Predicting Academic Difficulties: Does a Complex, Multidimensional Model Outperform a Unidimensional Teacher Rating Scale?

CLAIRE V. CROOKS, University of Western Ontario
Centre for Research on Violence Against Women and Children, Faculty of Education
RAY DeV. PETERS, Better Beginnings, Better Futures Research Project,
Queen's University

Abstract

The main objective of this study was to compare a multidimensional model (i.e., including social, cognitive, physical, behavioural, and language functioning) of academic readiness and a unidimensional teacher-rated screening questionnaire on the basis of accuracy in predicting future problems in children. The participants were 146 children (80 boys and 66 girls) living in disadvantaged communities in Ontario, Canada. Data were collected in junior kindergarten and at Grade 3. Constructs were assessed with multiple methods and multiple raters. Results for the multidimensional model indicated that social-behavioural functioning, and to a lesser extent, cognitive/language skills were most salient in differentiating between children with problems and those without. The multidimensional model had accuracy and sensitivity comparable to the unidimensional model, but the multidimensional model also had a high rate of false positives as indicated by poor specificity. Implications for identifying high-risk children are discussed.

Résumé

L'objectif principal de cette étude était de comparer un modèle multidimensionnel (c.-à-d. qui comprenait le fonctionnement social, cognitif, physique, comportemental et langagier) de la préparation universitaire et un questionnaire d'évaluation de la sélection des enseignants unidimensionnel sur la base de l'exactitude dans la prédiction de problèmes futurs chez l'enfant. On participé à l'étude 146 enfants (80 garçons et 66 filles) vivant dans des communautés désavantagées de l'Ontario, au Canada. Les données ont été récoltées auprès d'enfants de la maternelle et de troisième année. Les construits ont été évalués à l'aide de plusieurs méthodes et plusieurs évaluateurs. Les résultats du modèle multidimensionnel ont indiqué que le fonctionnement social-comportemental et dans une moindre mesure, les habilités cognitives/langagières étaient plus importants dans la différenciation entre les enfants avec problèmes et ceux qui n'en avaient pas. Le modèle multidimensionnel avait une exactitude et une sensibilité comparables au modèle unidimensionnel, mais le modèle multidimensionnel avait aussi un taux élevé de positifs faux tel qu'indiqué par la faiblesse de la spécificité. Les conséquences d'identifier les enfants à haut risque font l'objet de discussion.

Early identification of children "at risk" for academic failure facilitates more timely intervention, thus addressing problems before they become entrenched and require expensive, long-term intervention. The term academic readiness has been employed to identify those children who are not at risk, but rather are well poised for a successful transition to school. The notion of academic readiness has evolved into a multidimensional construct incorporating social-behavioural, emotional, physical, cognitive, and language functioning. This multidimensional complexity contrasts with earlier work in the area, which emphasized screening for difficulties with brief, unidimensional teacher rating scales. Indeed, earlier frameworks viewed emotional and behavioural problems as secondary results of learning difficulties, rather than as an integral part of the presenting problem (Tramontana, Hooper, & Selzer, 1988). The goal of the present study was to compare a multidimensional model of academic readiness and a unidimensional teacher-rated screening questionnaire on the basis of accuracy for predicting future problems in children.

At a policy level, the conceptualization of academic readiness in North America has been strongly influenced by two papers, both of which espouse a multidomain or holistic model of child functioning

(Doherty, 1997; Kagan, Moore, & Bredekamp, 1995). The five domains included in these approaches are physical well-being, emotional health, social competence, language skills, and cognitive skills and general knowledge. Implicit to this approach is the idea that not only must this multidimensionality be recognized theoretically, but that research and policy initiatives must also embrace the concept of academic readiness in the complete, multidimensional sense of the term (Doherty; Kagan et al., 1995). Furthermore, the outcome is similarly multidimensional, in that achievement and adjustment are embedded within the larger concept of academic success. Thus, both Canadian and American policy research emphasize a risk identification model, despite a lack of empirical validation. It is crucial that the risk identification model be evaluated, as lack of empirical support might provide indirect support for a more universal intervention approach. That is, if it is not feasible to identify children who will develop problems using broad-based measures, then services targeting all children in particular groups or communities might be a logical alternative.

Although the integrated multidimensional model of academic readiness has not been evaluated, each of the components has sound theoretical and empirical support for being related to adjustment in the school setting. The theoretical basis for the inclusion of the physical domain is two-fold. In the case of clear physical and/or neurological limitations, the child's developing academic skills may be directly affected by ability. In addition, conditions requiring frequent or lengthy school absences diminish the child's opportunities to acquire and practice new skills. In terms of some of the more subtle fine motor skills, children who lack these types of skills may be negatively affected in that they are not acquiring the fundamental building blocks of later tasks. Furthermore, given these motor deficits, they may come to view themselves and the school experience negatively from an early age (Doherty, 1997).

School readiness requires several types of skills that fall under the rubric of emotional functioning. Emotional maturity is required such that a child must exhibit self-control and the ability to delay gratification. Lack of self-control, as indicated by hyperactivity and /or inattention, has been linked to poor academic performance, particularly when combined with other behavioural deficits (Flanagan, Bierman, & Kam, 2003). From a deficit perspective, high levels of anxiety, or school phobia can result in absenteeism, or interfere with a child's ability to devote attention to the tasks at hand. A child must also have the emotional stability to cope with repeated attempts at a task, or momentary failure without becoming highly distressed and disruptive. Doherty's (1997) conceptualization of emotional health appears to encompass two domains – emotional functioning and emotional dysfunction – the latter of which tends to be equated with internalizing disorders such as anxiety and depression.

In the multidimensional model of academic readiness, social functioning incorporates an awareness of the general behaviour standards for a public place, respect for others (teachers and fellow students), an ability to inhibit inappropriate aggressive behaviour, and the ability to communicate feelings and wants in socially acceptable ways. Children who lack appropriate social and behavioural skills at the time of school entry may resort to less accepted means of having their needs met, including aggression towards others and disruptive behaviour in the classroom. The use of aggressive behaviour is associated with subsequent peer rejection (Coie & Kupersmidt, 1983; Dodge, 1983), which in turn affects the degree to which a child engages in the prosocial culture of the school. Indeed, an array of early behaviour problems (including internalizing, externalizing, hyperactivity-impulsiveness, and immaturity-dependence) predicts later victimization, even after accounting for concurrent behaviour problems (Schwartz, McFadyen-Ketchum, Dodge, Pettit, & Bates, 1999). More recent research has shown that the co-occurrence of these multiple social and behavioural deficits is particularly problematic for children's successful transition to school (Flanagan et al., 2003).

Although the manifestation of early behaviour problems predicts future difficulties, simply lacking appropriate social competence at school entry is also problematic. Canadian longitudinal research demonstrates that the lack of appropriate social skills at the time of school entry is one of the best predictors of delinquent behaviour in early adolescence (Tremblay et al., 1992). Other researchers have emphasized deficits in prosocial functioning as central to academic success. For example, Caprara and colleagues found that poor prosocial functioning at Grade 3 predicted academic functioning five years later, even after controlling for aggression and academic functioning at Grade 3 (Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000).

The importance of language skills to the successful adjustment to school is such that measures of language, such as the Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981), are often used as screening tools to identify children who may not be ready for school. Receptive and expressive language are essential both as the means by which information

and directions are imparted, and as a means for the child to express his or her own needs appropriately. Indeed, language impairment at the time of school entry has been found to predict a more general academic deficit, as well as specific language difficulties later in school (Fazio, Naremore, & Connell, 1996). As the formal education system continues to be rooted in verbal learning, this dimension is clearly an essential one to be included in any study of academic readiness.

Cognitive functioning in the mulitidmensional model of academic readiness describes the ways in which individuals perceive, organize, and analyze information provided by their senses (Doherty, 1997). The original work on academic readiness tended to focus on the area of cognitive skills to the exclusion of other domains, such as social and emotional functioning (Tramontana et al., 1988). The ability to adequately retain and retrieve information is a cornerstone of the learning process. There is considerable research indicating that a child's level of cognitive skills prior to school entry is predictive of later academic success (see Doherty, 1997). Research with a Canadian sample has shown that child IQ is related to subsequent school drop-out, even after controlling for family configuration, parental education, and child disruptiveness (Vitaro, Brendgden, & Tremblay, 1999).

In summary, there are two models for assessing academic readiness: the publicly supported but unevaluated multidimensional model, and a unidimensional model that relies on the use of simpler assessment measures. Although it is widely assumed that a complex multidimensional model represents a marked improvement over the unidimensional model, the extent to which either of these models provides the level of screening accuracy required for prudent decision-making has not been established.

Current Study

The current study was designed to predict academic outcomes at Grade 3 from child functioning at the time of school entry in children living in disadvantaged communities (i.e., to evaluate the academic readiness paradigm). There were three objectives in this investigation. First, an attempt was made to delineate a multidimensional model of child functioning. Second, a multidimensional model of child functioning at school entry was compared to a unidimensional teacher rating in regard to ability to predict problematic outcomes at Grade 3. Outcomes at Grade 3 included academic achievement and social-behavioural adaptation, as both of these domains have been identified as key components of academic success (Huffman, Mehlinger, & Kerivan, 2000).

Finally, the two models of academic readiness were compared on two epidemiological indices (sensitivity and specificity) to evaluate the utility of using either model to predict academic difficulties. Sensitivity refers to the proportion of children with a negative outcome who were considered to be at-risk at the initial screening. Is this case it is the proportion of children who have academic difficulties in Grade 3 who were identified as high risk at school entry (i.e., valid positives/[valid positives + false negatives]). Conversely, specificity refers to the proportion of children without negative outcomes who were considered low-risk at the initial screening. Thus, it is the proportion of students who do not develop difficulties at Grade 3 who were not identified as high risk at school entry (i.e., valid negatives/[valid negatives + false positives]). These study objectives can be summarized with the following two research questions:

Are each of the five child factors identified by Doherty (1997) and Kagan, Moore, and Bredekamp (1995) both independent and uniquely important in the prediction of school functioning?

How valid is identification at school entry based on a broadly defined holistic model for determining which individual children will struggle with the demands of the formal school setting?

Method

Participants

The participants were 147 children (81 boys and 66 girls) living in disadvantaged communities. The selected communities were all either demonstration or comparison sites of the Better Beginnings, Better Futures prevention project. Extensive documentation of the sample characteristics, recruitment, and procedures is outlined elsewhere by Peters and colleagues (Peters, Petrunka, & Arnold, 2003). Data were collected when the children were in junior kindergarten (mean age = 4.6 years, range = 3.3 to 5.4), as well as when they were in Grade 3 (mean age = 8.6 years, range = 7.6 to 9.5). Twenty-eight percent of the children were living in single-parent households, and 53% were living in households whose income was less than the Statistics Canada low income cut-off. With respect to ethnicity, 38% of the children were Anglophone Caucasion, 30% were French Canadian, 3% were Native American, and 30% were other. In addition, 30% of the children had primary caregivers who were born outside of Canada. Active parental consent was obtained and all protocols received clearance from the university research ethics board.

Procedure

Data were collected through a combination of interviews with parents, parent-report questionnaires, teacher-report questionnaires, and child administered measures. Parent interview questions included items such as the presence or absence of particular symptoms or difficulties. Parent self-report included questionnaires on children's social, behavioural, and emotional functioning. Teacher-report data included questionnaires rating children's social, behavioural, and emotional functioning, as well as general school functioning questionnaire. Child-administered measures included a subtest from an intelligence test and a test of receptive vocabulary. The specific measures are described in the following section.

Measures

Unidimensional model of functioning at school entry. The unidimensional model was operationalized with the ABC Kindergarten Teacher Observation Rating System (ABC), a 17-item teacher rating scale developed by the Toronto Board of Education to assess developmental maturity and to identify children who might have difficulties with the transition to school (Toronto Board of Education, 1990). This measure has items that tap the individual components of the holistic child model (i.e., social-behavioural functioning, language, etc.), but has a unidimensional factor structure (Arnold, 1995). Teachers rate a child's competence on a 5-point scale (ranging from "not yet apparent or becoming positive" to "extremely well developed") in areas such as, "Relationships with children: Beginning to relate in comfortable, helpful ways" and "Attention (teacher-directed): Focuses and sustains interest for varying periods of time depending on the nature of the task." The internal reliability for the ABC was high ($\alpha = .97$).

Multidimensional model of functioning at time of school entry. The multidimensional model was operationalized using multiple measures that mapped onto the constructs identified as components of academic readiness. Several of the variables are based on published scales, but have been adapted on the basis of confirmatory factor analyses of the Better Beginnings data. As many of the scales differ slightly from the published versions (based on the factor analysis results), reported internal reliability coefficients are for the entire Better Beginnings, Better Futures data set, unless otherwise noted (Arnold, 1995).

1) *Illness/Physical Limitations construct.* Parents were asked if their child had specific chronic health problems (e.g., asthma, heart or lung problems, kidney problems, epilepsy). In addition, parents were asked if their child had ever stayed overnight in a hospital, whether they had health difficulties that limited their activities, and the length of time that any activity-limiting conditions had existed.

2) *Emotional construct.* Measures for the emotional construct originated from a variety of sources. Shyness was measured with six items adapted from a social wariness subscale ($\alpha = .64$ for teachers and .61 for parents). Other items were taken from the Tri-Ministry Helping Children Adjust project. These three items, which were used in the Ontario Child Health Study to identify Emotional Disorder, were combined with the parent and teacher-rated anxiety items from the Preschool Social Behaviour Questionnaire (PSB-Q; Tremblay et al., 1992). Internal reliability for these combined anxiety items was .64 for teachers and .61 for parents.

3) *Social-Behavioural construct*. The Social-Behavioural construct was measured mainly by teacher and caregiver completion of the PSB-Q (Tremblay et al., 1992). There were 10 items on both the teacher- and parent-rated prosocial scales ($\alpha = .92$ and .78, respectively), 12 items on the parent-rated disruptive scale ($\alpha = .90$), and 14 items on the teacher-rated disruptive scale ($\alpha = .95$). In addition, a binary judgment of whether or not a child had a behaviour problem was collected from parents.

4) *Language Skills construct*. Language skills were measured with the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981), a non-verbal, multiple-choice test designed to evaluate the hearing vocabulary or receptive vocabulary of children and adults. The split-half reliability reported by the authors for the two- to six-year age group is .67, and the validity of the PPVT-R has been demonstrated. Parents were also asked to report whether or not their family doctor had ever identified the child as having learning problems.

5) *Cognitive Functioning construct*. The Block Design subtest from the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1967) was administered. This subtest requires children to reproduce stimulus designs using three or four blocks. The published manual for the WPPSI reports an alpha reliability coefficient of .82 (Wechsler). Parents' reports of doctor-identified developmental delay were also collected.

Grade 3 outcomes: Behaviour. Three scales from the Canadian National Longitudinal Survey of Children and Youth (NLSCY; Human Resources Development Canada) were used to measure behaviour problems at Grade 3; the Hyperactivity Disorder scale (7 items), the Conduct Disorder scale (6 items), and the Emotional Disorder scale (8 items). Internal reliability coefficients range from .82 to .93 for these scales. National NLSCY data were used to establish cut-off scores as close to the 90th percentile as possible, as this cut-off has been used previously with the NLSCY data to identify children with difficulties reaching clinical significance (Lipman, Offord, & Dooley, 1996; Offord & Lipman, 1996). These percentiles are in reference to all of the Grade 3 students in the national sample. In total, 48.3% (n = 71) of the children in the present sample were characterized as having at least one behaviour problem by a parent, teacher, or both.

Grade 3 outcomes: Achievement. Three measures were used to evaluate academic achievement outcomes. The math test was taken from the NLSCY protocol, and originated with the Canadian Achievement Test. The Peabody Picture Vocabulary Test was used to assess receptive vocabulary (Dunn & Dunn, 1981). A final indicator of academic performance was whether or not the child had been held back a year. Achievement cut-offs were based on established convention where possible; math scores were dichotomized at a point that separated the lowest functioning 10% from the rest of the normative sample, and a one standard deviation cut-off was used with the PPVT. These various cut-points delineate the bottom 10-15% of a normative sample, which coincides approximately with the percent of youth who do not graduate from high school by age 24 (i.e., 14%). Finally, if a child was at least one grade behind what would be predicted by the year he/she entered the formal school system, he/she was also considered to be having achievement difficulties. When all of the achievement measures were combined, it was found that 55.8% (n = 82) of the sample was classified as being below the cut-off for academic success on at least one of these three indices, as would be expected given that this was a high-risk community sample. When the areas of behavioural adaptation and achievement were combined, 70% of the sample was identified as having problems in one or both areas.

Results

Latent constructs were generated to represent difficulties in independent dimensions of child functioning (i.e., Physical health, Emotional well-being, Social functioning, Language skills and Cognitive functioning). Separate confirmatory factor analyses were performed to refine each of the factors individually with AMOS 3.6 (Arbuckle, 1997). Based on the factor analyses, several adjustments were required to the individual factors that were proposed in the Method section. Standardized beta weights for all of the confirmatory analyses are presented in Table 1.

TABLE 1
Standardized Beta Weights for Latent Constructs

Construct	
Item/Scale	Standardized Beta Weight
Social-Behavioral Problems	
Teacher-rated disruptiveness	.76
Parent-rated disruptiveness	.55
Teacher-rated prosocial	-.43
Parent-reported presence of behavior problem	.53
Parent-rated Emotional Problems	
Shyness	.62
Anxiety	.75
Teacher-rated Emotional Problems	
Shyness	.82
Anxiety	.59
Cognitive and Language Problems	
Peabody Picture Vocabulary Test score	-.59
WPPSI Block Design Subtest score	-.54
Parent report of MD-diagnosed learning problems	.38
Parent report of MD-diagnosed developmental delay	.49
Illness/Physical Limitations	
Specific health difficulties (e.g., asthma, heart problems)	.46
Child has stayed in a hospital overnight	.22
Health-related limitations of normal activities	.97
Chronicity of health limitations	.98

Illness/Physical Limitations Factor

Although the illness/limitations factor was generated with four measures, in reality it was more of a dichotomous concept as the two indicators of physical limitations had extremely high loadings (i.e., .97 and .98). A histogram of the construct confirmed this dichotomy. The goodness of fit and comparative fit indices were both .99, with a chi-square of 4.53 (df = 2, $p < .01$).

Emotional Functioning Construct

Attempts to load teacher- and parent-rated shyness and anxiety on a single factor were unsuccessful. The nature of the correlation matrix highlighted the problem. The correlations among these measures

TABLE 2
Logistic Regression Results for Child Variables Predicting Problematic Grade 3 Outcomes (N = 146)

Variable	*B*	*S.E. B*	Wald Statistic (df = 1)	Significance of Wald	Change in -2 Log Likelihood (df = 1)	Significance of the Change
Step 1						
Social-behavioral	.99	.27	13.48	< .01	18.83	< .01
Step 2						
Social-behavioral	.90	.27	10.80	< .01	14.74	< .01
Cognitive/language	.68	.26	6.73	< .01	7.63	< .01

TABLE 3
Classification Accuracy for Multidimensional and Unidimensional Models

	Predicted Outcome Group		
	n (% within actual outcomes group predicted correctly)		
Actual Outcome Group	No problems	One or both problems	Overall classification accuracy
Multidimensional Model			
No problems	12 (27.9%)	31 (72.1%)	73.3%
One or both problems	8 (7.8%)	95 (92.2%)	
Unidimensional Model			
No problems	15 (34.9%)	28 (65.1%)	75.5%
One or both problems	8 (7.7%)	96 (92.3%)	

indicated that measures correlated within source (e.g., parent-rated shyness and parent-rated anxiety) or within construct (e.g., parent-rated shyness and teacher-rated shyness), but not both. As the correlations within source were much larger than those within construct, two factors were created (i.e., teacher-rated emotional problems and parent-rated emotional problems). Once these two factors were delineated, the model converged successfully. The goodness of fit index was .96, the comparative fit index was .90, and the chi-square was 16.11 ($df = 3$, $p < .01$). The decision to maintain both of these separate factors in subsequent analyses was supported by the low correlation between teacher rated emotional and parent-rated emotional functioning ($r = .29$).

Social-Behavioural Construct. One of the proposed indicators for the social-behavioural construct, (i.e., parent ratings of prosocial behaviour) did not load with the others, and was removed from the model. The remaining four indicators (teacher rated prosocial and disruptiveness, parent-rated disruptiveness and parent-rated presence of a behaviour problem) clustered together well to form a Social-Behavioural construct. The goodness of fit index was .96, the comparative fit index was .88 and the chi-square was 24.02 ($df = 2$, $p < .01$). Although parent-rated prosocial behaviour did not load with these other indicators, it was still included in the predictive model based on the theoretical importance of the construct. Parent-rated prosocial behaviour scores were standardized to be comparable to the latent construct scores for inclusion in the logistic regression model.

Cognitive and Language Constructs

Initial attempts to differentiate between a cognitive functioning and a language skills factor were unsuccessful. Standardized beta weights were as low as .01 (for the WPPSI block design loading on the cognitive factor). Once the indicators of cognitive functioning and language skills were combined into one confirmatory analysis, a unified latent construct emerged. The revised model had a goodness of fit statistic of .98, a comparative fit index of .92, and a chi-square of 7.70 ($df = 2$, $p < .05$).

The five resulting factors (Social-Behavioural Problems, Cognitive/Language Problems, Teacher-rated Emotional Problems, Parent-rated Emotional Problems, and Illness/Physical Limitations) were found to have low to moderate correlations with each other (-.19 to .28), with one exception, which was a correlation of .57 between the Social-Behavioural Problems construct and the Parent-rated Emotional Problems construct. The parent-rated prosocial

behaviour variable (included separately) was significantly correlated with the parent-rated emotional functioning construct (.23). The rest of the factors were negligibly correlated with parent-rated prosocial behaviour (ranging from -.03 to .09). These low correlation values support the inclusion of parent-rated prosocial behaviour as a separate construct.

Predicting Achievement and/or Behavioural Difficulties

Logistic regression analyses of the multidimensional constructs as well as the unidimensional measure on Grade 3 outcomes were used to generate classification summaries. In addition, forward stepwise logistic regression used with the multidimensional model served to evaluate the contribution of the individual constructs in the multidimensional model to the prediction of Grade 3 outcomes. A dichotomous Grade 3 outcome was used (i.e., the presence of any significant behaviour *and/or* achievement problem). The logistic regression analyses and all subsequent analyses were conducted with SPSS 12.0.

Results of the logistic regression analysis with the multidimensional model indicated that the Social-Behavioural Problems and the Cognitive/Language Problems factors were the only factors that uniquely contributed to the prediction of outcome at Grade 3. Beta weights were .90 (SE = .27) for the Social-Behavioural construct and .68 (SE = .26) for the Cognitive-Language construct. Both of the predictors added unique variance as demonstrated by significant changes in the Log Likelihood of the model. Results are outlined in Table 2.

These results were used to determine the classification accuracy for both models. That is, if decisions to identify a child as "at risk" were to be made on the basis of either of these measurement models, how accurate would these decisions be? The classification summaries for both models are fairly similar, as shown in Table 3. The overall classification for the multidimensional model was 73.3% compared to 75.5% for the unidimensional model.

Applying a Risk Identification Model

Epidemiological indices allow models to be compared in terms of the different types of prediction errors being made. In particular, the use of false positive and false negative rates to calculate indices of sensitivity and specificity generates a much more complete picture of the effectiveness of academic screening than does an overall classification rate on its own (Gredler, 1997). When compared on sensitivity (i.e., the proportion of children who were initially identified who had difficulties in Grade 3), both the multidimensional and unidimensional models indicated high sensitivity (i.e., over 0.9). Specificity indices (i.e., not identifying those children who will not have subsequent problems at high risk) indicated that both of these models tended to overidentify the extent to which children would have difficulties. In the multidimensional model, only one in three of the children who were performing adequately in Grade 3 escaped being identified as high risk at the time of school entry (i.e., specificity = .28). The unidimensional model had better specificity (i.e., .35); however, the majority of students not having problems in Grade 3 were still misclassified.

Discussion

This study grew out of the observation that research designed to influence policy in the field of academic readiness is progressing rapidly in a direction that is guided by two untested assumptions: 1) an inclusive and holistic child model is necessary to explain the development of academic functioning; and, 2) a risk model is an appropriate paradigm for identifying and understanding those children who will not be successful in the formal school setting. These assumptions in turn led to the current research questions:

Are each of the five child factors identified by Doherty (1997) and Kagan, Moore, and Bredekamp (1995) both independent and uniquely important in the prediction of school functioning?

How valid is identification at school entry based on a broadly defined holistic model for determining which individual children will struggle with the demands of the formal school setting?

Underlying these questions is the issue of whether or not the databases being designed to address these policy issues, with such complex and comprehensive models, contain sufficient depth to achieve their purpose.

Is the Five-Factor Model of Child Functioning Empirically Sound?

One of the major purposes of this project was to empirically evaluate a multidimensional model of academic readiness. In contrast to the original five factors, the results of the present study found that cognitive and language functioning merged into one factor, and emotional functioning required two factors differentiated by rater. The social-behavioural construct emerged as predicted.

Cognitive functioning and language skills could not be teased apart into two separate constructs. This

enmeshment likely reflects both the interrelated nature of the concepts, as well as limitations of the data. Administered cognitive tests were limited in scope (e.g., one subtest of the WPPSI rather than the whole test), and the main language test (i.e., the Peabody Picture Vocabulary Test-Revised) was one that is often used as an indicator of language and cognitive ability. This difficulty may be even more challenging in larger sets such as the NLSCY, which have very few child-administered measures due to the cost incurred by including these measures.

The emotional functioning construct was problematic for a different reason. In this case, the research does not support the notion of a unidimensional construct that has a linear relationship with overall adjustment. Although emotional functioning is discussed as a cohesive realm of functioning in the five-factor model (Doherty, 1997), current trends in the literature have focused on breaking the domain into more refined concepts with better clinical utility. The notion of emotional functioning as a heterogeneous construct was supported by the disparate ratings from parents and teachers. The conundrum is that emotional functioning concepts are good candidates to be included in a transactional, holistic child approach because emotional regulation as demonstrated by an inhibited or disinhibited style might mean something different in conjunction with different patterns of social and cognitive functioning. Including multiple components of emotional functioning in the model would more accurately capture the domain, but would also add greatly to the complexity of the model and difficulty interpreting relationships among the variables. The emotional functioning domain is a clear example of an area where using a general, overinclusive concept to represent several related components likely obscures important relationships among the variables.

Finally, the physical health and well-being concept was clearly a multidimensional concept from the start. When physical limitation was chosen as a construct that might affect academic functioning, it did not emerge as an important predictor of Grade 3 functioning. Thus, either the construct is not central to the concept of academic readiness (as implicated in the multidimensional model), or the salient part of the construct was not adequately represented in measures available in the current study.

The extent to which there were difficulties mapping the data onto the five factors of this model likely reflects both problems with the model and inadequacies in the data. The two are inextricably related. The data set being used to evaluate the model was representative of large research projects being compiled to generate policy data in areas such as academic readiness, so the problems with the data in this project would likely be present in other efforts to examine academic readiness in large samples.

Although reasonable representations of the five factors were generated, they did not all emerge as uniquely contributing to the prediction of Grade 3 difficulties. In attempting to understand the lack of significance for the physical health domain, the problem may be related to the lack of a normal distribution for that dimension. That is, while physical health may not be a crucial factor for the majority of children (i.e., most children may have "good enough" physical health), there could be a relatively small group of children for whom physical well-being is a crucial indicator of academic performance (e.g., children suffering from a chronic illness that might result in frequent school absences). This infrequent, but meaningful, subgroup would be obscured with a typical variable-oriented analysis. In comparison, the lack of significance of the emotional functioning domains could be a byproduct of the nonlinear nature of emotional inhibition; that is, both overregulation (i.e., an anxious presentation) *and* underregulation (i.e., impulsivity) are associated with problematic outcomes.

Clinical Implications of Using General Measures of Child Functioning for Risk Identification

Avaluating relative success of an attempt to predict which children are at risk for academic difficulties depends on the purpose that is under consideration. From a developmental research standpoint, both models were able to predict quite well which children would have trouble at Grade 3. Classification rates ranging from 70-75% are acceptable in many areas of research. The clinical utility of implementing the model, however, conveys a different picture.

The use of the epidemiological indices of sensitivity and specificity indicated that even if the multidimensional model could be operationalized as a screening procedure, a substantial number of children who would not develop problems might be misclassified as high-risk. In addition, the complexity of the model appeared superfluous when compared to a simple unidimensional screening measure. The unidimensional ABC had a comparable classification accuracy and better specificity (i.e., fewer false positive errors) than the multidimensional model. In addition, it has been noted that teacher ratings are less expensive and easier to administer than a multiinformant, multimethod screening model (Flanagan, et al., 2003). The relative success of the brief teacher-rated questionnaire compared to the more complex

model is consistent with the general literature on the validity of teacher ratings. An earlier review of the convergence between teacher ratings of achievement and objective scores (e.g., standardized testing) found teacher ratings to have reasonably high validity, irrespective of methodological differences in the studies included in the review (Hoge & Coladarci, 1989).

The reality of this sample of children living in disadvantaged communities was that 70% of these children were having difficulties in achievement, adjustment, or both by Grade 3. Does it make sense to use a risk-identification model when the majority of children will develop problems? When the prevalence of problems in Grade 3 is this high in a specific sample, it may be more prudent to shift from a risk identification paradigm to one of universal intervention. That is, rather than utilizing precious resources attempting to identify children who will be at risk, services could be offered to all of the children within the community. One potential direction for early universal intervention is the provision of high-quality preschool.

A U.S. study that selected 401 preschool programs (half for-profit and half not-for-profit) in four states and collected extensive data on the quality of these programs as well as teacher-child relationships, identified a positive impact for high-quality programs (Peisner-Feinberg et al., 1999). Results of the study suggested that the quality of preschool care for children had a significant impact on children's cognitive and social functioning over time. In this study, children who attended child care with higher-quality practices had better language and math skills from the preschool years into elementary school (Peisner-Feinberg et al.). The relationship between quality preschool and both cognitive and behaviour outcomes was particularly strong for children whose mothers had less education, suggesting that there was a differential impact of high-quality programs, in that more benefit was experienced by children who would be traditionally considered a high-risk group. Given the strong message from this well-designed study, there is growing evidence that universal approaches may be more efficacious, particularly in light of our limited ability to predict which children will have difficulties.

As long as the risk model is still in use, it is essential to incorporate indices that are used to evaluate success of identification of high-risk individuals (sensitivity, specificity, etc.) into our analyses. What constitutes adequate sensitivity and/or specificity can only be determined within the context of the possible outcomes. Depending on the cost (financial, emotional, resources, etc.) of the proposed outcome, and potential benefits and risks associated with it, different levels of sensitivity and specificity are acceptable. For example, if measures of academic readiness are being used to designate children who should be held back a year, than specificity under .3 is problematic (Gredler, 1997). In this scenario, only 30% of the children being held back would be accurately identified as high-risk, and the other 70% of children deemed high risk would be retained in kindergarten unnecessarily. If the identification is merely being used to indicate the need for ongoing developmental surveillance by school personnel (i.e., a much less intrusive outcome), then a lower specificity may be acceptable. These issues of sensitivity, specificity, and associated risks and benefits of proposed interventions should be included every time the risk/screening model is used, although in current practice such analyses rarely occur. The danger of the risk model is that it is extremely appealing at an intuitive level, while the underlying assumptions are often overlooked. For example, a survey of screening practices in 755 school districts in New York State found that 30% of respondents used locally developed instruments, and in 20% of districts, parents were encouraged to delay their child's kindergarten entry based on screening (Costenbader, Rohrer, & Difonzo, 2000). This study illustrated the potential danger of locally designed screening tools that lack the necessary psychometric properties being utilized by people who do not have the necessary appreciation of screening error rates.

Limitations

In the current study, the measurement of some of the individual child functioning factors was not optimal. In particular, the lack of a measure of expressive language may have detracted from the predictive validity of that construct. Previous work conducted to predict reading ability has highlighted the importance of multiple child-administered measures (Kirby & Parrila, 1999). However, while the measures were similar to those in most large-scale longitudinal projects, the inclusion of more child-administered tests might have improved the specificity of the models. Thus, while relying on the Block Design subtest of the WISC-III to define a cognitive functioning factor provides a narrow conceptualization of cognition, it is also a realistic reflection of the way such concepts are operationalized in large-scale research. Another limitation stems from the high-risk nature of the sample. Given that the sample was high risk in general, there may have been some limitations relating to a restricted measurement range with some of the variables. Future directions could include a similar analysis with a more representative sample.

In conclusion, although a multidimensional approach to academic readiness could logically be expected to provide better prediction than a unidimensional approach due to the increased complexity, the current analysis suggests that it is neither required nor particularly helpful in the first level of screening that occurs when children enter the formal school system. Thus, while it was possible to delineate multiple factors of child functioning at the time of school entry, this enterprise did not lead to an improved prediction model of academic functioning in Grade 3. The discrepancy between these findings and the popularity of applying risk prediction models indiscriminately to the general child population supports the need for further research in this area. Furthermore, models of screening need to be evaluated with respect to sensitivity and specificity, to evaluate their utility for decision making.

The authors acknowledge the assistance of Dr. Katreena Scott and the entire Better Beginnings, Better Futures research team.

Correspondence and reprint requests may be addressed to the first author at CRVAWC, Room 1118, Faculty of Education, University of Western Ontario, London, Ontario, Canada N6G 1G7 (E-mail: ccrooks@uwo.ca).

References

Arbuckle, J. L. (1997). *Amos user's guide version 3.6*. Chicago, IL: Small Waters Corporation.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science, 11*, 302-306.

Coie J. D., & Kupersmidt, J. B. (1983). A behavioral analysis of emerging social status in boys' groups. *Child Development, 54*, 1400-1416.

Costenbader, V. K., Rohrer, A. M., & Difonzo, N. (2000). Kindergarten screening: A survey of current practice. *Psychology in the Schools, 37*, 41-50.

Dodge, K. A. (1983). Behavioral antecedents of peer social status. *Child Development*, *54*, 1386-1389.

Doherty, G. (1997). *Zero to six: The basis for school readiness*. Research paper released by the Applied Research Branch of Human Resources Development Canada. Oxford, UK.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Fazio, B. B., Naremore, R. C., & Connell, P. J. (1996). Tracking children from poverty at risk for specific language impairments: A 3-year longitudinal study. *Journal of Speech, Language, and Hearing Research, 39*(3), 611-624.

Flanagan, K. S., Bierman, K. L., & Kam, C-M. (2003). Identifying at-risk children at school entry: The usefulness of multibehavioral problem profiles. *Journal of Clinical Child and Adolescent Psychology, 32*, 396-407.

Gredler, G. R. (1997). Issues in early childhood screening and assessment. *Psychology in the Schools, 34*, 99-106.

Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research, 59*, 297-313.

Huffman, L. C., Mehlinger, S. L., Kerivan, A. S. (2000). *Risk factors for academic and behavioral problems at the beginning of school*. Report commissioned by the Child Mental Health Foundation and Agency Network. Bethesda, MD.

Human Resources Development Canada/ Statistics Canada (1996). *Growing up in Canada: National longitudinal survey of children and youth*. Ottawa, ON: Statistics Canada.

Human Resources Development Canada/ Statistics Canada (undated). *National longitudinal survey of children and youth: User's handbook and microdata guide*. Ottawa, ON: Statistics Canada.

Kagan, S. L., Moore, E., & Bredekamp, S. (Eds.). (1995). *Reconsidering children's early development and learning: Toward common views and vocabulary*. Report released by the National Education Goals Panel. Washington, DC.

Kirby, J. R., & Parrila, R. K. (1999). Theory-based prediction of early reading. *Alberta Journal of Educational Research, 45*, 428-447.

Lipman, E. L., Offord, D. R., & Dooley, M. D. (1996). What do we know about children from single-mother families? Questions and answers from the National Longitudinal Survey of Children and Youth. In HRDC/Statistics Canada, *Growing up in Canada: National longitudinal survey of children and youth*. Ottawa, ON: Statistics Canada.

Offord, D. R., & Lipman, E. L. (1996). Emotional and behavioral problems. In HRDC/Statistics Canada, *Growing up in Canada: National Longitudinal Survey of Children and Youth.* Ottawa, ON: Statistics Canada.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., et al. (1999). *The children of the cost, quality, and outcomes study go to school: Executive summary.* Chapel Hill, NC: University of North Carolina and Chapel Hill, Frank Porter Graham Child Development Center.

Peters, R. DeV., Petrunka, K., & Arnold, R. (2003). The Better Beginnings, Better Futures Project: A universal, comprehensive, community-based prevention approach for primary school children and their families. *Journal of Clinical Child and Adolescent Psychology, 32*, 215-227.

Schwartz, D., McFadyen-Ketchum, S., Dodge, K. A., Pettit, G. S., & Bates, J. E. (1999). Early behavior problems as a predictor of later peer group victimization: Moderators and mediators in the pathways of social risk. *Journal of Abnormal Child Psychology, 27*(3), 191-201.

SPSS (2003). Standard Version 12.0.0. Copyright SPSS, Inc. 1989-2003. Chicago, IL.

Toronto Board of Education (1990). *ABC Kindergarten Teacher Observation Rating System.* Toronto, ON: Toronto Board of Education.

Tramontana, M. G., Hooper, S. R., & Selzer, S. C. (1988). Research on the preschool prediction of later academic achievement: A review. *Developmental Review, 8*, 89-146.

Tremblay, R. E., Masse, B., Perron, D., Leblanc, M., Schwartzman, A. E., & Ledingham, J. E. (1992). Early disruptive behavior, poor school achievement, delinquent behavior, and delinquent personality: Longitudinal analysis. *Journal of Consulting and Clinical Psychology, 60*, 64-72.

Tremblay, R.E., Vitaro, F., Gagnon, C., Piche, C., & Royer, N. (1992). A prosocial scale for the preschool behavior questionnaire: Concurrent and predictive correlates. *International Journal of Behavioral Development*, *15*(2), 227-245.

Vitaro, F., Brengden, M., & Tremblay, R. E. (1999). Prevention of school dropout through the reduction in disruptive behaviors and school failure in elementary school. *Journal of School Psychology, 37*(2), 205-226.

Wechsler, D. (1967). *Manual for the Wechsler Preschool and Primary Scale of Intelligence.* San Antonio, TX: The Psychological Corporation.

Received November 3, 2003
Revised July 20, 2004
Accepted October 7, 2004