and Cognitive Development domain includes 26 questions answered in a yes/no format (the child does or does not have the skill), that address the ability to use language correctly, basic literacy and numeracy skills, interest and memory, and more complex literacy skills. Finally, the Communication and General Knowledge domain includes nine items evaluating the child's ability to clearly communicate needs and thoughts, understand others, and articulate clearly, as well as

---

[1] Although a formal test administration manual was not yet available for the EDI at the time of this study, the information presented here for the instrument is based on that provided by test developers (see http://www.offordcentre.com/readiness/pubs/publications.html). For the version of the EDI used in this study, item response formats varied (Y/N, 3-point and 5-point scales) as described here. However, readers should be aware that in the current version of the EDI, item format has been modified, with 3-point, Likert scales for all items.

aspects of general knowledge, with several items scored on 5-point scales, one on a 3-point scale, and one on a 2-point scale.

For each domain, responses to relevant items were averaged to create a composite score, with higher scores indicating more positive teacher evaluations in each domain. In addition, we also created an overall EDI score for each child, computed as the average of the five domain scores. Although an overall EDI score is not typically computed (see Ford et al. 2007 for an exception), we wanted to compare overall readiness assessments across EDI and comparison measures, some of which also provide overall readiness scores.

Comparison measures. Four different measures of readiness, designed for use with kindergarten children were selected to address the convergent and discriminant validity of the EDI. Each addressed some aspect of readiness as tapped by the EDI, although none provided a complete match. Accordingly, it is important to underscore the fact that these comparison measures do not represent indices of criterion validity. Rather, they are more accurately seen as indices against which the convergent and discriminant validity of the EDI can be assessed, within the general framework of construct validity. Comparison measures included two standardized, well-established and frequently used indices of school readiness—the Early Screening Instrument-Kindergarten—Revised (ESI-K; Meisels et al. 1997) and the School Readiness Composite (SRC) of the Bracken Basic Concepts Scale-Revised (Bracken 1998). Each assesses aspects of language, cognition, communication, and general knowledge; the ESI-K also includes assessment of some physical skills.

The ESI-K provided the broadest assessment of school readiness, given its focus on a variety of cognitive and physical arenas, and was expected to yield the strongest correlation of all the comparison measures with the overall EDI score. With regard to reliability, the ESI-K has been shown to demonstrate high inter-rater reliability (.97–.99) and test–retest reliability (.68–.98 over 7–10 days) (Kimmel and Paget 2001). With regard to validity, the ESI-K has been shown to correlate highly with the McCarthy Scales of Children's Abilities (.73) (Meisels et al. 1997; Kimmel and Paget 2001). Given its content, ESI-K scores were expected to correlate more strongly with EDI scores in the Language/ Cognition and Communication/General Knowledge domains and, to a lesser extent, with the Physical Health/Well-Being domain than with the Social Competence and Emotional Maturity domains. Particularly relevant to the present study were ESI-K subscale scores that addressed more specific areas of readiness, given their overlap with domains tapped on the EDI, including the ESI-K subscales for (1) Visual Motor Skills and (2) Gross Motor Skills (both of which overlapped with EDI's Physical Domain), as well as the ESI-K subscales for (3) Visual Sequential Memory, (4) Auditory Memory, (5) Language and Cognition, (6) Number Concepts (all overlapping with the ESI-K Language/Cognition domain), and the ESI-K subscale for (7) Verbal Expression (which overlapped with the ESI-K Communication/General Knowledge domain).

The Bracken SRC is a measure of receptive language that taps children's understanding of basic concepts that teachers typically expect children to know when they enter school. According to the Mental Measurements Yearbook (McKnight and Schwarting 2004), the Bracken demonstrates strong reliability (e.g., split half reliability coefficients from .78 to .97; test retest reliability of .88) and validity (e.g., correlations with WPPSI-R in the range of .76–.88; correlations with the Differential Abilities Scale in the range of .69–.79) (see also Bracken 1998). Given its focus on language and general world knowledge, we expected scores on the Bracken SRC to be most closely related to EDI evaluations of Language/Cognition and/or Communication/General Knowledge. Subscales from the Bracken SRC provided more specific assessments, targeting the children's knowledge of

letters, numbers and counting (as assessed by the EDI Language/Cognition domain), colors, sizes, and comparisons.

Phonological awareness, the ability to manipulate the smallest units of sound in language, is currently considered one of the best predictors of later success in reading and writing (e.g., Heath and Hogben 2004; Savage and Carless 2004). Although phonological awareness is not tapped directly by the EDI, several items in the Language and Cognition domain include teacher evaluations of children's early reading and writing skills. Accordingly, we included the Comprehensive Test of Phonological Processing (CTOPP, Wagner et al. 1999) as a third comparison measure. Herford (2003) reports that the CTOPP is highly reliable (internal consistency estimates range from .70 to .96; test–retest reliability estimates range from .68 to .97) and that CTOPP scores are significantly related to concurrent reading assessments (correlations range from .19 to .70) and to Woodcock Reading Mastery Test performance both concurrently (.42–.71) and 1 year later (.80). Performance on the CTOPP was expected to be most closely related to EDI evaluations in the Language/Cognition domain.

Finding direct, child-based measures of social competence and emotional maturity was more difficult as, to date, child social competence has typically been assessed via adult report measures, much like the EDI. However, given that we wanted to avoid inflated associations due to shared method variance, we selected the Relationship Questionnaire, a measure developed by Selman and colleagues as part of the Group for the Study of Interpersonal Development (GSID Relationship Questionnaire, Schulz and Selman 2000), the only measure available to provide a direct, child-based assessment of kindergarten children's social competence and emotional maturity. Given its recent development, psychometric data on the Relationship Questionnaire are limited, although it has been used effectively in program evaluation research with preschool-aged children (Schonert-Reichl, personal communication, December, 2004). For this measure, children were asked to determine what is the best thing to do in response to a series of hypothetical social situations. Children's responses to these situations were summed to yield an overall score as well as subscales for (1) Conflict Resolution, (2) Perspective-Taking and (3) Interpersonal Understanding. Although these aspects of children's social and emotional development theoretically underlie the skills and behaviors tapped by the EDI, we acknowledge that they do not provide a direct assessment.

3 Results

To provide a context within which to consider the findings of the present study, we first examined relationships among the comparison measures considered in the present study. As shown in the top spanner of Table 1, significant but modest correlations were found among the comparison measures, with lower correlations observed for the Relationships Questionnaire. The modest correlations observed are not surprising, given that the measures each assess somewhat different aspects of school readiness.

Next, we computed correlations between total readiness scores as assessed by each of the comparison measures and the overall EDI score for the entire sample and for boys and girls separately. As shown in the middle spanner of Table 1, significant but moderate correlations were observed between EDI overall scores and each of the four comparison measures, despite the variations in the developmental areas tapped by the measures and the breadth of coverage of the EDI. Generally, the correlations between EDI overall scores and the comparison indices of readiness were somewhat stronger for girls than for boys, but

Table 1 Correlations between EDI total and domain scores and overall validity measures ($N = 238$–257)

	ESI-K	Bracken SRC	CTOPP composite	Relationship Questionnaire (average level)
Comparison measures				
ESI-K	–			
Bracken SRC	**.33***	–		
CTOPP Composite	**.35***	**.37***	–	
Relationship Questionnaire (overall average level)	**.17***	**.16***	**.16***	–
EDI total score				
EDI total				
Full sample	**.49***	**.46***	**.37***	**.25***
Boys	**.44***	**.38***	**.34***	**.21***
Girls	**.53***	**.60***	**.41***	**.23***
EDI domain subscales				
Physical health/well being	**.41***	.41***	.34***	.21***
Social competence	.45***	.31***	.27***	**.18***
Emotional maturity	.27***	.28***	.11*	**.16***
Language/cognition	.44***	**.34***	.41***	.16**
Communication/general knowledge	.41***	**.49***	**.42***	.26***

* $p < .05$

** $P < .01$

*** $p < .001$

only the correlation between the EDI and the Bracken SRC total scores differed significantly as a function of gender ($z = -2.31$, $p = .01$).

Correlations were also computed between specific EDI domain scores and both total and subscale comparison measure scores, with expectations of stronger associations among total scores and between measures with greater conceptual overlap and low or non-significant correlations between measures with less conceptual overlap (i.e., evidence of convergent/discriminant validity). Results, as presented in the bottom spanner of Table 1, indicated that correlations between EDI Total Scores and total scores on the comparison measures were typically equal to or stronger than those between domains and subscales. However, statistical comparisons indicated that none of the differences in these correlations was significant.

To further explore the convergent and discriminant validity of the EDI, subsequent analyses examined the correlations between EDI overall and domain scores and subscales of the comparison measures, again expecting higher correlations among conceptually overlapping scales. Results are presented in Table 2 for each of the five domains of the EDI. Highlighted in bold are correlation coefficients in the cells where the strongest associations were expected as a result of conceptual overlap across EDI domains and comparison measure subscales.

As shown in Table 2, there is mixed evidence for the convergent validity of the EDI Physical Health/Well-being domain score. For example, the Gross Motor subscale and the Visual-Motor subscale of the ESI-K were significantly correlated with the EDI Physical Health/Well-Being domain scores (.27, .26), although scores on the ESI-K Visual-Motor

Table 2 Correlations between EDI total and domain scores and comparison measures and subscales (N = 249–256)

Comparison measures and subscales	EDI total score	EDI domain scores				
		Physical health/ well-being	Social competence	Emotional maturity	Language/ cognition	Communication/ general knowledge
ESI-K						
Total readiness	**.49***	.41***	.45***	.27***	**.44***	**.41***
Gross-motor	.23***	**.27***	.22***	.15**	.13 *	.20**
Visual motor	.31***	**.26***	.33***	.10	.41***	.19**
Verbal expression	.43***	.32***	.33***	.35***	.29***	**.39***
Language/cognition	.46***	.33***	.38***	.29***	**.38***	.45***
Number concepts	.29***	.14*	.31**	.16**	**.33***	.22***
Visual-sequential memory	.10	.06	.11*	.00	**.10**	.12*
Auditory memory	.18**	.16**	.16**	.10	**.11***	.20**
Bracken						
Readiness composite	.55***	.42***	.44***	.21***	**.65***	**.53***
Letters	.25***	.19***	.23***	.05	**.42***	**.16****
Numbers/counting	.39***	.31***	.33***	.13*	**.57***	**.29***
CTOPP						
Phonological awareness	.37***	.34***	.26***	.12*	**.40***	.42***
Relationship Questionnaire						
Total	.25***	.21***	**.18****	**.16****	.16**	.26***
Perspective taking	.20***	.15**	**.14***	**.13***	.19**	.19**
Interpersonal understanding	.10	.08	**.06**	**.10**	.00	.14*
Conflict resolution	.22***	.23***	**.18****	**.12***	.17**	.22***

* $p < .05$

** $p < .01$

*** $p < .001$

subscale were also significantly associated with the Language/Cognition (.41) and Social Competence (.33) domain scores. Contrary to expectations, EDI Physical Health domain scores were significantly correlated with scores on the Bracken, the CTOPP and the Relationship Questionnaire.

Also contrary to expectations, the correlations between EDI Social Competence and Emotional Maturity scores and scores on the Relationship Questionnaire were low in magnitude and some were non-significant (range: .06–.18, see Table 2). In fact, none of the correlations between subdomains of the EDI and the Relationship Questionnaire exceeded that between EDI Total Scores and the total Relationship Questionnaire score ($r = .25$), although the differences between the correlations were not significant. Contrary to expectations, total scores on the Relationship Questionnaire were significantly correlated with EDI Physical Health/Well-Being and Communication/General Knowledge scores.

Also unexpectedly, EDI Social Competence scores were significantly more strongly related to total scores on the Bracken and ESI-K than to overall scores on the Relationship Questionnaire ($t = -3.41$, $p < .001$ and $t = -3.59$, $p < .001$, respectively). EDI Social Competence scores were also more strongly correlated with CTOPP total scores than with scores on the Relationship Questionnaire ($r = .26$ vs. .18), although not significantly so. A similar pattern was evident for the EDI Emotional Maturity Domain, although differences between correlations were not statistically significant.

EDI Communication and General Knowledge scores were significantly related to all comparison measure scores. As expected, significant correlations were observed between EDI Communication/General Knowledge domain and the ESI-K Verbal Expression and Language/Cognition subscales, the Bracken SRC and the CTOPP. In each case, however, correlations of similar magnitude were observed between comparison measures and scores in the EDI Language/Cognition domain, which may not be surprising given the strong correlation observed between scores obtained in these two EDI domains (r (265) $= .55$, $p < .001$).

Finally, within the Language/Cognition domain, a somewhat mixed pattern of correlational results emerged. Consistent with hypotheses, EDI Language/Cognition scores correlated significantly with the Bracken SRC and subscales and the CTOPP total score. Also, EDI Language/Cognition scores were significantly correlated with the Language/Cognition and Number Concepts subscales of the ESI-K. Unexpectedly, similar modest but significant correlations were also observed between EDI Language/Cognition domain scores and ESI-K Visual-Motor and Verbal Expression subscales. Although the EDI Language/Cognition domain includes items tapping memory, correlations between domain scores and the ESI-K memory subscales were weak.

Given the modest magnitude of the validity correlations observed and the breadth of coverage of the EDI, subsequent analyses addressed the question of how much variance in the EDI overall scores could be accounted for by this set of child-based comparison measures, all of which purport to assess different aspects of school readiness. To this end, a multiple regression analysis was conducted, predicting EDI overall scores from the four comparison measure total scores (ESI-K, Bracken SRC, CTOPP and Relationship Questionnaire), with all predictors entered simultaneously. Results indicated that 36% of the variance in EDI Total Scores was predicted, with each of the four comparison measures making significant and unique contributions ($R = .60$, $R^2 = .36$, $F(4,232) = 32.76$, $p < .001$).

Although thus far our analyses have been at the level of individuals, a final set of correlational analyses was conducted at the classroom level to explore the relationships between EDI scores and the two standardized readiness comparison measures for classrooms in which we had at least 10 pairs of scores. Results indicated considerable variability across teachers in the degree to which their readiness assessments on the EDI corresponded with two standardized readiness indices. Specifically, among those teachers ($n = 12$ out of 27) for whom data were available on at least 10 of their students (range 10–21), correlations between EDI Total Scores and ESI-K total scores ranged from .17 to .95. Similarly, correlations between EDI Total Scores and the Bracken School Readiness Composite ranged from $-.04$ to .66 across teachers.

4 Discussion

The aim of the present study was to examine the convergent and discriminant validity of the EDI, with interest in associations with direct, child-based assessments, including three

standardized readiness assessments with well-established psychometric properties and a newly developed measure of children's social competence. This was an important task, as previous validation studies of the EDI have tended to rely on other teacher reports as comparison measures, which capitalize on shared method variance and potentially inflate correlations among those measures.

To summarize, in a sample of kindergarten children from socioeconomically and culturally/ethnically diverse neighborhoods who were evaluated by 27 teachers in 16 schools across three districts, we found that overall EDI scores (averaged across the five EDI domains) were significantly correlated with both standardized measures of school readiness (ESI-K, Bracken SRC) and with direct, child-based indices of early social (Relationship Questionnaire) and academic competencies (CTOPP, phonemic awareness) thought to be important to early school success. The magnitude of these correlations was moderate, but not unlike those observed among the comparison measures in the present sample (Table 1) and in previous research. In fact, the correlations observed between overall EDI scores and the ESI-K Total Readiness score (.49) and the Bracken School Readiness Composite (.55) were comparable to those reported by Ford et al. (2007) assessing a much smaller sample (.46 and .39, respectively). The magnitude of these correlations is particularly noteworthy given (a) the difficulty of assessing readiness in young children, (b) the challenge of identifying comparison measures which effectively tap the breadth of coverage of the EDI, and (c) the fact that the present study compared teacher evaluations on the EDI with direct, child-based assessments across different domains of development that emerge at different rates.

Higher correlations have been reported in studies using comparison measures that capitalize on shared method variance (e.g., comparing teacher reports with teacher reports). For example, Brinkman et al. (2007) reported higher correlations between EDI scores and teacher-reports than between EDI scores and parent-reports or direct assessment measures. By way of comparison, reported concurrent correlations between the standardized comparison measures and other standardized indices of related readiness constructs, as published in the technical manuals for the ESI-K and the Bracken measures, were somewhat higher than those observed in the present study, ranging from .57 to .86, and represent upper level expectations for validity correlations, based on associations between two standardized, child-based measures with considerable conceptual overlap. One of the strengths of the present investigation was the effort to evaluate the validity of teacher evaluations in terms of actual child performance, eliminating the problem of shared method variance. Taken together, these findings provide correlational evidence that support the construct validity of the overall EDI score with respect to direct, expert-administered school readiness assessments, suggesting that the overall EDI score may provide a reasonably valid index of general school readiness.

However, correlational evidence for the convergent and discriminant validity of the EDI domain scores with respect to direct, expert-administered school readiness assessments was more equivocal. First, as shown in Table 1, the correlations observed between standardized school readiness measures (ESI-K, Bracken SRC) and overall EDI scores, with one exception, were of greater magnitude than those observed for each of the five domains scores of the EDI, but the correlations did not differ significantly from one another. For the other two comparison measures, which were research-based indices tapping more restricted aspects of readiness, results were even more mixed. Specifically, indices of phonological awareness (CTOPP) were more strongly associated with EDI scores in conceptually related domains (Language/Cognition, Communication/General Knowledge) than with overall EDI scores or scores in less-related domains, but not significantly so.

Direct, child-based assessments of social competence (Relationship Questionnaire) showed weak but significant correlations with all of the EDI domain scores (correlations ranged from .16 to .26). Interestingly, the strongest correlation observed with the Relationship Questionnaire was not with the EDI Social Competence or Emotional Maturity domains, as hypothesized, but with the Communication/General Knowledge domain. Perhaps children with strong communication skills responded more competently in discussions with an examiner, as required with the Relationship Questionnaire.

In interpreting these results, we recognize that our validity index of social competence, the Relationship Questionnaire, is itself a relatively new measure for which there is little psychometric information. The assessment of children's social competence generally has been an area fraught with difficulties and to date, despite numerous conceptual papers on the topic (e.g., Masten et al. 1995; Pellegrini and Glickman 1990, Schneider 1993; Topping et al. 2000) there is no clear or agreed upon definition of, nor consensus about the breadth of the construct. Thus, it is perhaps not surprising that the validity correlations in the two social-emotional domains of the EDI are particularly weak in the present study.

Further analyses, as presented in Table 2, examining the correlations between each of the EDI Domain scores and conceptually-related comparison measures and subscales, provided only limited support for the convergent and discriminant validity of the EDI Domain scores. The strongest case for convergent validity was observed for the EDI Language/Cognition domain, an area that is typically included in traditional readiness assessments. In the present study, scores in the EDI Language/Cognition domain correlated most strongly with conceptually similar indices of readiness including the Bracken SRC (.65) and its subscales, the CTOPP assessment of phonological awareness (.40), and the ESI-K total score (.44) and its subscales of Language/Cognition and Number Concepts (.38, .33). Regarding discriminant validity, the correlation between EDI Language/Cognition domain scores and scores on the Relationship Questionnaire (social competence, $r = .16$) were significantly lower than those observed between EDI Language/Cognition domain scores and the Bracken SRC ($t = -8.27$, $p < .001$), but not significantly lower than the correlation observed with the CTOPP or the ESI-K scale and subscales. Moreover, among the subscales of the ESI-K, EDI Language/Cognition domain scores were significantly related to scores on the ESI-K Visual-Motor ($r = .41$) and Language/Cognition ($r = .38$) subscales, as expected, and less strongly related to other ESI-K subscales, but the differences among these correlations were all nonsignificant, making the case for discriminant validity less clear.

The EDI Communication/General Knowledge domain scores correlated significantly with all of the comparison measures, including the ESI-K total (.41), the Bracken SRC (.53), and the CTOPP (.42), but also the Relationship Questionnaire (.26), with no significant differences across these correlations. Within the ESI-K subscales, as expected, the EDI Communication/General Knowledge domain scores were most strongly associated with scores on the ESI-K Language/Cognition (.45) and Verbal Expression (.39) subscales, and significantly but less strongly with the other subscales (.12–.20, see last column in Table 2) although again, the differences among all of these correlations were nonsignificant.

For the EDI Physical Health/Well Being domain, significant but moderate correlations were observed with conceptually-related ESI-K Gross-Motor (.27) and Visual-Motor (.26) subscales, as expected. However, these correlations were not significantly stronger than those observed with less conceptually-related measures, including the overall Bracken School Readiness score (.42), the CTOPP (.42), and with the ESI-K subscales for Verbal Expression (.32) and Language/Cognition (.33), calling into question the discriminant

validity of this domain score. It is recognized, however, that the comparison measures we used did not provide extensive coverage of this EDI domain.

Finally, with regard to the EDI Social Competence and Emotional Maturity domains, correlations observed with the Relationship Questionnaire overall and subscale scores were typically low and sometimes non-significant, and did not differ significantly from the relatively stronger correlations obtained for these two EDI domains and the overall ESI-K Readiness total (.45 and .27) and the overall Bracken SRC (.44 and .21, respectively). However, it is important to recognize that, whereas the Relationship Questionnaire was designed as a direct assessment of children's verbal understanding of socially competent behavior, few of the items in the EDI Social Competence domain actually tap overall "social competence", but instead reflect such things as openness and eagerness to follow classroom rules and motivation or curiosity in learning. Similarly, the items included in the Emotional Maturity domain appear to be more closely aligned with interpersonal behavior and skills (prosocial behavior, aggression, anxiety, inattention/hyperactivity). It is likely that the low correlations between scores on the Relationship Questionnaire and both EDI Emotional Maturity and Social Competence reflect the fact that the selected scales measure constructs that do not fully overlap. Given the challenges inherent in direct, child-based assessments of social and emotional competence currently, it remains for future research to evaluate further the validity of the EDI Social Competence and Emotional Maturity domains.

Taken together, results of the present study provide correlational evidence that supports the convergent validity of the overall EDI score with respect to direct, child-based school readiness assessments. However, the correlational evidence for the convergent and discriminant validity of the EDI domain scores, when related to corresponding domains of direct, child-based school readiness assessments, is less compelling. One limitation of the present investigation was the ability to identify measures that adequately assessed performance across the varied domains and skills tapped by the EDI. Together, the comparison measures only accounted for about one-third of the variance in overall EDI scores, owing in part to the limits of the selected comparison measures to adequately match the breadth of the EDI. Future research on the validity of the EDI would benefit from considering a broader range of comparison measures that might be evaluated at a more micro/item level with teacher ratings on the EDI. A second potential limitation concerns the time lag between administration of individual assessments and teacher completion of the EDI. In a few cases, this lag was as long as 3 months, during which time children's skills and competencies may have changed. Ideally, assessments should be closer in time. However, given the task of completing comprehensive individual assessments of very young children with limited attention spans, approximating simultaneous assessment remains a challenge.

As used currently, interpretation of the EDI is typically made at the level of domain scores and, with few exceptions (e.g., Ford et al. 2007), an overall readiness score is not computed. The present results indicated that only one of the five EDI domain scores (Language/Cognition) demonstrated sufficient convergent and discriminant validity to enable interpretation. Evidence for the convergent and discriminant validity of the EDI Emotional Maturity and Social Competence domain scores was particularly weak. The degree to which this reflects a problem with the EDI itself, or with the identification of appropriate and psychometrically sound comparison measures, or method effects (e.g., rater familiarity, assessment versus rating, context effects) remains uncertain, especially for the Social Competence, Emotional Maturity and Physical Health/Well Being domains.

An important issue arising from the present findings concerns the implications of the lack of validity of EDI domain scores at the individual level for the EDI's validity at the

aggregate level (see Zumbo and Forer in press). Although it is possible to have low individual level reliability of measurement but higher aggregate level score reliability, convergent and discriminant validity of EDI domain scores at the individual level are desirable, because valid individual level scores are likely a sufficient condition for the validity of aggregated domain scores (which is how the EDI is intended to be and currently is being used). Establishing validity of EDI domain scores at both the individual and aggregate levels remains a challenge for future research, especially with regard to the assessment of social and emotional development in young children.

References

Bracken, B. C. (1998). *Bracken basic concept scale-revised*. San Antonio, TX: The Psychological Corporation, Harcourt Brace and Company.

Brigance, A. H. (1997). *Brigance K and 1 screen, revised*. North Billerica, MA: Curriculum Assoc.

Brinkman, S. A., Silburn, S., Lawrence, D., Goldfeld, S., Sayers, M., & Oberklaid, F. (2007). Investigating the validity of the Australian early development index. *Early Education & Development, 18*(3), 427–451.

Chambers, J., & Windschitl, P. (2004). Biases in social comparative judgments: The role of non-motivated factors in above-average and comparative-optimism effects. *Psychological Bulletin, 130*, 813–838.

Clancy, C. H., & Pianta, R. C. (1993). The metropolitan readiness test as a descriptor and predictor of children's competence in kindergarten through grade two. *Journal of Psychoeducational Assessment, 11*(2), 144–157.

Connor, S. & Brink, B. (1999). *Understanding the early years: Community impacts on child development*. Human Resources Development Canada, Working paper W-99-6E. (available at http://www.sdc.gc. ca/en/cs/sp/sdc/pkrf/publications/nlscy/uey/1999-000092/page00.shtml).

Doherty, G., & Stuart, B. (1997). The association between childcare quality, ratio and staff training: A Canada-wide study. *Canadian Journal of Research in Early Childhood Education, 6*(2), 127–138.

Duku, E. & Janus, M. (2004). *Stability and reliabilityof the early development Instrument: A popu-lation-based measure for communities (EDI)*. Paper presented at the annual Department of Psychiatry and Behavioural Neurosciences Research Day, McMaster University, Hamilton, ON.

Dunn, L. M., & Dunn, L. (1981). *Peabody picture vocabulary test-revised. Manual for Forms L and M*. Circle Pines, MN: American Guidance Service.

Ford, L. Dahinten, V. S., Merkel, C., Chong, S. L., & Moraes, S. (March, 2007). *Investigating the validity of the early development instrument at the child level*. Paper presentation at the annual meeting of the National Association of School Psychologists.

Forget-Dubois, N., Lemelin, J., Boivin, M., Dionne, G., Séguin, J. R., Vitaro, F., et al. (2007). Predicting early school achievement with the EDI: A longitudinal population-based study. *Early Education & Development, 18*(3), 405–426.

Glascoe, F. P. (1997). *Technical report for the Brigance screens. Anonymous*. Billlerice, MA: Curriculum Assoc.

Graue, E. (2006). The answer is readiness–now what is the question? *Early Education and Development, 17*, 43–56.

Guhn, M., Gadermann, A., & Zumbo, B. D. (2007a). Does the EDI measure school readiness in the same way across different groups of children? *Early Education & Development, 18*(3), 453–472.

Guhn, M., Janus, M., & Hertzman, C. (2007b). The early development instrument: Tranlating school readiness assessment into community actions and policy planning. *Early Education & Development, 18*(3), 369–374.

Heath, S., & Hogben, J. (2004). Cost-effective prediction of reading difficulties. *Journal of Speech, Language and Hearing Research, 47*(4), 751–766.

Herford, D. (2003). The comprehensive test of phonological processing. *Mental Measurements Yearbook, 15*, 226–229.

Hubley, A. M., & Zumbo, B. D. (1996). A dialectic on validity: Where we have been and where we are going. *Journal of General Psychology, 123*(3), 207–215.

Irwin, L., Siddiqi, A. & Hertzman, C. (2007). World Health Organization commission on the social determinants of health early childhood development: A powerful equalizer www.earlylearning.ubc.ca/globalknowledgehub/WHO_ECD_Final_Report.pdf.

Janus, M. (2006). *The Early Development Instrument (EDI): The national and international perspective.* Paper presented at the annual meeting of the American Educational Research Association as part of a symposium organized by Martin Guhn entitled, "Translating school readiness assessment into community actions and policy planning: The Early Development Instrument Project". San Francisco, CA, April 2006.

Janus, M. & Offord, D. (2000). Readiness to learn at school. *ISUMA, 1* (2), 71–75. (available at www.offordcentre.com/readiness/pubs/publications.html).

Janus, M., & Offord, D. (2007). Development and psychometric properties of the early development instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioural Science, 39*(1), 1–22.

Janus, M., Walsh, C., Viveiros, H., & Offord, D. R. (2002). *Community, neighborhood, and 5-yeaar-olds readiness to learn at school.* Paper presented at the head start conference, Washington, DC.

Keating, D. P. (2007). Formative evaluations of the early development instrument: Progress and prospects. *Early Education & Development, 18*(3), 561–570.

Keating, D. P., & Hertzman, C. (1999). *Developmental health and the wealth of nations: Social biological, and educational dynamics.* New York, NY: Guilford Press.

Kershaw, P., Forer, B., Irwin, L. G., Hertzman, C., & Lapointe, V. (2007). Toward a social care program of research: A population-level study of neighborhood effects on child development. *Early Education & Development, 18*(3), 535–560.

Kershaw, P., Irwin, L., Trafford, K., & Hertzman, C. (2005). *The British Columbia Atlas of child development* (1st ed.). Victoria, British Columbia, Canada: Human Early Learning Partnership and Western Geographical Press.

Kimmel, E., & Paget, K. D. (2001). The early screening inventory. *Mental Measurements Yearbook, 14*, 450–453.

Konold, T. R., & Cox, R. C. (2005). Empirically-derived, person-oriented patterns of school readiness in typically-developing children: Description and prediction to first-grade achievement. *Applied Developmental Science, 9*(4), 174–187.

Masten, A. S., Coatsworth, J. D., Neemann, J., Gest, S. D., Tellegen, A., & Garmezy, N. (1995). The structure and coherence of competence from childhood through adolescence. *Child Development, 66*, 1635–1659.

McCarthy, D. (1972). *McCarthy scales of children's abilities.* San Antonio, TX: Psych. Corporation.

McKnight, T., & Schwarting, G. (2004). Bracken school readiness assessment. *Mental Measurements Yearbook, 16*, 155–157.

Meisels, S. J. (1987). Uses and abuses of developmental screening and school readiness testing. *Young Children, 42*(4–6), 68–73.

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. Brooks.

Meisels, S. J., Marsden, D. B., Wiske, M. S., & Henderson, L. W. (1997). *Early screening inventory: Revised. examiner's manual.* Ann Arbor, MI: Rebus Inc.

National Research Council and Institute of Medicine (2000). *From neurons to neighbourhoods: The science of early childhood development.* In: Committee on integrating the science of early childhood development, J. Schonkoff. & D.Phillips (Eds.). Washington, DC.: National Academy Press.

Nurss, J. R., & McGauvran, M. E. (1995). *Metropolitan readiness tests* (6th ed.). San Antonio, TX: Harcourt Brace Educational Measurement (Psychological Corporation).

Pellegrini, A. D., & Glickman, C. D. (1990). Measuring kindergartners' social competence. *Young Children, 45*, 40–44.

Savage, R., & Carless, S. (2004). Predicting curriculum and test performance at age 7 from pupil background, baseline skills and phonological awareness at age 5. *British Journal of Educational Psychology, 74*(2), 155–171.

Sayers, M., Coutts, M., Goldfeld, S., Oberklaid, F., Brinkman, S., & Silburn, S. (2007). Building better communities for children: Community implementation and Evaluation of the Australian early development index. *Early Education & Development, 18*(3), 519–534.

Schneider, B. (1993). *Children's social competence in context; The contrib-utions of family school and culture*. Tarrytown, NY: Pergamon Press.

Schulz, L. H., & Selman, R. L. (2000) *The meaning and measurement of social competence from a developmental perspective*, Working paper 153. New York: Russell Sage.

Sturner, R. A., Funk, S. G., & Green, J. A. (1984). Predicting kindergarten school performance using the McCarthy scales of children's abilities. *Journal of Pediatric Psychology, 9*(4), 495–503.

Topping, K. J., Bremner, W. G., & Holmes, E. A. (2000). Social competence: The social construction of the concept. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence*. San Francisco: Jossey-Bass.

Vernon-Feagans, L., & Blair, C. (2006). Measurement of school readiness. *Early Education and Development, 17*(1), 1–5.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: PRO-ED.

Zumbo, B. D., & Forer, B. (in press). Testing and measurement from a multilevel view: Psychometrics and validation. In J. Bovaird, K. Geisinger, & C. Buckendahl (Eds.), High stakes testing in education—science and practice in K-12 Settings [Festschrift to Barbara Plake]. Washington, DCs: American Psychological Association Press.