

EXAMINING INTERACTIONS BETWEEN SOCIAL AND CULTURAL
CHILD AND NEIGHBORHOOD CHARACTERISTICS
ON CHILDREN'S DEVELOPMENTAL OUTCOMES:
STUDIES FROM THE POPULATION-BASED
EARLY DEVELOPMENT INSTRUMENT PROJECT

by
Martin Alex Guhn

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Human Development, Learning, and Culture)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2009

© Martin Alex Guhn, 2009

Abstract

This dissertation presents three manuscripts: The first manuscript presents a theoretical framework that integrates Bronfenbrenner's bioecological theory with theories from the school readiness and neighborhood effects literatures, and considerations from validity theory. The framework is applied to the Early Development Instrument project, for which data on Kindergarten children's developmental outcomes in the physical, social, emotional, cognitive, and communicative domains are obtained via teacher ratings on the Early Development Instrument (EDI, Janus & Offord, 2007). These data are linked, at a population level, to children's personal characteristics, health outcomes, academic achievement, and to social and cultural characteristics of their neighborhoods. The first manuscript illustrates how a comprehensive, integrative theoretical framework rooted in Bronfenbrenner's bioecological theory can lead to specific research design, research analysis, and validation research recommendations for the EDI project. The second manuscript presents differential item functioning analyses that examine whether the EDI measures (the construct of) *school readiness* in the same way across different groups of children (according to gender, English-as-a-Second-Language (ESL) status, and Aboriginal background). The analyses detected no item bias, thus presenting a critical prerequisite for further analyses that employ the EDI for examining group differences. The third manuscript presents a multilevel analysis that explores whether the theories of *double jeopardy* and *relative functionalism* can predict and explain interaction effects between socioeconomic status and gender or first language background (English, Punjabi, and Cantonese) on children's EDI scores. The findings show that (i) girls are rated higher than boys on all EDI domains, and that this gender gap is consistent across the

socioeconomic spectrum, (ii) group differences between ESL and non-ESL children on all domains (except for communication) can be attributed to group differences in socioeconomic status, (iii) the socioeconomic gradients for the English and Cantonese groups are equally pronounced, but significantly flatter for the Punjabi group, and (iv) the Cantonese group receives the relatively highest scores on the cognitive domain, and the relatively lowest scores on the communication domain. The concluding chapter highlights the theoretical and empirical significance of the research presented in this dissertation, addresses its strengths and limitations, and discusses implications for future analyses.

Table of Contents

Abstract.....	ii
Table of Contents.....	iv
List of Tables.....	vii
List of Figures.....	viii
Acknowledgements.....	x
Widmung.....	xi
Co-authorship Statement.....	xii
1. Introduction.....	1
School Readiness.....	1
The EDI Project.....	2
Theoretical Challenges.....	5
Two Empirical Studies.....	8
Purpose of Dissertation Research.....	11
References.....	13
2. A Bioecological Approach to Theory and Validity of a Population-Based Child Development Measure.....	19
Purpose of the Chapter.....	21
The Theoretical Construct of <i>School Readiness</i>	27
School Readiness: Links between Purposes, Theory, and Assessment.....	36
Sociological Views and Neighborhood Effects.....	39
The Bioecological Theory of Human Development.....	45

The EDI from a Bronfenbrennerian Perspective	50
Implications for Validation Research and Practice.....	57
Conclusion	60
References.....	63
3. Does the EDI measure school readiness in the same way across different groups of children?.....	74
Method	81
Results.....	88
Discussion.....	96
References.....	103
4. Language Background, Gender, and Socioeconomic Status Interactions Related to Kindergarten Children’s Development.....	107
Background.....	109
Method	123
Results.....	129
Discussion.....	139
References.....	149
5. Concluding chapter	154
Connections between the Presented Theoretical Framework and Empirical Studies.	154
Strengths and Limitations of the Presented Research.....	156
Working Hypotheses for Future Research.....	160
Future Directions	162
Contribution of the Dissertation to the Research Literature.....	165

References..... 167

List of Tables

Table 2.1 Developmental Domains of School Readiness.....	30
Table 3.1 Mean Group Differences (on a 10-point Scale) and Effect Sizes (Cohen’s d) ¹ for Group Comparisons.....	88
Table 3.2 Effect Sizes for EDI Items with Uniform DIF in the ESL versus non-ESL comparison.....	89
Table 4.1 Descriptive statistics (means, standard deviations) of the EDI scores and socioeconomic status, for girls and boys of the non-ESL group, and the ESL (Punjabi-, Cantonese-speaking) and non-ESL groups.....	128
Table 4.2 Summary of multilevel modeling results (coefficients, standard errors, and significant levels) for the gender and language group analyses	130
Table 4.3 Summary of multilevel modeling results (coefficients, standard errors, and significant levels) for the gender and language group analyses	133

List of Figures

Figure 2.1 Three nomological networks representing common disciplinary approaches to human development.....	24
Figure 2.2 Disciplinary foci in human development research.....	31
Figure 2.3 Illustration of the bioecological theory of human development as integrative framework.....	54
Figure 2.4 A Bronfenbrennerian nomological network for human development.....	56
Figure 3.1 Item ‘Ability to use English effectively’, on the Communication Skills and General Knowledge domain, displaying DIF in ESL comparison.....	91
Figure 3.2 Differential Domain Functioning (Communication Skills and General Knowledge domain) due to (cumulative) item-level DIF.....	92
Figure 3.3 Item ‘Gets into physical fights’, on the Emotional Maturity domain, displaying DIF in gender comparison.....	94
Figure 3.4 No Differential Domain Functioning (on Emotional Maturity domain), despite item-level DIF.....	95
Figure 4.1 Gender by socioeconomic status interactions (Social competence domain).	131
Figure 4.2 Language group by socioeconomic status interactions (Physical health and well-being domain).....	134
Figure 4.3 Language group by socioeconomic status interactions (Social competence domain).....	135
Figure 4.4 Language group by socioeconomic status interactions (Emotional maturity domain).....	135

Figure 4.5 Language group by socioeconomic status interactions (Language and cognitive development domain).....	136
Figure 4.6 Language group by socioeconomic status interactions (Communication skills and general knowledge domain).....	136
Figure 4.7 Scatterplots for the relationship between developmental outcome and socioeconomic status for the English, Punjabi, and Cantonese groups (Social competence domain).....	138

Acknowledgements

My sincerest thanks go to my supervisor Hillel Goelman, and my committee members, Bruno D. Zumbo and Kim Schonert-Reichl, who have taught, motivated, supervised, guided, and supported me throughout my time at UBC, throughout my dissertation research over the past five years, through full and wonderful years in Vancouver, and through the difficult times when they arose.

I would also like to sincerely thank Clyde Hertzman and the Human Early Learning Partnership for the continuous personal, intellectual, technical, and financial support.

Also, I am grateful to my friends and my family, my siblings Anna and Jakob, and especially my parents Annette and Ekkehard, and most of all, Anne, who have been (t)here for me not only during the years of my PhD, but also always before. Much of what has motivated, energized, led up to, and gone into this dissertation is rooted in their unconditional love, and the experience of shared joy, creativity, curiosity, laughter, questions, play, conversations, hope, thoughts, patience, and closeness.

Widmung

Voller Dankbarkeit und Liebe

für meine Eltern,

& meine Anne!

Co-authorship Statement

In this dissertation, the introduction and the concluding chapter have been written by myself. I have been the lead author of chapter 2, which has been co-authored with Hillel Goelman, who contributed to the conceptual, theoretical, and semantic coherence and clarity of the chapter. Likewise, I have been the lead author of Chapter 3, which has been co-authored with Anne Gadermann and Bruno D. Zumbo, who contributed to the formulation of the research question, the syntax for the statistical analyses, and assisted with the interpretation of the empirical results. Finally, I have been the lead author for chapter 4, which is a co-authored manuscript with Anne Gadermann, Clyde Hertzman, and Bruno D. Zumbo, who contributed with methodological advice and assistance with the interpretation of statistical findings.

1. Introduction

“There is nothing as practical as a good theory.”

(Kurt Lewin, quoted in Bronfenbrenner, 1978, p. 43)

School Readiness

Over the past two decades, the construct of *school readiness* has had increased visibility in the arenas of public education, educational and developmental research, and politics in North America. This has been spurred by a combination of different factors. For decades, research has shown that the transition to school represents a critical period with respect to children’s social, emotional, and cognitive adjustment and development (e.g., Alexander & Entwisle, 1982; Pianta & Cox, 1999). In addition, research from multiple disciplines has accumulated evidence showing the critical relationships between experiences during the early years and later developmental outcomes in the domains of education, health, well-being, and social competencies (e.g., Doherty, 1997; Keating & Hertzman, 1999; Shonkoff & Phillips, 2000). At the same time, large inequalities with regard to children’s early developmental outcomes as reflected in school readiness as well as health assessments have existed with some consistency in the North American context, and these inequalities in developmental outcomes are highly correlated with inequalities in socioeconomic status (e.g., Berliner, 2005; Jencks & Mayer, 1990; Kohen, Brooks-Gunn, Leventhal, & Hertzman, 2002).

Making references to this research knowledge, political speeches and government documents in Canada and the US have, over the past years, gained new momentum in

proclaiming the fundamental importance of school readiness, its relation to the early years, and its significance for later social, academic, and professional accomplishments (e.g., Governor General of Canada, 1997; National Education Goals Panel, 1995). This renewed and increased political focus on school readiness issues has, in turn, affected policies which govern practices in schools and communities (Kagan & Rigby, 2003), particularly with regard to issues of accountability related to children's developmental outcomes in general, and their school readiness and academic achievement (or improvement) in particular (e.g., No Child Left Behind in the US). The combination of these factors has facilitated the emergence of numerous initiatives (e.g., the Understanding the Early Years project in Canada; Nosbush, 2006) aimed at fostering, assessing, and better understanding children's school readiness. One such initiative, briefly described in the following section, is theoretical and empirical work pertaining to the Early Development Instrument, which is presented in this dissertation.

The EDI Project

In British Columbia, the Human Early Learning Partnership is an interdisciplinary research consortium based at the University of British Columbia, with the mission "to create new knowledge to help children thrive". One of the main research initiatives at the Human Early Learning Partnership is a project to which I will refer in this context as the *EDI (Early Development Instrument) project*. The EDI project is, in essence, a population-level effort to monitor children's well-being in a number of developmental domains, and to examine in what way children's developmental trajectories are related to characteristics of the environments in which they grow up.

Within this research project, the *Early Development Instrument* (EDI; Janus & Offord, 2007), a teacher-administered school readiness assessment tool, represents a primary component of the data collection initiative (hence the name for the project). With regard to the EDI, a number of points are important to note. First, the EDI is designed to assess children's developmental status as reflected in their *school readiness* in five developmental domains: physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge). Second, the EDI is referred to as a *community tool*, and is associated with the purpose of collecting information on children's development (at a community level) that can be used to inform discussions, decisions, and practices with regard to fostering the well-being of children. It is equally important to note, however, that—according to the EDI developers—information obtained via the EDI should ideally be used in combination with other indicators of children's development and information on the social, demographic, and socioeconomic characteristics of the environment (Janus & Offord, 2007). In this approach, Janus and Offord (2007) clearly follow recommendations for a community approach toward supporting children's school readiness as delineated by Love, Aber, and Brooks-Gunn (1994).

A community, or population-health approach to school readiness is, in some ways, a procedure that breaks new ground, and presents a number of challenges. It has thus been a priority to obtain evidence for the validity of this approach. As part of this process, traditional notions of test reliability and validity have been examined (see Janus et al., 2007): The internal consistencies (Cronbach's α) for the EDI's five domains range from $\alpha = .84$ to $\alpha = .96$, test-retest reliabilities for the five domains from $r = .82$ to $r =$

.94, and inter rater reliabilities for the five domains from $r = .53$ to $r = .80$ in a study with children's school and daycare teachers, and from $r = .36$ to $r = .64$ in a study with the children's parents and teachers. Correlations between EDI domain scores and other childhood development or competence measures show convergent and discriminant validity. For example, the communication skills domain of the EDI has a correlation of $r = .57$ with the Peabody Picture Vocabulary Test, whereas the other four domains have much lower correlations ($r = .05$ to $r = .26$). The predictive validity (for later academic achievement) of the EDI has also been examined, showing that the EDI does as well as other school readiness measures in predicting achievement scores or grades. (For further details on validity studies, see Guhn, Janus, & Hertzman, 2007; Janus et al., 2007). As will be described in more detail later, this dissertation and other ongoing work address further aspects of validity. Important questions in this regard are (i) whether the EDI validly captures the construct of school readiness, in a theoretically coherent way (cf. Cronbach & Meehl, 1955), (ii) whether social consequences resulting from the EDI measurement procedure are in line with the purposes of the EDI (cf. Messick, 1995), and (iii) whether the EDI is a valid school readiness measure at the community level (Zumbo & Forer, in press).

In order to complement the information from the EDI with information on community characteristics, the Human Early Learning Partnership has undertaken a concerted data collection effort linking children's developmental trajectories to a number of social and cultural person and context characteristics (e.g., children's gender, family socioeconomic status) at a population level. At the moment, the database consists of data on five children's developmental domains for two entire cohorts (2003 & 2006) of

Kindergarten children in British Columbia (each cohort containing records for approximately 40,000 children). For one pilot cohort (of approximately 10,000 children), the database now also contains a longitudinal component, namely information on the (same) children's academic achievement in grade 4 and grade 7 (cf. Lloyd & Hertzman, 2009). In addition to these data at the individual child level, the database contains a wide range of demographic and socioeconomic indicator variables that are aggregated at a neighborhood level. The database thus allows one to examine the relationship between children's developmental status and demographic and socioeconomic characteristics of their neighborhoods. It is envisioned that, eventually, the database can be complemented by (that is, linked to) information on children's developmental outcomes at an earlier age than Kindergarten and a later age than grade 7, in a number of developmental domains (e.g., health, educational attainment, social competences, well-being, community participation). This brief description illustrates how the EDI project will, in the long run, be able to not only monitor children's developmental outcomes over time, but also to examine to what extent developmental trajectories and patterns for different groups of children are related to neighborhood characteristics.

Theoretical Challenges

With regard to its theoretical approach, the Human Early Learning Partnership draws upon Bronfenbrenner's bioecological theory of human development (e.g., Bronfenbrenner, 1979; Bronfenbrenner, 1992; Bronfenbrenner & Morris, 2006) for the EDI project. In addition, the EDI (Janus & Offord, 2007) is also informed by theoretical considerations drawn from the school readiness literature and the neighborhood effects

literature (e.g., Chase-Lansdale, Gordon, Brooks-Gunn, & Klebanov, 1997; Doherty, 1997; Love, Aber, & Brooks-Gunn, 1994; Meisels, 1999).

Bronfenbrenner's bioecological theory of human development is an appropriate theoretical framework for the EDI project, given its ecological nature. One of the propositions that the theory has consistently been emphasizing is that children's developmental outcomes are influenced by a range of interacting processes that occur between the developing child and the persons and objects in its environment. Based on this proposition, it has thus been argued that certain human development research questions need to be addressed by simultaneously taking into account the characteristics (i) of those processes (e.g., parenting style), (ii) of the developing person (e.g., gender, temperament), (iii) of developmentally relevant context factors (e.g., social expectations, quality of the parental work environment), and, finally, characteristics pertaining to time (e.g., when and in what order those processes occur; Bronfenbrenner & Morris, 2006). The theory is widely recognized for its conceptual appeal (reference). At the same time, the complex nature of the theory—and the phenomena of human development it is trying to capture—make it difficult to translate it into concrete research designs. Also, the theory has considerably evolved since it was first formulated in the 1970s (Bronfenbrenner, 1979; Bronfenbrenner, 1992, Bronfenbrenner & Morris, 2006). Therefore, the theory will be reviewed in some detail in Chapter 2 (pages 44-48) and then applied to the EDI (pages 49-59).

In light of the EDI's intended use and purpose, it is similarly critical to firmly base the EDI in the school readiness and neighborhood effects literatures. This scenario entails great challenges. It calls for an explicit recognition of a complex theoretical

framework (i.e., Bronfenbrenner's bioecological theory of human development) and of the complex theoretical research literatures on school readiness and neighborhood effects. More importantly, it also raises the question of whether these different theoretical frameworks and approaches are, in fact, compatible with and complementary to each other. Bronfenbrenner himself anticipated that a comprehensive application of his theory's principles to research designs, to interpretations of empirical findings, and to translation of research knowledge into practice would be challenging (Bronfenbrenner & Morris, 2006). Similar conclusions have been stated with respect to the school readiness construct (e.g., Meisels, 1999) and theories pertaining to neighborhood effects (e.g., Jencks & Mayer, 1990; Sampson, Morenoff, & Gannon-Rowley, 2002). This is not necessarily indicative of any shortcomings of the theories, but simply a reflection of the complex nature of human development. Contributing to this challenge is the fact that Bronfenbrenner's bioecological theory of human development is, at once, both a developmental theory and a meta-theory or conceptual framework—or, in Bronfenbrenner's words, a 'perspective', or 'paradigm'. That is, on the one hand, the theory contains very specific hypotheses regarding human development, and on the other hand, general principles that pertain to methodological considerations for conducting developmental research as well as to further development and refinement of the theory itself.

Given these theoretical challenges, and in light of the scope and complexity of the EDI project, it is critical to revisit the bioecological theory of human development, and to integrate it with theoretical and methodological considerations from the school readiness and neighborhood effects literature, in order to maximize its utility for applied

developmental research. Accordingly, chapter 2 proposes a theoretical framework for the EDI project, integrating the bioecological theory of human development with theoretical considerations from the school readiness literature (e.g., Doherty, 1997; Graue, 1995) and the neighborhood literature (e.g., Brooks-Gunn, Duncan, Klebanov, & Sealand, 1993; Sampson, Morenoff, & Gannon-Rowley, 2002). The chapter also examines the theoretical and methodological considerations drawn from seminal writings on construct validity (Cronbach & Meehl, 1955) and test validity (Messick, 1995). Based on the theoretical integration, chapter 2 also addresses a number of specific recommendations for research design, research methodologies, and validation research pertaining to the EDI project. Chapter 2 thus presents part of an ongoing attempt to operationalize Bronfenbrenner's bioecological theory of human development for the EDI project and the construct of school readiness (cf. Rimm-Kaufman & Pianta, 2000).

Two Empirical Studies

Chapters 3 and 4 of this dissertation present two empirical studies pertaining to the EDI project. Both empirical studies are theoretically linked to the theoretical framework proposed in the first manuscript. Also, the two empirical studies are related to each other with regard to their research question, in that they both explore gender differences as well as English-as-a-Second-Language (ESL) versus non-ESL differences on the EDI. The first empirical study presents a set of differential item functioning analyses that examined whether the EDI measures (the construct of) school readiness in the same way across different groups of children (according to gender, ESL status, and children Aboriginal background). A differential item functioning analysis, in essence,

examines whether examinees (i.e., those who obtain scores on a measurement scale) from different groups have different probabilities of success on (or endorsing) an item after matching on the underlying construct that the item is intended to measure (Zumbo, 1999). Such an analysis is critical if one intends to conduct meaningful group comparisons with regard to the construct that is measured by the scale, as it allows one to assess whether group differences on the EDI are potentially related to the way the construct is measured. In other words, a DIF analysis provides information on whether the EDI is potentially biased against certain subpopulations of Kindergarten children.

A couple of key findings from that study are presented here already, because they substantially informed the research questions examined in the subsequent empirical analyses (which will be presented in chapter 4). In the gender comparison, the differential item functioning analyses showed that only one EDI item—on physical aggression—had differential item functioning (i.e., boys are rated as more physically aggressive by their Kindergarten teachers than girls, even after controlling for the children’s overall level of school readiness). For the ESL versus non-ESL comparison, the analyses indicated that ESL children—at comparable levels of their respective overall school readiness—obtain systematically lower scores on six of eight items on the communication skills and general knowledge domain. For the items on the other four developmental domains assessed by the EDI, no differential item functioning was found for the ESL analysis. At the same time, the EDI data showed significant overall group differences between girls and boys as well as between ESL and non-ESL children on all five EDI domains—and according to the DIF analyses it could be ruled out that those differences were caused by item bias.

The second empirical study, a set of multilevel analyses presented in chapter 4, therefore explored a further aspect of the gender differences as well as differences between ESL and non-ESL children. This was done, because a better understanding of gender differences and ESL versus non-ESL differences is of particular relevance for educators who would like to adequately respond to such differences. The focus of the multilevel analyses was on the following questions. First, are the gender differences consistent across all socioeconomic groups? Previous research in this area provides mixed evidence, in that there is no consistent pattern with regard to how gender and socioeconomic status interact with regard to children's early developmental outcomes (e.g., Duncan et al., 2007; Entwisle, Alexander, & Olson, 2005). As noted by Entwisle et al. (2005), this issue has not been given rigorous theoretical attention, and their study represents a rare exception in that regard. Therefore, this study is aimed at producing additional empirical evidence to answer this question and to also add to the theoretical discussion regarding this topic.

The second question is whether the ESL versus non-ESL group differences are consistent across the socioeconomic spectrum. Empirical evidence in regard to this question is even more difficult to interpret than with regard to gender differences, because the operationalization or definition of ESL status is not a universal one. In fact, immigrant status, minority status, ethnic background, and first language have all been used as criterion for group comparisons in previous studies. Adding to the complexity is the fact group composition is, naturally, very context-dependent with regard to its ethno-cultural nature (e.g., Hispanic, Asian-American), size (e.g., percentage of the overall population) and its characteristics (e.g., homogeneity versus heterogeneity of the group).

This analysis subdivides the one heterogeneous group of all ESL children into two more homogeneous groups (Punjabi and Cantonese, which are the largest ESL groups in British Columbia) to investigate whether ESL versus non-ESL group differences are consistent across different ESL groups.

In the concluding chapter, several key issues are addressed. First, the strengths and limitations of the research presented in this dissertation are discussed. Based on the discussion, working hypotheses and research ideas, as well as suggestions pertaining to the research design and analyses of future studies are presented. In particular, a research design for the EDI project is sketched out, according to which one could systematically examine how developmentally relevant processes interact, over time, with children's gender and language background and the social and cultural characteristics of their environmental contexts in regard to their developmental outcomes. The last chapter then concludes with a discussion of the theoretical and empirical contributions and their significance of the presented research to the research literature.

Purpose of Dissertation Research

This introductory chapter concludes with an attempt to respond to the question, 'What is the significance of this research?' In British Columbia, as in many other regions around the globe, families, educators, community practitioners, policy makers, and others are continuously confronted with raising children in contexts, within which gender roles, expectations, and opportunities continue to change, and within which (im)migration

patterns of different ethno-cultural groups continue to enrich the multicultural makeup of society. As postulated by Bronfenbrenner, creating scientific knowledge that can be translated and applied to decisions and practices intended to support families, schools, and communities in their efforts to foster children's well-being can thus never be considered as 'done', but remains an ongoing necessity (Bronfenbrenner & Morris, 2006; cf. Messick, 1995). In this regard, the EDI project is well positioned to significantly continuously contribute to a better understanding of children's development in British Columbia. It is hoped that the empirical work presented in this dissertation contributes to this undertaking, and that, in paying homage to Urie Bronfenbrenner and his mentor, Kurt Lewin, the proposed integrated theoretical framework will function as *'a theory good enough to be practical'*.

References

- Alexander, K. L., & Entwisle, D. R. (1988). Achievement in the first two years of school: Patterns and processes. *Monographs of the Society for Research in Child Development, 53*, 1-157.
- Berliner, D. C. (2005). Our impoverished view of educational reform. *Teachers College Record*. Retrieved on November 20, 2006, from <http://www.tcrecord.org/content.asp?contentid=12106>
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U. (1992). Ecological systems theory. In R. Vasta (Ed.), *Six theories of child development: revised formulations and current issues* (pp. 187-249). London, UK: Jessica Kingsley Publishing.
- Bronfenbrenner, U., & Morris, P. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Development, Volume 1, Theoretical Models of Human Development*, (6th ed.), 793-828. Hoboken, NJ: John Wiley & Sons.
- Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K., & Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology, 99*, 353–395.
- Chase-Lansdale, P. L., Gordon, R.A., Brooks-Gunn, J., & Klebanov, P. K. (1997). Neighborhood and family influences on the intellectual and behavioral competence of preschool and early school-age children. In G. J. Brooks-Gunn & J. L. Duncan (Eds.), *Neighborhood Poverty* (pp. 79-118). New York, NY: Russell

Sage Foundation.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests.

Psychological Bulletin, 52, 281-302.

Doherty, G. (1997). *Zero to six: The basis for school readiness* (No. R-97-8E). Ottawa,

ON: Human Resources Development Canada, Applied Research Branch, Strategic Policy.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P.,

et al. (2007). School readiness and later achievement. *Developmental Psychology*, 43, 1428–1446.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap

of being poor and male. *The Sociology of Education*, 80, 114-138.

Governor General of Canada (September, 1997). *Speech from the Throne*. Retrieved on

November 20, 2007, from <http://www.gg.ca/media/doc.asp?lang=e&DocID=190>

Graue, M. E. (1995). Connecting visions of authentic assessment to the realities of

educational practice. In T. A. Romberg (Ed.), *Reform in school mathematics and authentic assessment*, pp. 260–275. Albany, NY: State University of New York Press.

Guhn, M., Janus, M., & Hertzman, C. (2007). The Early Development Instrument:

Translating school readiness assessment into community actions and policy planning. *Early Education & Development*, 18, 369-374.

Janus, M., Brinkman, S., Duku, E., Hertzman, C., Santos, R., Sayers, M., et al. (2007).

The Early Development Instrument: A population-based measure for communities. A handbook on development, properties, and use. Hamilton, ON:

- Offord Centre for Child Studies. Retrieved on June 10, 2008, from http://www.offordcentre.com/readiness/pubs/2007_12_FINAL.EDI.HANDBOOK.pdf
- 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*, 1–22.
- Jencks, C., & Mayer, S. (1990). The social consequences of growing up in a poor neighbourhood. In L. E. Lynn & M. G. H. McGeary (Eds.), *Inner-city poverty in the United States* (pp. 111–186). Washington, DC: National Academy Press.
- Kagan, S. L., & Rigby, E. (2003). *Policy matters: Setting and measuring benchmarks for state policies. Improving the readiness of children for school: Recommendations for state policy*. Washington, DC: Center for the Study of Social Policy.
- Keating, D. P., & Hertzman, C. (Eds.) (1999). *Developmental Health and Wealth of Nations*. New York, NY: Guilford.
- Kohen, D. E., Brooks-Gunn, J., Leventhal, T., & Hertzman, C. (2002). Neighborhood income and physical and social disorder in Canada: Associations with young children's competencies. *Child Development, 73*, 1844-1860.
- Kohen, D. E., Hertzman, C., & Brooks-Gunn, J. (1998). *Neighborhood influences on children's school readiness*. Hull, QB: Human Resources Development Canada.
- La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research, 70*, 443-484.

- LaPointe, V. R., Ford, L., & Zumbo, B. D. (2007). Examining the relationship between neighborhood environment and school readiness for Kindergarten children. *Early Education & Development, 18*, 473-495.
- Lewit, E. M., & Schuurmann Baker, L. (1995) School readiness. *The Future of Children, 5*, 128-139.
- Love, J. M., Aber, J. L., & Brooks-Gunn, J. (1994). *Strategies for assessing community progress toward achieving the First National Educational Goal*. Princeton, NJ: Mathematica Policy Research.
- Meisels, S. (1999). Assessing readiness. In R. C. Pianta, & M. J. Cox (Eds.), *The transition to Kindergarten* (pp. 39-66). Baltimore, MD: Paul H. Brookes.
- Messick, S. (1995). Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.
- NAEYC (1996). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8*. Washington, DC: National Association for the Education of Young Children. Retrieved on November 20, 2007, from <http://www.naeyc.org/about/positions/pdf/PSDAP98.pdf>
- National Education Goals Panel (1995). *The National Education Goals report*. Washington, DC: US Government. Retrieved on November 20, 2007, from http://eric.ed.gov:80/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/41/7e.pdf

- Nosbush, L. (2006). *Community report for children. The Understanding the Early Years final report*. Prince Albert, SK: Saskatchewan Rivers School Division. Retrieved on November, 2007, from <http://www.srsd119.ca/uey/>
- Pelletier, J., & Corter, C. (2005). Design, implementation, and outcomes of a school readiness program for diverse families. *The School Community Journal*, 89-116.
- Pianta, R. C., & Cox, M. J. (Eds.) (1999). *The transition to Kindergarten*. Baltimore, MD: Paul H. Brookes.
- Pianta, R. C., Cox, M. J., & Snow, K. L. (Eds.) (2007). *School readiness and the transition to Kindergarten in the era of accountability*. Baltimore, MD: Paul H. Brookes.
- Piotrkowski, C. S. (2004). A community-based approach to school readiness in Head Start. In E. Zigler & S. Styfco (Eds). *The Head Start debates* (pp. 129-142). New Haven: Yale University Press.
- Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An ecological perspective on the transition to Kindergarten: A theoretical framework to guide empirical research. *Journal of Applied Developmental Psychology*, 21, 491-511.
- Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing “neighborhood effects”: social processes and new directions in research. *Annual Review of Sociology*, 28, 443-478.
- Schrag, R. D. A., Styfco, S. J., & Zigler, E. (2004). Familiar concept, new name: Social competence/school readiness as the goal of Head Start. In E. Zigler & S. J. Styfco (Eds.), *The Head Start debates*. Baltimore, MD: Paul H. Brookes.

- Shonkoff, J., & Phillips, D. A. (Eds.) (2000). *From neurons to neighborhoods. The science of early childhood development*. Washington, DC: National Academy of Sciences.
- Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press.
- Zigler, E., & Valentine, J. (1979). *Project Head Start: A legacy of the war on poverty*. New York, NY: The Free Press.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.
- 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, D.C.: American Psychological Association Press.

2. A Bioecological Approach to Theory and Validity of a Population-Based Child Development Measure¹

'Kasserian ingera?'—'How are the children?'—is a traditional Massai greeting that reflects societal concern for children's well-being. This concern is similarly reflected in numerous current national and international indicators of children's well-being, which ask the same question: "How are the children?"

The traditional response to the Massai greeting is: "All the children are well". However, a range of indicators of child development consistently shows that not all children are well. In the Canadian context, surveys and statistics show that the prevalence of mental health/psychiatric disorders among children is about 10-20 % (Breton et al., 1999; Offord et al., 1987; Spady, Schopflocher, Svenson, & Thompson, 2001), that approximately 13% of all children are obese (Tremblay & Willms, 2000), and that provincial child poverty rates range from approximately 5-20% (CCSD, 2004).

Assessing and obtaining statistics on children's well-being or developmental status is not, however, an end in itself. Rather, it is one component in describing and understanding phenomena in human development. The scope and complexity of understanding human development is illustrated by Bronfenbrenner's concept of the *process-person-context-time model*. In essence, this model proposes that understanding human development requires the theoretically guided study of what *processes* lead, over *time*, to what developmental outcome for what *person* in what specific *context*. Such an

¹ A version of this chapter has been submitted for publication. Guhn, M. & Goelman, H. A bioecological approach to theory and validity of a population-based child development measure.

understanding, however, is still only a part of a larger objective, namely that of generating an understanding of human development that can, eventually, be used to make adequate decisions and implement practices, so as to increase the number of children who are doing well (Bronfenbrenner & Morris, 2006).

In line with Bronfenbrenner's holistic view argument, creating useful research knowledge is the purpose of the Early Development Instrument (EDI; Janus & Offord, 2000; 2007; Janus et al. 2007), a measure administered by teachers in the middle of the Kindergarten year to assess Kindergarten children's developmental status as reflected in their *school readiness* in five domains: physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge. One of the conceived usages of the EDI is, through linking EDI data with other child development data (e.g., pre-school screenings; school achievement) and other social indicators (e.g., socioeconomic status; social capital) at the neighborhood level, to provide communities with knowledge on children's development and their social and ecological context. This knowledge, in turn, is intended to inform families, practitioners, educators, researchers, and policy makers with respect to their practices and decisions in support of children's and families' well-being (see also Nosbush, 2006).

The EDI has been implemented at a population-level in several jurisdictions in Canada and abroad. In British Columbia, for example, the implementation has been conducted by the Human Early Learning Partnership at the University of British Columbia, in collaboration with the Ministries on Health, Education, and Child and Family Development. To date, data from the EDI have been collected for two cohorts of

about 40,000 Kindergarten children (per cohort) across British Columbia, and these data have been linked to demographic and socioeconomic census data, aggregated at a neighborhood level. In the following, I will refer to this research initiative in British Columbia as the *EDI project (BC)*².

Purpose of the Chapter

One important question is whether the EDI provides valid and reliable data on children's developmental status regarding their school readiness. This question can be approached from numerous directions, as 'validity' refers to a broad category of concepts. That is, there are numerous types of validity, such as construct validity, measurement validity, predictive and concurrent, convergent and discriminant validity, content validity, and test validity (see Kane, 2006, for an overview). In this paper, issues pertaining primarily to construct validation and test validation³ of the EDI are addressed. Given the numerous, and sometimes drastically different (compare, for example, Borsboom, Mellenbergh, & van Heerden, 2004, with Messick, 1995) interpretations and the evolving meaning of the terms (Kane, 2006), the definitions used in this context are provided first. These definitions primarily refer (i) to Cronbach and Meehl's (1955) work on theoretical construct validity and validation, and (ii) to Messick's (1995) work on integrating test-use-based social consequences into a holistic view of test validity.

In their seminal paper on construct validity in psychological tests, Cronbach and

² For information on the EDI project in British Columbia, please visit www.earlylearning.ubc.ca. For information on EDI research and projects in Ontario, see www.offordcentre.com/readiness/index.html, and for Australia, see www.rch.org.au/australianedi and www.ichr.uwa.edu.au/research/highlights/early/aedi.

³ Here, the term 'validation' is used, as it refers to the process and methods via which continuous, context-specific research seeks to establish construct and test 'validity'.

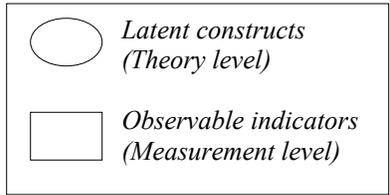
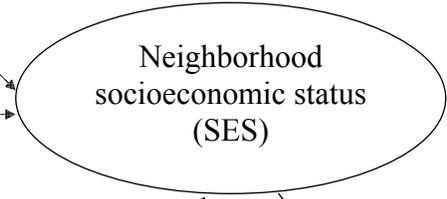
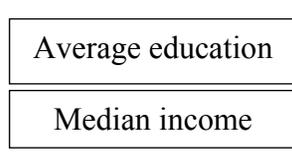
Meehl (1955) define construct validation as the process of testing hypotheses set forth by theory—a nomological network of associations and propositions—pertaining to the construct. As psychological constructs themselves are, by definition, unobservable, construct validation needs to occur via measurements of observable indicators (of the respective constructs) within the nomological network, and the relationships between these observables need to adhere to the theoretical predictions. Alternatively, in some cases, the nomological network might have to be refined or expanded in order to accommodate accumulating measurement evidence. Accordingly, theory building and construct validation are iteratively intertwined (see also Smith, 2005).

Figure 2.1 illustrates three examples of nomological networks for the school readiness construct. In each of the three network examples, theoretically proposed relationships (represented as arrows) among the latent constructs (the theory level; represented as circles) and the observable indicators (the measurement level; represented as rectangles) are shown⁴. The three examples represent three different approaches to conceptualizing and studying school readiness: Example 1 (top of figure) represents a scenario, which is often employed in sociology, in which a developmental outcome (in this case, school readiness) is examined in its relationship to neighborhood characteristics (in this case, socioeconomic status of the neighborhood, simultaneously taking individual families' socioeconomic status into account). In example 2 (middle of figure), which shows a typical scenario from an educational, psychological, or intervention approach,

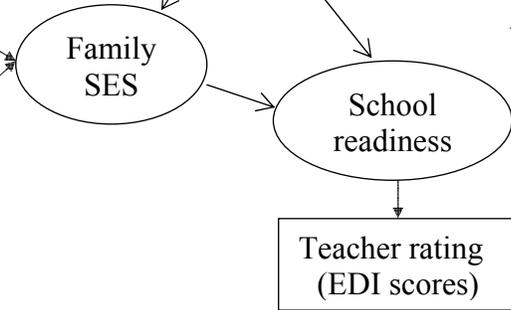
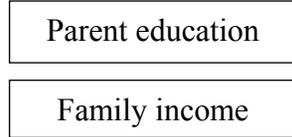
⁴ The illustrations adhere to conventional structural equation model, or path analysis, notation, and follow Lennox and Bollen's (1991) recommendations for distinguishing between *composite indicators*, such as socioeconomic status, and *effect indicators*, such as school readiness and academic achievement, for both of which the arrows represent, respectively, the direction of association. For associations among constructs, arrows with open-shaped arrow heads are used, and for relationships between measures and constructs, arrows with solid heads are used.

the developmental outcome is studied in the context of developmentally relevant experiences; in this case, parenting and/or teaching practices. In example 3, the nomological network represents a longitudinal approach from a maturational, or individual differences perspective, conceptualizing the developmental outcome as the outcome of genetic and biological factors, as well as a long-term predictor of later developmental outcomes (for example, academic achievement).

Neighborhood level¹

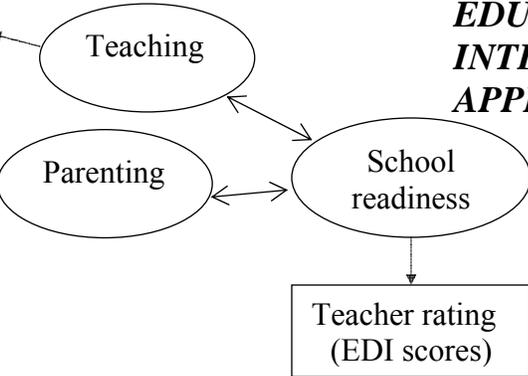
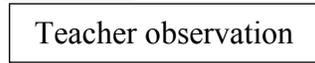


Family level¹



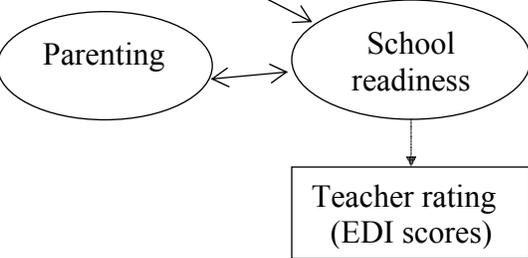
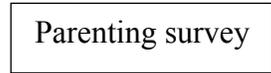
SOCIOLOGICAL APPROACH

School level¹



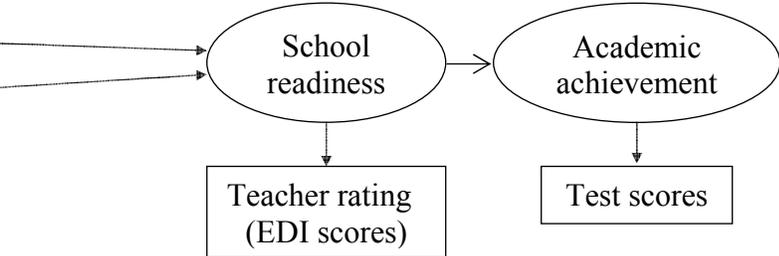
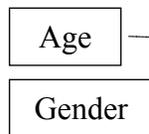
PSYCHOLOGICAL-EDUCATIONAL-INTERVENTIONAL APPROACH

Family level¹



MATURATIONAL-INDIVIDUAL DIFFERENCES APPROACH

Child level¹



¹Level of measurement at which data are obtained

Figure 2.1 Three nomological networks representing common disciplinary approaches to human development.

In the school readiness literature(s), all three approaches, and variations of them, can be found. Taking a sociological approach, there are a small number of Canadian studies that examined the relationship between socioeconomic neighborhood characteristics and school readiness (e.g., Kohen, Hertzman, & Brooks-Gunn, 1998) and early child development outcomes (Kohen, Brooks-Gunn, Leventhal, & Hertzman, 2002). Prominent examples from the educational preschool intervention field are the *Head Start* program (Zigler, Gilliam, & Jones, 2006) and the Perry/High Scope Preschool program from the US (Schweinhart, Barnes, & Weikart, 1993), and the *Better Beginnings, Better Futures* program in Canada (bbbf.queensu.ca), all of which have been examined with respect to their effects on children's school readiness. A number of longitudinal studies have examined the relationships between school readiness and later academic achievement (see Duncan et al., 2007, for a meta-analysis), with some of them including a wide range of additional variables (e.g., child behavior, parental support, family socioeconomic status) as control or predictor variables in their design (e.g., Alexander & Entwisle, 1988; Entwisle, Alexander, & Olson, 2007, for the US context, and the National Longitudinal Study of Children and Youth (NLSCY) for the Canadian context (Statistics Canada, 2007).

Similarly, school readiness as measured by the EDI has been studied within these theoretical approaches. For example, Lapointe, Ford, and Zumbo (2007) as well as Kershaw, Irwin, Trafford, and Hertzman (2005) used a sociological approach to examine the relationship between socioeconomic neighborhood characteristics and child development as assessed with the EDI; Pelletier and Corter (2005) conducted a study that related preschool intervention program efforts to EDI measures; and Lloyd and Hertzman

(2009), in a longitudinal study, linked children EDI scores to their later academic achievement. From a construct validity perspective, it is, of course, important, to clearly delineate to what degree these approaches are theoretically compatible, and maybe even complementary, or contradictory. As will be shown later, Bronfenbrenner's bioecological theory of human development provides an apt conceptual framework for integrating these different approaches.

Another approach to validity has been prominently advocated by Messick (1995). According to Messick, validity, in its broadest, most comprehensive sense, includes the notion that the interpretations, actions, and social consequences that are based on measurements of the construct—as reflected in test scores—are also valid. From this view, construct validity thus also depends on whether any consequences of the measurement are in line with the purpose of the test. Clearly, this approach to validity requires, in general, that a test's purpose, its underlying theoretical assumptions, its uses, and the social context's values are explicitly stated. For the EDI, this approach has not yet been systematically addressed. Therefore, the following discussion of the EDI's validity includes an explication of the EDI's underlying theoretical assumptions, as well as of the purpose for and context within which the EDI is (intended to be) used.

Previous writings about the EDI have addressed, respectively, (i) some theoretical aspects pertaining to the EDI (Janus & Offord, 2000, 2007), (ii) the purpose for and context within which the EDI may be used (Janus & Offord, 2000; Nosbush, 2006), and (iii) specific validity issues of the EDI (Guhn, Janus, & Hertzman, 2007). As the theory, purpose, practice, and—ultimately—validity of the measurement of school readiness are

all inherently connected, it is, however, necessary to formulate a framework that explicitly formulates the interconnections between these strands. Accordingly, the purpose here is to (i) explicate and develop underlying theoretical assumptions of the EDI, to (ii) expand the scope of the theoretical discussion by integrating theoretical considerations from the child development, sociology, neighborhood effects, and validity literatures, to (iii) explicitly link the theoretical foundations to the purpose and (assessment) practice of the EDI, and to (iv) illustrate how Bronfenbrenner's bioecological theory of human development can serve as a unifying conceptual framework to inform validity and validation issues with respect to the EDI. To this end, the forthcoming discussion is organized around the following six sections: (i) The theoretical construct of *school readiness*; (ii) School readiness: Links between purposes, theory, and assessment; (iii) Sociological views and neighborhood effects; (iv) The bioecological theory of human development; (v) The EDI from a Bronfenbrennerian perspective; (vi) Implications for validation research and practice.

The Theoretical Construct of *School Readiness*

The EDI has been “designed to provide communities with an informative, inexpensive and psychometrically sound tool to assess outcomes of early development as reflected in children's school readiness” (Janus & Offord, 2007; p. 1). The EDI assesses those ‘outcomes of early development as reflected in children's school readiness’ in five domains: (i) Physical health and well-being, (ii) Social competence, (iii) Emotional maturity, (iv) Language and cognitive development, and (v) Communication skills and general knowledge. In the words of the EDI authors (Janus & Offord, 2007), competence

in these five domains is considered essential for children to be “ready to benefit from educational activities offered in the school environment” (p. 4).

Given the centrality of the school readiness construct in the EDI, it is necessary to define this construct, not only according to the EDI’s definition, but also within the context of the school readiness research literature. The number and heterogeneity of current theories and practices regarding the construct(s) of *readiness*, *readiness to learn*, *school readiness*⁵, or *readiness to learn at school* mirror the historical and cultural diversity inherent to theories and practices pertaining to learning and schooling themselves. This diversity of theories pertaining to learning, schooling, and school readiness is not only indicative of the complexity of these issues, but it also hints at the societal relevance of these constructs. After all, a society’s collectively endorsed perceptions of school readiness define at what age formal schooling starts, can influence admission decisions for individual children, can affect educational practices at the pre-school and early school level, and can even shape the social structuring of family and community life. Not surprisingly, debates about school readiness have therefore been highly contentious, as opinions on education are based on a blend of people’s personal values, political ideologies, cultural norms, social habits, knowledge, research evidence, intuition, experience, economic constraints, and feasibility considerations. A comprehensive review of these debates is neither feasible, nor desirable in this context.

⁵ The author recognizes the multiple – and often conflicting – definitions of the term “(school) readiness”. In this paper, the term school readiness is used to refer to the teachers’ assessment of their Kindergarten class children’s developmental status in five different domains. For a current review of theoretical and empirical approaches to school readiness the reader is directed to Pianta, Cox, and Snow (2007). With regard to the EDI, Janus and Offord (2007) explicitly differentiate school readiness from readiness (to learn from birth), and consider school readiness to consist of a set of competences that will allow children to benefit from educational activities offered in the school environment (Janus & Offord, 2000). In this paper, the term *school readiness* is consistently used.

Instead, it is attempted here to delineate the field according to a number of long-standing debates and contested key questions that, in various combinations, are at the core of most school readiness theories:

- (i) What developmental domains or characteristics are facets of school readiness?
- (ii) To what extent is school readiness a result of biological maturation (*nature*), of experiences (*nurture*), or of a combination of both?
- (iii) To what extent is school readiness an individual (child) characteristic, a group (e.g., aggregated school or neighborhood) characteristic, or an (interdependent) combination of both?
- (iv) Can school readiness be defined in terms of objective, absolute criteria (i.e., observable, measurable norms; standardized references), or is school readiness a subjective and/or relative concept (i.e., person-, context-, and time-dependent)?
- (v) Is school readiness conceived of as the outcome of early development, or the predictor of later school achievement and other developmental outcomes, or primarily a key transition within a child's developmental trajectory from a life-span perspective? In other words, does school readiness have a retrospective or prospective focus, or a combination of both?

In the following, these questions are delineated by presenting a number of different positions and as yet unresolved issues. Then, critical issues regarding school readiness measurement and assessment are discussed, and Bronfenbrenner's bioecological theory of human development is illustrated. Based on those theoretical considerations, the paper

then attempts to position the EDI with regard to the five core questions.

With respect to the first question, about the developmental domains of school readiness, Table 2.1 aligns the results from a review of the school readiness literature (Doherty, 1997; cf. National Education Goals Panel, 1995) with results from a national parent and teacher survey (US) on school readiness (Lewit & Schuurmann Baker, 1995), as well as with Gardner’s *multiple intelligences*, which are defined as societally valued capacities for problem solving and creative action (1999), in order to illustrates how these respective examples map onto each other.

Table 2.1 Developmental Domains of School Readiness

Doherty (1997): School readiness domains ¹	Gardner (1999): Multiple intelligences	National Household Education Survey (1995): Teacher and Parent perceptions ⁵ of school readiness
Cognition and general knowledge	Logical-mathematic (Spatial)	Can count to 20 or more ‘P’ Able to use pencils or paint brushes ‘P’
Language use	Linguistic	Knows letters of the alphabet ‘P’
Social knowledge and competence ²	Interpersonal	Takes turns and shares ^{‘T’/‘P’} ; Can follow directions ^{‘T’} ; Is not disruptive of the class ^{‘T’} ; Is sensitive to other children’s feelings ^{‘T’} ; Communicates needs, wants, and thoughts verbally (in child’s primary language) ^{‘T’/‘P’}
Emotional health and a positive approach to new experiences ³	Intrapersonal	Sits still and pays attention ‘P’ Enthusiastic and curious in approaching new activities ^{‘T’/‘P’}
Physical well-being and motor development	Bodily-kinesthetic (Spatial, Musical) ⁴	Physically healthy, rested, and well-nourished ‘T’

¹ Research in this area has frequently endorsed the view that either the cognitive domain or the social-emotional domain is the primary domain of school readiness and, accordingly, a recent meta-analysis (La Paro & Pianta, 2000) solely distinguishes between these two broad domains.

^{2,3} The National Education Goals Panel’s (1995) categorization of school readiness domains is equivalent, except that it labeled these two domains as ‘Social and emotional development’ and ‘Approaches to learning (motivation; independence, etc.)’, respectively.

⁴ In later versions of the theory, Gardner included naturalist, spiritual, existential, and moral intelligences.

⁵ Characteristics of school readiness considered important by more than 50% of teachers (indicated by ‘T’) and parents (indicated by ‘P’) (from Lewit & Schuurmann Baker, 1995).

Questions (ii) and (iii)—about school readiness as a result of nature versus nurture, and school readiness as individual versus group characteristic—are theoretically related, and thus an overview of different approaches to them is jointly illustrated in Figure 2.2.

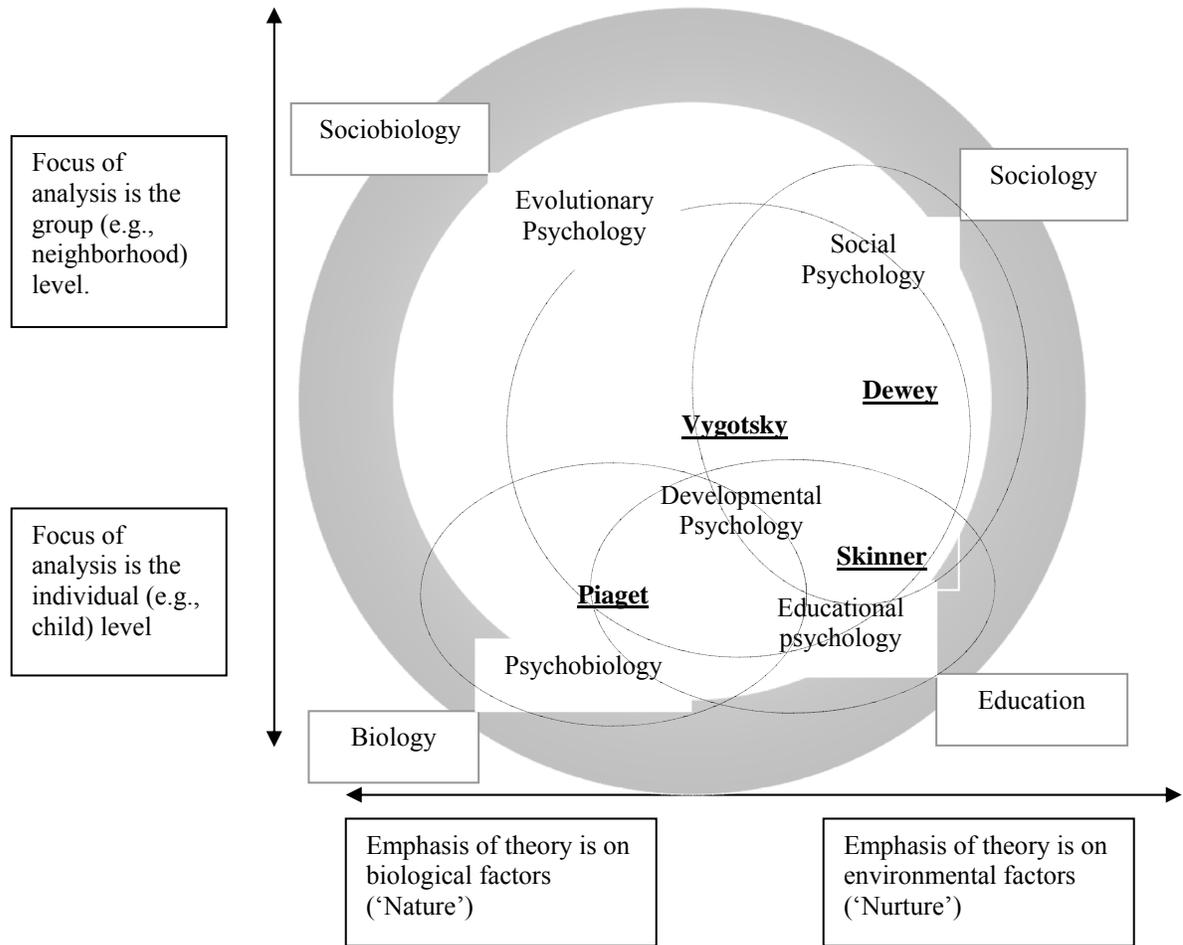


Figure 2.2 Disciplinary foci in human development research

The figure depicts how biology, education, sociology, and sociobiology, as well as the respectively associated sub-disciplines of psychology, have, traditionally, focused on either individual or group characteristics (vertical axis), and either biological or environmental factors (horizontal axis) with respect to examining human behavior and development. Likewise, theories of development (e.g., Piaget's *genetic epistemology*; Vygotsky's *social constructivism*) and learning (e.g., Skinner's *behaviorism*; Bandura's *social (cognitive) learning theory*), growing out of the different disciplinary traditions, have emphasized either individual or group processes, and either biological or environmental factors.

As a consequence, these disciplinary differences in emphasis can be found in the school readiness theories that are associated with these child development and learning theories. Maturationist views (e.g., Gesell) on school readiness emphasize, as the name implies, the importance of biological maturation—'the gift of time'—for school readiness. Empiricist views emphasize the need for children to learn specific skills before school entry. Constructivist views (e.g., Piaget) combine these approaches by emphasizing the notion that teaching practices and learning experiences need to be developmentally appropriate, in that they fall within the biologically delimited capacities of the child. Social constructivist views endorse a similar, interactive approach, though they emphasize the role of the social context; for example, Vygotsky's concept of the *zone of proximal development* (1936/1978) refers to the difference between a child's current developmental status and the potential that can be reached under guidance by other (e.g., teachers, peers). Yet other views emphasize the primacy of social relationships for children's development and learning (e.g., attachment theory; see

Watson & Ecken, 2003). Finally, some sociological theories of development and learning (e.g., Dewey's transactional view) emphasize the notion that development is not only affected by the social context, but can, in fact, not even be interpreted appropriately without an understanding of the social context. In line with this view, some authors conceive school readiness as a construct that characterizes entire ecologies (e.g., a school or a community), rather than solely individual children as school ready (e.g., Piotrkowski, 2004; Kagan & Rigby, 2003)⁶. As will be discussed later, Bronfenbrenner's bioecological theory of human development defines an intellectual space that allows one to view these different theories as mutually informative and complementary, rather than as competing alternatives.

The fourth question, about school readiness as defined by subjective, relative versus objective, absolute criteria, clearly addresses ontological and epistemological issues, and thus enters thorny terrain. For the purpose of this discussion, it suffices to point out that most conceptions of school readiness contain a blend of both the objective, absolute, and the subjective, relative approaches, with implications for the assessment of readiness. For example, certain cognitive competences, such as knowledge of basic numbers and vocabulary, and skills such as holding a crayon represent fairly objective, absolute criteria. Not surprisingly, such items feature prominently in widely used school readiness tests (e.g., *School Readiness Test*, Anderhalter & Perney, 2004), as they are relatively easily to obtain. On the other hand, certain social, emotional, or communicative competences, such as getting along with peers, being interested and motivated to learn,

⁶ Cf. Meisels (1999) for this categorization of school readiness theories.

and being able to communicate needs are inherently context-dependent (i.e., teacher- and peer-dependent), and thus subjective and relative by nature. This, in fact, has prompted some authors to argue against school readiness tests, and in favor of so-called authentic assessment. Authentic assessment refers to a practice that occurs over time, is embedded within the student-teacher relationship, and is solely to be performed if it evidently supports the learning and development of the child (see Graue, 1995).

From a theoretical perspective, this is of interest insofar as it affects to what extent and for which domains school readiness may be universal or may differ from context to context. From a measurement perspective, this is even more important, as it raises not only the question of how the subjective, relative aspects of school readiness can be assessed, but also of whether or how such assessments can be compared across context.

The fifth question asks whether school readiness is considered to be (i) a developmental outcome, and thus a retrospective concept, (ii) an informative snapshot of children's current developmental status, (iii) a predictor of further development, and thus a prospective concept, or (iv) primarily one of many developmental markers on a longitudinal, life-long trajectory. Examples for all of the above exist. A retrospective approach is typically taken in cases where interventions, preschool, or child care are evaluated in terms of their effectiveness to increase children's school readiness (e.g., Zigler & Valentine, 1979). The snapshot-in-time approach represents a common component of diagnosis that is intended to inform parents, teachers, counselors, and others about the current status of a child, with the intent to inform, respectively, their parenting, teaching, and counseling. In some cases, the snapshot approach, when used

inappropriately—for example, by solely administering poorly validated school readiness tests—has led to placement and retention practices considered inadequate (e.g., Ellwein, Walsh, Eads II, & Miller, 1991). A predictive approach has been a focus of psychological research examining the predictive validity of school readiness assessments with regard to, for example, later academic achievement (see meta-analyses by Duncan et al., 2007; La Paro & Pianta, 2000).

Finally, one prominent example of a longitudinal study that examines school readiness as part of broader developmental trajectories is the High/Scope Perry Preschool study (Schweinhart et al.; 1993), in which children were followed from the early years into adulthood. It should be noted that longitudinal studies can also occur at an ecological level, meaning that ecological units (e.g., neighborhoods, countries), rather than individuals are tracked. For example, the OECD Programme for International Student Assessment (PISA; www.pisa.oecd.org) assesses school achievement of consecutive cohorts of 15-year olds to allow for (national and international) comparisons over time. Of course, these two approaches can be combined, as is done by the NLSCY (Statistics Canada, 2007), which tracks the development of numerous consecutive cohorts of children from the early years through adulthood. This combined approach allows one to study whether diverse historical contexts are systematically related to differences in developmental outcomes between cohorts.

This approach lies at the core of the *life course paradigm* (Elder, 1994) or *life span* approach to studying human development (e.g., Baltes, Reese, & Pipsett, 1980). Bronfenbrenner integrated concepts pertaining to this approach into the bioecological theory, within his conceptualization of the *chronosystem*. As will be discussed later,

Bronfenbrenner's chronosystem paradigm allows one to aptly situate the EDI with respect to developmental time.

Now that a number of key perspectives on school readiness have been identified, it is important to discuss the relationship between school readiness theory, assessment, and the underlying educational goals and purposes of both. Following that discussion the paper will explore where the EDI fits with regard to theory, its approach to assessment, and its underlying purpose. Before addressing this question specifically, however, the following section discusses an inseparably connected issue, namely the relationship between school readiness theory, assessment, and the underlying educational goals and purposes of both, school readiness theory and assessment.

School Readiness: Links between Purposes, Theory, and Assessment

Education is an inherently practical as well as purposeful endeavor. Therefore, school readiness theory, educational practices (during the pre-school and school age period), and their purposes cannot be viewed independently. Likewise, assessment practices, whether they are formal or informal, idiosyncratic or universally alike, cannot be perceived separately from educational practices and purposes (cf. Schrag, Styfco, & Zigler, 2004). This argument implies that theoretical conceptions of school readiness should go hand in hand with clearly explicated educational purposes, and should provide a framework, principles, or guidelines that inform educational practices to accomplish defined purposes. If, for example, the goal is that all children be 'ready for school' at school entry (Governor General of Canada, 1997; cf. National Education Goals Panel, 1995, for the US context), an adequate, comprehensive school readiness theory should

not only define criteria of school readiness (cf. Table 2.1), but should also delineate the processes and practices that lead to and foster school readiness.

With respect to school readiness assessment practices, the field has seen, across time and across contexts, a wide range of assessment recommendations, from universally mandated or widely practiced school readiness screening to complete abandonment of formal school readiness tests (for examples, see Cuban, 1992; Shepard, 1997). Evidently, an evaluation of whether any type of school readiness assessment is valid (and successful) depends on the educational purpose the assessment is designed to accomplish or intended to promote.

It was stated earlier that one of the purposes of the EDI, through linking EDI data with other child development data (e.g., pre-school screenings; school achievement) and other social indicators (e.g., socioeconomic status; social capital) at the neighborhood level, is to provide communities with knowledge on children's development and their social and ecological context. The knowledge on children's development thus created, furthermore, is intended to facilitate the translation of research knowledge into action, as it can be used to inform families, practitioners, educators, researchers, and policy makers with respect to their practices and decisions in support of children's and families' well-being (Janus et al., 2007; Janus & Offord, 2000, 2007; Nosbush, 2006). By doing so, the EDI solely refers to a fairly general approach as far as educational practices are concerned. This issue will be revisited in the section on implications for validation research and practice, where an attempt is being made to expand the discussion on what theory-based recommendations for approaches to practice the EDI could provide.

Given the nature and scope of the EDI's purposes, an elaboration of the theoretical aspects of the EDI is necessary. Most importantly, this discussion needs to address issues deriving from the proposition that EDI scores are solely to be used at an aggregate level (e.g., the neighborhood level). Thereby, the EDI shifts the focus from the individual child to the community or population (Janus et al., 2007), so that "school readiness" may actually be perceived as a neighborhood or population characteristic. In other words, rather than asking, 'Is this child ready to learn at school?', the EDI addresses the questions, 'Are the children *in this neighborhood/population* ready to learn at school?', and, 'Does this neighborhood/population support its children in a way that they are ready to learn at school?'.

The EDI's use of a neighborhood or population perspective on school readiness (Janus & Offord, 2000; 2007) has been informed by writings that view child development from an ecological perspective (e.g., Love, Aber, & Brooks-Gunn, 1994). This ecological perspective on child development is based on theories that have their roots in social and developmental psychology and (urban) sociology⁷. In recent decades, the ecological perspective has led to the emergence of the so-called *neighborhood effects literature* (Jencks & Mayer, 1990; Sampson, Morenoff, & Gannon-Rowley, 2002), has increasingly permeated the field of population health (Heymann, Hertzman, Barer, & Evans, 2006), and is at the core of the evolving discipline of *community psychology* (Dalton, Elias, & Wandersman, 2007). In the following, the most relevant tenets of the underlying theories are reviewed.

⁷ With respect to Figure 2, this perspective would be placed in the upper right quadrant—(emphasizing group characteristics and the role of experiences on development)—whereas traditionally widely employed school readiness tests would have to be placed in the lower left quadrant—(emphasizing individual characteristics and the developmental primacy of genetics and maturation).

Sociological Views and Neighborhood Effects

It has been a longstanding debate to what extent individual and environmental factors shape human development. This nature-nurture debate has to some degree focused on the relative importance of these factors. However, according to Hebb, this focus is tantamount to asking, "which contributes more to the area of a field, its length or its width" (cited in Hebb, Lambert, & Tucker, 1971, p. 213). In recent decades, the debate has therefore shifted toward determining in what ways individual and environmental factors *interact* with respect to human development. Of course, theories of how these interactions shape development have existed for many decades before this notion has become more widely accepted. In developmental psychology, for example, Vygotsky's concept of the *zone of proximal development* (1936/1978) lays out how the social and cultural context of a developing child influences its cognitive development. In social psychology, Lewin's *field theory* (1951) delineates how individual experiences and situational characteristics interact with respect to human behavior. Dewey, in his interdisciplinary writings on philosophy, psychology, education, and sociology, describes the *transactional* approach, a method to understand human development and behavior by studying ecological systems in their entirety, rather than by studying them as a combination of individual, separate components (Bentley & Dewey, 1949/1991). Finally, many of the core tenets of Bronfenbrenner's ecological model on human development—which will be described in more detail below—were first formulated in the 1970s (1977; 1979). These tenets primarily emphasize the importance of exploring and understanding the *interactions* of those processes that occur within and between different levels of environmental systems, and that are most relevant for human development.

In the neighborhood effects literature, the focus is, as the label suggests, on understanding the environmental effects on development at the neighborhood system level. This focus is based on the assumption that the neighborhood is a meaningful ecological entity, and that the social processes and mechanisms at this level can significantly influence human development and behavior, in addition to the processes and mechanisms that occur at other ecological levels. This notion is prominent in theories that are associated with the so-called Chicago School (or Ecological School) of Sociology⁸ (Sampson et al., 2002), which is philosophically rooted in *Pragmatism* (with John Dewey as one of the main proponents), and from which *Symbolic Interactionism* emerged (Blumer, 1969; Mead, 1934). Due to these philosophical roots, this theoretical branch of the neighborhood effects literature contains explicated epistemologies (i.e., theories of knowing/knowledge) and ontologies (i.e., theories of being/reality) that suggest certain research methodologies for the study of developmental phenomena of interest.

In terms of “*how* neighborhoods bring about change in a given phenomenon of interest” (Sampson et al., 2002, p. 447), a variety of assumptions and hypotheses exist. Sampson et al. (2002) have categorized the social processes and mechanisms that are assumed to underlie neighborhood effects on development into four classes. The first class, *social ties and interactions*, refers to the strength, frequency, and density of social relationships within neighborhoods. The second, *norms and collective efficacy*, pertains to the degree to which neighbors share expectations and trust, and act based on them for a common, public good. The third, *institutional resources*, captures the quality and quantity

⁸ The (University of) Chicago School is typically associated with ‘sociology’, but it needs to be noted that several of its influential scholars, like Dewey, Mead, and Blumer, came from backgrounds of philosophy and education, (social) psychology, and sociology, respectively.

of institutionally provided services that address community needs. The fourth, *routine activities*, encompasses (social) activities that are patterned according to land use, such as transportation, shopping, and recreation. Empirical evidence supporting this classification exists (e.g., Sampson, Raudenbush, & Earls, 1997), but is sparse, mainly due to methodological challenges and infeasibility of data collection. Thus, major challenges in understanding *how* the theoretically proposed social processes and mechanism at the neighborhood level affect development remain (Sampson et al., 2002; Jencks & Mayer, 1990).

Another branch of the neighborhood effects literature that has, in contrast, produced a vast amount of empirical data evolved from the area of *social capital* research. Social capital is a concept that has most prominently been discussed in the writings of Bourdieu (1985), Coleman (1988), and Putnam (2000); and it must be noted that their respective views of the concept are not equivalent, but variations on a theme. In the present context, it suffices to say that social capital generally refers to the value of social networks, in the sense that such social networks can produce tangible (e.g., material resources) and intangible (e.g., social support) benefits for the members of the networks.

Empirically, social capital, in its numerous facets, is related to a number of other constructs (e.g., school outcomes, socioeconomic status, and health). In contrast to the situation regarding the theories on social mechanisms underlying neighborhood effects mentioned above, however, theoretical rigor has not kept pace with the wide-spread use of ‘social capital’ concepts in empirical research, where the operationalization of the term has been inconsistent (for a review, see Sampson, Morenoff, & Earls, 1999). As a result, questions as to how the processes and mechanisms related to social capital actually work

also remain a matter of debate. In fact, a theoretically problematic phenomenon is that ‘social capital’ in some contexts is correlated to arguably negative developmental outcomes (e.g., Caughy, O’Campo, & Muntaner, 2003), whereas it is typically perceived as a generally beneficial resource, at least for those inside the social network (Putnam, 2000).

A concept that has received even more attention in the neighborhood effects literature is that of *socioeconomic status* (typically defined as a composite of education, income, and employment status) or *poverty*. Here, the situation is less ambiguous empirically: Low socioeconomic status or poverty have, certainly in their extreme forms, almost universally detrimental effects on children’s development (Berliner, 2005; Brooks-Gunn, Duncan, Klebanov, & Sealand, 1993; Jencks & Mayer, 1990; Wilson, 1987). However, it remains unclear to what extent and via which processes poverty at the neighborhood level—after controlling for poverty at the individual level—exerts its influence on child development⁹. In fact, in 1990, Jencks and Mayer had concluded that most neighborhood effect studies have employed a *black box model* as far as *processes* at the neighborhood level are concerned.

The situation has improved insofar as it could be systematically shown that constructs such as “informal social control, trust, institutional resources and routines, peer-group delinquency, and perceived disorder” are related to socioeconomic status and to developmental outcomes in “theoretically meaningful ways” (Sampson et al., 2002, p. 473). Nonetheless, the challenge regarding the question of what processes and causal

⁹ It is well-understood, however, how poverty at the individual and family level goes hand in hand with developmentally detrimental processes, such as poor nutrition, enduring stress, exposure to toxic environments, and lack of developmentally positive experiences and opportunities (Berliner, 2005).

relationships underlie these constructs remains.

This situation arises, apart from the conceptual intricacies and ambiguities involved in studying neighborhood effects of social capital or poverty on child development, from the fact that neighborhood studies are, due to their inherent design, confronted with complex methodological challenges. Two of these challenges are addressed here now, and they both have to do with the fact that neighborhood-level characteristics, such as social capital or neighborhood poverty, are conceptualized, by definition, as characteristics of a group.

The first challenge is one of measurement, and one that affects social indicators in general. Measurement of social indicators can be done in numerous ways. One can measure a characteristic at an individual level, and then aggregate. Depending on the type of aggregation, this can lead to quite different results (e.g., mean income versus Gini coefficient¹⁰ versus poverty percentage rate), with important implications for the interpretation of the aggregate measures. For other types of social indicators, the measurement itself occurs at a group level. Social capital, for example, is not the sum or average of individuals' social capitals, but the quantity (and quality) of connections between people within the group. The problem thus becomes how one should devise the groups. This entails conceptual and interpretational challenges that then lead to the second major challenge.

Apart from the theoretical considerations for choosing specific groupings, and the measurement problems related to it (e.g., exclusive versus overlapping groups; socially or

¹⁰ The Gini coefficient is a statistical index used to describe the distribution of income in a population, where 0 means absolute equality of income (i.e., everyone has the same) and 1 means absolute inequality (i.e., one person has everything). A definition and world map showing each country's Gini coefficient can be found at en.wikipedia.org/wiki/Gini_coefficient.

geographically defined groups; sampling issues)¹¹, there are challenges in soundly analyzing and interpreting group data. These challenges can occur whether group-level data (e.g., neighborhood poverty) are analyzed by itself or in combination with individual-level data (e.g., family poverty). At the group level, two common sources of error have been referred to as the *modifiable areal unit problem* and the *ecological fallacy*. The modifiable areal unit problem designates the phenomenon that the way in which data are aggregated into groups (units) can lead to drastically different statistical outcomes at the group (unit) level (Taylor, Gorard, & Fitz, 2003). The ecological fallacy refers to an interpretational error, namely that of incorrectly inferring relationships at an individual level from ecological (i.e., group-level) correlations (Piantidosi, Byar, & Green, 1988).

When group-level data are analyzed together with individual-level data, multilevel statistical procedures are employed to account for the *nested* structure of the data (Raudenbush & Bryk, 2002). These methods allow one to simultaneously account for variance at the individual level (within group) and the group level (between group). Even though these methods allow one to avoid the ecological fallacy (or, its reverse counterpart, the *atomistic fallacy*), multilevel models do not solve the modifiable areal unit problem. In other words, the decision of how to devise groups (e.g., geographically defined neighborhoods versus socially defined neighborhoods; small versus large groups) and their meaningfulness and interpretability remain conceptual and theoretical challenges. With regard to examining school readiness in the neighborhood context, it is thus necessary to define neighborhoods in terms of the neighborhoods' hypothesized

¹¹ See Raudenbush and Sampson (1999) for a discussion of methodological and conceptual issues pertaining to the assessment of ecological settings.

effect on child development, and to address potential effects of the statistical and conceptual challenges when interpreting school readiness data in a multilevel (i.e., neighborhood) context¹².

At this point in the discussion, the question arises as to how concepts relevant to an ecological study of school readiness, from the diverse strands of child development, education, psychology, and sociology and the neighborhood effects literature, can be integrated into a coherent framework. This integration aims to illustrate how such a framework might substantially contribute to a better understanding of the processes that are causally related to neighborhood effects. Previous work, faced with the conceptual challenges described here, has repeatedly referred to Bronfenbrenner's bioecological theory of development (e.g., Brooks-Gunn et al., 1993; Rimm-Kaufman & Pianta, 2000). Nonetheless, a number of Bronfenbrenner's key theoretical propositions that have the potential to critically advance a better understanding of the field (still) seem to be underused (cf. Bronfenbrenner & Morris, 2006). Therefore, the following section elaborates on Bronfenbrenner's theory, including its implications for the EDI and school readiness theory, in some detail.

The Bioecological Theory of Human Development

Urie Bronfenbrenner is recognized as one of the intellectual fathers of the US's federal compensatory preschool program *Head Start* (NHSA, 2005), and considered one of the most influential thinkers on human development, most notably for his

¹² Zumbo and Forer (in press) elaborate theoretically on school readiness as a multilevel construct, and refer to the EDI project for illustrative examples.

bioecological theory of human development (or *bioecological model*). The theory, for the inspiration and foundation of which Bronfenbrenner credits his mentor Kurt Lewin (see above), was first formulated in the late 1970s (Bronfenbrenner, 1977; 1979), and has since substantially been revised and elaborated (e.g., Bronfenbrenner, 1992; 2001; Bronfenbrenner & Morris, 2006).

Central to the bioecological model are epistemological propositions, which Bronfenbrenner defines within the concept of *science in the discovery mode*. These propositions, in essence, delineate how a well-constructed combination of theory and corresponding research designs “stand a chance of yielding new, more differentiated, more precise”, and “more valid scientific knowledge” on human development (Bronfenbrenner, 2001, p. 4). Equally central to the bioecological model is its objective to produce knowledge that is applicable for the benefit of children, for example, by informing social policies and programs. Due to the complexity of human development and of translating knowledge into practice, these objectives are, as Bronfenbrenner notes, not easily accomplished, and may sometimes not even be feasible; yet, some routes toward these objectives are deemed more promising than others.

First of all, Bronfenbrenner proposes to focus on the examination of *proximal processes*, which are considered the “primary engines of development” (Bronfenbrenner & Morris, 2006, p. 798), and are defined as “processes of progressively more complex reciprocal interaction between an active, evolving bio-psychological organism and the persons, objects, and symbols in its immediate external environment” (p. 797).

Second, Bronfenbrenner stresses the notion that “in human development, the main effects are likely to be found in the interactions” and that the “interactions to be examined

be theory based” (p. 802). It is noteworthy that the term interaction is used in different meaning in these two propositions. In the first, it is used to refer to processes occurring between the developing person and its environment. In the second, it refers to the concept of interactions as used in the field of statistics, where interactions refer to situations in which two (or more) variables have multiplicative effects—in addition to additive effects¹³—on an outcome of interest. So, in regard to human development, such interactions are present when the effects of developmentally relevant factors are moderated (i.e., amplified, attenuated, or reversed) by other developmentally relevant factors with regard to a developmental outcome.

Third, Bronfenbrenner recommends a comprehensive research model to accomplish this, namely the process-person-context-time model. This model conceptualizes how, ideally, a research study should simultaneously examine those proximal *processes*, *person* characteristics, *context* characteristics, and characteristics pertaining to *time* (e.g., historical time; timing in life) that are considered relevant for a specific developmental outcome, and this model also calls for a research design and statistical analyses that allow one to detect statistical interactions.

Bronfenbrenner has repeatedly used one study as an example to illustrate this model (Bronfenbrenner & Morris, 2006, p. 799ff.). In the study, it is examined how the *process* of responsive versus unresponsive mothering affects children’s behavior problems (i.e., the developmental outcome variable) at age 2 and age 4 (i.e., the *time*/chronological age variable), and this is done with respect to the children’s birth weight (normal, low, and very low; i.e., the *person* characteristic) and the mother’s socioeconomic status (high,

¹³ In the statistical jargon of the social sciences, additive effects are typically referred to as *main effects*.

middle, and low; i.e., the *context* characteristic). The results show a clear main effect, namely that responsive mothering consistently leads to less behavior problems for both age groups. Interactions were also found, in that the positive effect of responsive mothering is more pronounced for 4 year olds than 2 year olds, and among the 4 year olds, it is particularly pronounced in the low SES context. Another (statistical) interaction occurred between the birth weight and SES variables, in that—in the low SES group—the normal birth weight children benefited more than low birth weight children, and those more than very low birth weight children, whereas all of the children (i.e., all three birth weight groups) in the high and middle SES groups benefited equally from responsive mothering.

These interactions are interesting findings in themselves. More important in this context, however, is the following: The data for this study had previously been published (by Drillien, cited in Bronfenbrenner & Morris, 2006), but only main effects of responsive mothering were reported. Bronfenbrenner, however, had expected, based on his theoretical proposition, to also find specific interactions; which were only found in his re-analysis that statistically examined the hypothesized interactions.

Even though Bronfenbrenner makes a conceptual distinction between developmental outcomes, person characteristics, context characteristics, process variables, and time variables, these components need to be, in terms of human development, understood to be functioning within an interactional, holistic ecological system. This notion of interdependence is reflected in Bronfenbrenner's concept of *nested systems*. Since Bronfenbrenner's own definitions have evolved over time, the most recent definitions are provided here. Bronfenbrenner defines four nested systems: First, the *microsystem*, which

refers to all processes, or “activities, roles, and interpersonal relations experienced by the developing person” (p. 148) that typically take place within a person’s immediate environment (e.g., in the home, at school). The *mesosystem* refers to the “linkages and processes taking place between two or more settings containing the developing person” (p. 148; e.g., the processes occurring between home and school). The *exosystem* “encompasses the linkages and processes taking place between two or more settings, at least one of which does *not* ordinarily contain the developing person, but in which events occur that influence processes within the immediate setting that does contain that person” (p.148; e.g., processes taking place between a child’s home and a parent’s workplace). The *macrosystem* “consists of the overarching pattern of micro-, meso-, and exosystems characteristic of a given culture, subculture, or other broader social context [...]” (p.149).

In addition to these ecological systems, Bronfenbrenner has defined the so-called *chronosystem*. The chronosystem refers to a life-span perspective on development, stating that developmental effects of proximal processes may critically depend on *when* and *in which order* they happen in a person’s life, as well as on when they happen within the historical context.

What is critical about Bronfenbrenner’s nested systems is that each system refers to processes, all contain persons with their own distinct person characteristics, all have their own context characteristics, and all occur within their own time characteristics. Accordingly, developmental outcomes are not to be understood in terms of a process-person-context-time model that exclusively operates within one system. Rather, developmental outcomes result from the interactions of process-person-context-time variables within each of the systems, as well as from the interactions between these

systems.

The complexity of conceptualizing, let alone methodologically realizing, an examination of the interactions within and between systems grows, of course, exponentially with each additional layer. In acknowledgement of the fact that this complexity could easily become overwhelming in the search for new knowledge on human development, Bronfenbrenner thus strongly recommends that developmental science operating in the discovery mode is explicitly theory-based (Bronfenbrenner & Morris, 2006). After having reviewed some of the core tenets of Bronfenbrenner's theory, it is now time to discuss how these apply to the EDI and the questions raised throughout the paper.

The EDI from a Bronfenbrennerian Perspective¹⁴

In the section 'the theoretical construct of school readiness theories', five key questions about the school readiness construct were discussed. These questions are now revisited from a Bronfenbrennerian perspective.

(i) *What developmental domains are facets of school readiness?* The school readiness-related literature provides, as sketched out above, a number of theories proposing a set of developmental domains (Doherty, 1997; Love et al., 1994) or skills (Lewit & Schuurmann Baker, 1995), which are considered to be critical for doing well in school, an approach that is also taken by the EDI. Gardner's theory of multiple

¹⁴ Rimm-Kaufman and Pianta (2000) have presented an 'ecological perspective on the transition to Kindergarten', in which they use Bronfenbrenner's theory as a framework for guiding empirical research. Their approach primarily focuses on other aspects than the approach presented here, but there are several interesting parallels as well as complementary ideas between their paper and this chapter.

intelligences that is juxtaposed with the school readiness theories proposes a broader spectrum of capacities, but rather than claiming that those are essential for doing well in school, the intelligences are defined as capacities needed to do well in societally valued endeavors, such as becoming a teacher, athlete, scientist, writer, craftsperson, or musician.

The phrase ‘societally valued’ is of particular relevance for the discussion. As Bronfenbrenner (1992) notes, conceptions of (cognitive) capacities have frequently been assumed to be “invariant across place and time”; an assumption that, in fact, characterizes a “wide range of measures” and assessments (p. 121). According to a number of cross-cultural studies, this assumption does not, however, necessarily hold. Rather, it must be acknowledged that certain capacities (e.g., math skills) are highly context-dependent, leading Bronfenbrenner to conclude that “the context, in which [a capacity] takes place is not simply an adjunct to the cognition, but a constituent of it” (p. 127).

A further point illustrates why this distinction is important in this context. In one place, Bronfenbrenner (1988) attempts a definition of a developmental outcome, and differentiates between development (as a process), developmental outcomes (patterns of mental organization resulting from development), and behaviors (as indications of developmental outcomes). Referring back to the previously discussed process-person-context-time model, it might help to think of human development as a construct that includes both sides of an equation, in which a given developmental outcome is equal to the sum of the developmentally relevant process, person, context, and time characteristics, plus the developmentally relevant (first and higher order statistical) interactions among those processes.

As mentioned before, the role of processes in this equation is essential. Without understanding the processes that lead to person- and context-specific developmental outcomes, Bronfenbrenner argues, it might be possible to predict developmental outcomes, but not to explain what brings them about.

What does this imply with respect to the EDI? First, the developmental domains specified by the EDI represent developmental outcomes, not development itself. So, in order for the EDI to accomplish its purpose to provide communities with applicable knowledge on human development, it can only do so in combination with information on developmentally relevant context and process characteristics. Second, the developmental domains reflect, at least to some degree, societal and cultural values and expectations—after all, competences in the five developmental domains of the EDI are considered necessary in order to benefit from the experiences offered in school, which is an institution rooted in social and cultural norms. These two points have been explicitly acknowledged in previous writings on the EDI (e.g., Janus et al., 2007; Janus & Offord, 2000; Nosbush, 2006). Accordingly, the five domains of the EDI need to be viewed as an attempt to capture those domains that are—according to the current school readiness literature—fairly representative of current social and cultural norms (cf., Doherty, 1997; Duncan et al., 2007; La Paro & Pianta, 2000; Meisels, 1999; National Education Goals Panel, 1995; Rimm-Kaufman & Pianta, 2000).

(ii) Is school readiness a child characteristic, an environmental characteristic, or a combination of both? Is school readiness a result of biological maturation, of experiences, or of a combination of both? These two questions were addressed jointly above, primarily by means of an illustration (Figure 2.2). In that illustration, it was proposed to what degree certain disciplines (e.g., psychology, biology, sociology, education) emphasized personal versus contextual, individual versus group, and biological versus environmental factors with respect to studying human development and behavior. Historically, the different disciplinary approaches have, at times, been pitted against each other, to the degree that different schools of thought are deemed incompatible with each other (as, for example, in the nature-nurture debate). Regardless of whether such juxtapositions of disciplinary approaches were adequate or not, Bronfenbrenner's theory makes a strong case for integrating different disciplinary approaches for more fully understanding human development. This notion is illustrated in Figure 2.3, a revised version of Figure 2.2.

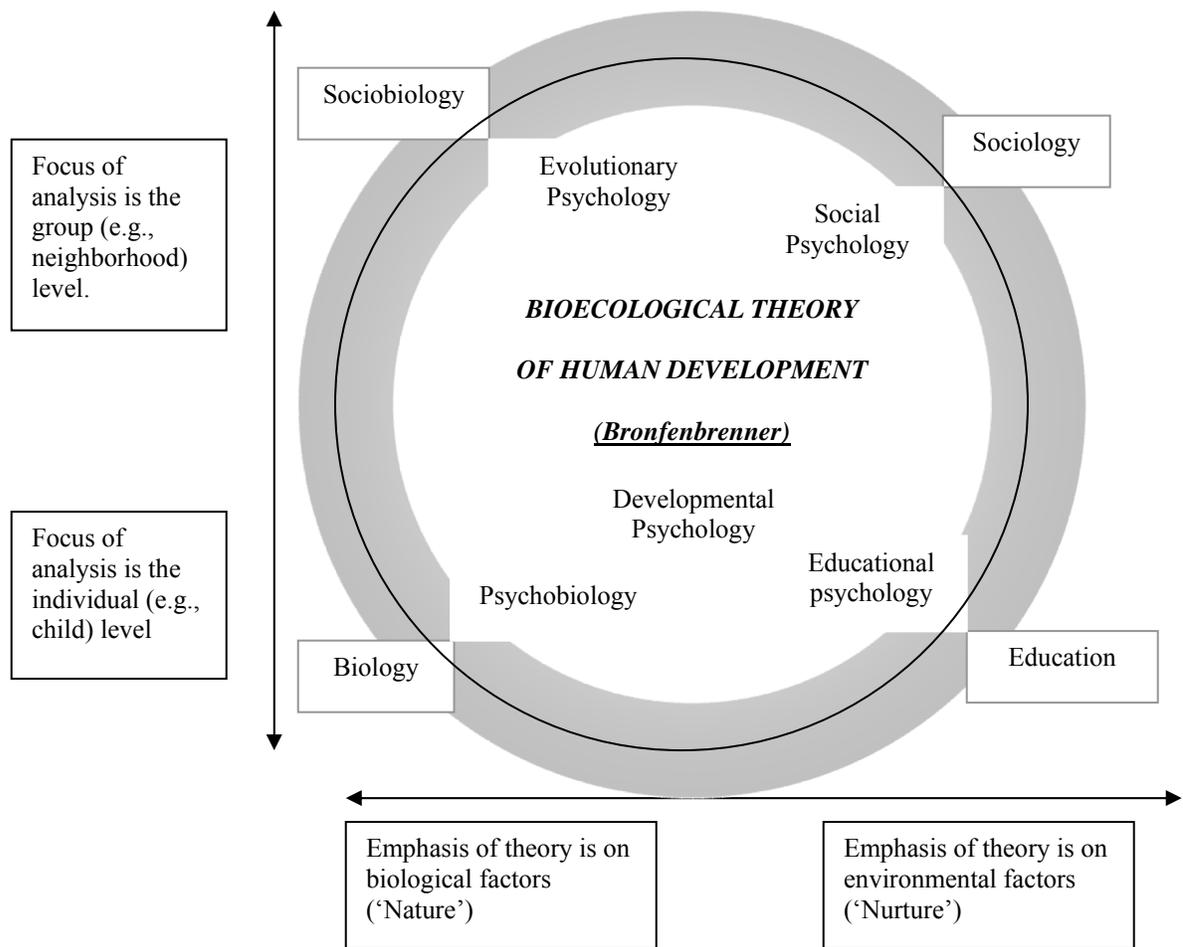


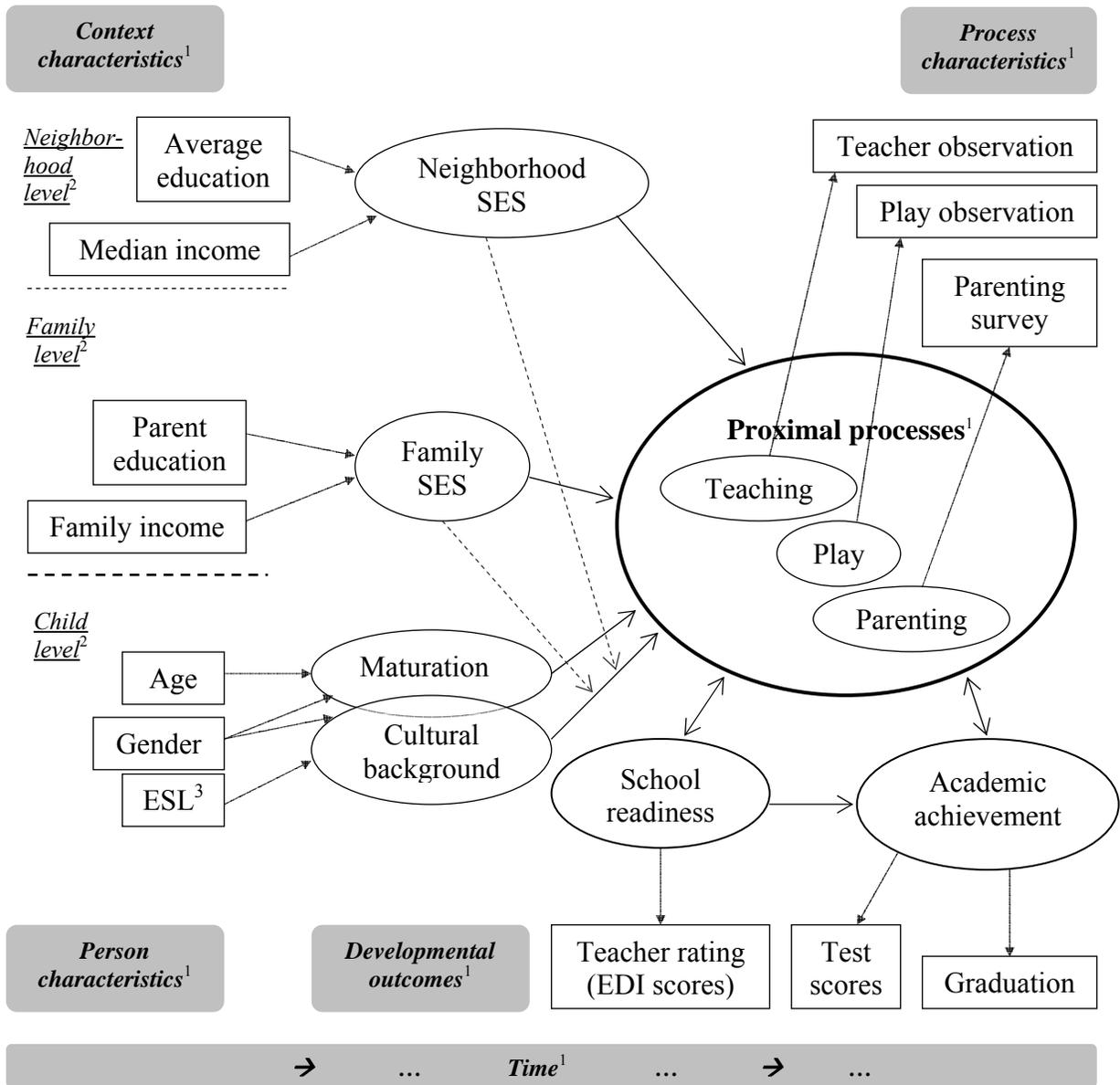
Figure 2.3 Illustration of the bioecological theory of human development as integrative framework.

Bronfenbrenner's integrative framework allows one to view the different disciplinary lenses that have traditionally been used to understand human development as complementary lenses, all of which are necessary jointly for one to see a more coherent, more holistic picture that is made up of different disciplines.

(iii) Can school readiness be defined in terms of objective, absolute criteria, or is school readiness a subjective and relative concept? Should school readiness be perceived as the outcome of early development, or the starting point (i.e., predictor) of later school achievement and other developmental outcomes, or primarily a key transition within a child's developmental trajectory from a life-span perspective?

Bronfenbrenner argues that (i) human development can only be understood with respect to the interactions of process, person, context, and time characteristics, and also that (ii) human development researchers should attempt to identify principled patterns, which can be objectively described and embedded within a cogent theoretical framework. The critical point is that the latter needs to be done within a (developmentally relevant) ecology in its entirety. In other words, the relationship between a certain developmental outcome and a certain person characteristic is, by definition, relative with regard to developmentally relevant process, context, and time characteristics, and their interactions. However, once the developmentally relevant characteristics of the process-person-context-time model are taken into account, such a system can be described in theoretically predictable, patterned terms.

With regard to the EDI, the implications are illustrated in Figure 2.4, which represents an integration of the nomological networks presented in Figure 2.1. (Throughout the paper, the importance of examining interactions has been emphasized. Therefore, in the integrated network presented in Figure 2.4, two dashed arrows were inserted to graphically convey an example of this notion. It must be noted that there are different ways of graphically representing interactions.)



¹For definitions, please refer to the description of the *bioecological theory of human development* and the *process-person-context-time model*.

²Level of measurement at which the data are obtained

³ESL: English-as-a-Second-Language

Figure 2.4 A Bronfenbrennerian nomological network for human development¹⁵

¹⁵ It must be noted that the nomological network in its entirety is primarily of theoretical and conceptual relevance to the EDI project, rather than actual data analyses, as data on children's family-level SES or on proximal processes are not available at a population level.

As has hopefully become apparent, these integrative figures are a reiteration of Bronfenbrenner's principles and themes, applied to the EDI context. They emphasize the interactive, ecological nature of human development, the integration of pertaining empirical and theoretical strands, and highlight the essential importance of the interactions between the numerous factors that jointly shape human development over time.

Implications for Validation Research and Practice

It has been stated repeatedly throughout this paper that the so-called proximal processes are key factors in shaping human development. This notion is not only core to Bronfenbrenner's theory, as well as to the sociological theories on neighborhood processes mentioned above (Sampson et al., 2002), but also a key tenet in what is known as *developmentally appropriate practice* (Bredekamp, 1986; NAEYC, 1996). Developmentally appropriate practice (DAP) refers to a set of principles that aim to guide practices in early childhood education, parenting, and child care settings, with the underlying purpose to foster children's development. In this context, the concept of DAP is of particular interest, (i) because its theoretical foundations show considerable overlap with Bronfenbrenner's theory, (ii) because it is a framework widely (though not unanimously) endorsed by child development researchers, practitioners, and administrative organizations in Canada (e.g., Cohen, Kiefer, & Pape, 2005) and the US (e.g., NAEYC, 1996), and (iii) because DAP specifically addresses the part of the human development equation that is the most difficult to capture, namely the developmentally relevant proximal processes.

With regard to conducting validation research on the EDI, complementing Bronfenbrenner's theory with the principles delineated in DAP seems very promising, as the DAP principles describe in much more detail than Bronfenbrenner's theory the nature of the proximal processes that are considered developmentally beneficial. Furthermore, the principles outlined by DAP are rooted in prominent theories of the child development literature (e.g., Piaget, Vygotsky, Bronfenbrenner), and their validity has been corroborated by empirical evidence¹⁶. For example, the recommendations regarding the developmental importance of child-initiated as well as adult-guided play have been found to be related to numerous desirable outcomes in children, such as increased conceptual understanding and social competences (NAEYC, 1996).

The implications for future validation research are obvious, at least in principle: Ideally, one could simultaneously collect process, person, context, and time characteristic data for groups or entire populations of children, in order to see whether certain combinations of those multiple factors lead to theoretically predicted and systematic patterns of developmental outcomes. Logistically, the implementation of such research at a large scale is, of course, enormous, and confronted with major measurement challenges, as proximal processes are notoriously complicated and time-consuming to reliably and validly assess. Nevertheless, such research design is, according to Bronfenbrenner, the most powerful with respect to "yielding new, more differentiated, more precise", and "more valid scientific knowledge" on human development (p. 4, Bronfenbrenner, 2001).

¹⁶ While DAP has gained much currency in both the theory and practice of early childhood programs, it has not been universally embraced. Some scholars have critiqued DAP as relying too much on universal notions of child development and the imposition of western ideas of child development on minority groups and on non-western societies. For a discussion, readers are referred to Kessler and Swadener (1992).

In light of the situation that the EDI project (BC) is already of very large scale in its current form—as it is repeatedly collecting and linking EDI, school achievement, and demographic and socioeconomic neighborhood-level census data for entire cohorts in British Columbia—it can be argued that adding a well-designed research component pertaining to proximal processes to the current design could possibly result in an improved cost-benefit ratio, because it might exponentially expand the opportunities to address research questions on child development¹⁷. In fact, a major goal of this theoretical integration has been to corroborate such a recommendation. To this end, it is illustrated how a Bronfenbrennerian conceptualization of the school readiness construct and a Bronfenbrennerian approach toward validation issues pertaining to the EDI and the associated research projects might lead to specific research policy recommendations. Among these recommendations, the one that is particularly emphasized is to increase the focus on integrating proximal processes into the EDI research, as it remains the one major missing piece in the process-person-context-time research model, which Bronfenbrenner considers to be the ideal design for human development research.

A second, related recommendation emphasized throughout the paper is to design EDI research in a way that it specifically targets the examination of a priori theoretically hypothesized interactions between the process-, person-, context-, and time factors. Thus, examining interactions is clearly not just an analytical issue—that is, an issue of statistically modeling interaction terms—but most and foremost a design issue: one has to include those variables that are theorized to be interacting with respect to the developmental outcome under question.

¹⁷ This recommendation applies to EDI research projects in general.

A third recommendation refers to an issue foreshadowed in previous sections, namely the question of whether the practical, social consequences of the EDI measurements are in line with the purposes of the EDI. As mentioned, this issue is essential for a test's validity, as defined by Messick (1995). A pragmatic approach to this issue is to simply monitor whether children continue to do well or better throughout the implementation of the EDI. Obviously, such an approach by itself does not, however, allow one to make any causal inferences with respect to the relationship between children's development and the EDI project. A more comprehensive approach is to integrate the processes that belong to the EDI (e.g., dissemination of EDI research findings; discussion of EDI findings among community groups) into the overall research design, in order to examine whether the EDI does affect decisions and practices that, in turn, affect children's development—an aspect of EDI-related research that is still in its beginning.

Conclusion

Three main recommendations have been made in this paper: First, the EDI project (BC), which routinely collects early child development data on entire cohorts of Kindergarten children in British Columbia, and which links these data to school achievement data and demographic data at the individual level, as well as to socioeconomic, cultural, and demographic data at the neighborhood level, should, ideally, collect (developmentally relevant) proximal process data. That way, the one research component that is missing as far as the process-person-context-time model is concerned—the model that Bronfenbrenner recommends for conducting comprehensive

human development research—would be added to the research design. In that regard, the theoretically rich literatures on developmentally appropriate practice (NAEYC, 1996) and on social neighborhood processes (see Sampson et al., 2002) provide invaluable insights on what processes to examine in such research program.

Second, Bronfenbrenner suggests that in human development, the “main effects are likely to be in the interactions” (Bronfenbrenner & Morris, 2006, p. 802). Accordingly, the suggested research program should, ideally, include those process-, person-, context-, and time variables that are theoretically predicted to interact with regard to the developmental outcome under question, both in the design and the analytical methods of each study.

Third, in order to fully validate the EDI, the accompanying validation research must address to what degree the practical and social consequences of the EDI (e.g., interpretation of findings; decisions based on EDI results) are in line with the purpose of the EDI—addressing an aspect of validity that is proposed by Messick (1995). This would require a research process that is ongoing, and community-based or community-oriented.

These three recommendations have been presented as the result of an attempt to integrate the school readiness, neighborhood effects, measurement, and validity literature, using Bronfenbrenner’s bioecological theory of human development as a conceptual framework. Bronfenbrenner’s theory is widely cited in the child development literature, and also in the neighborhood effects literature. However, Bronfenbrenner himself noted that his own theory—despite its recognition—had not been resulting in a noticeable

increase of research employing his design recommendations for studying human development. In fact, Bronfenbrenner once wrote that whereas there had been “too much research on development ‘out of context’, [there now is] ... a surfeit of studies on “context without development”” (Bronfenbrenner, 1986, cited in Bronfenbrenner & Morris, 2006, p. 794).

Now, about twenty years later, one can find numerous research studies in the neighborhood effects literature that studies child development in context. What remains missing, to date, is an increased focus on those re-occurring, proximal processes that are considered to be most relevant developmentally, and an increased emphasis on study designs and analytical methods that can examine the interactions between those proximal processes and relevant person-, context-, and time variables. As soon as developmental research accomplishes a much deeper understanding of these proximal processes, the final major challenge will, most likely, be to develop an understanding of how to implement processes that are beneficial in one context into another context in a way that the transfer acknowledges the ecological differences between the two settings (cf. Elias, Zins, Graczyk, & Weissberg, 2003). Creating such knowledge that can then be translated into practice is, as stated, the ultimate purpose of developmental research as envisioned by Bronfenbrenner, and also the ultimate purpose of the EDI. The EDI research project seems well positioned to make substantial contributions in this direction, and it is hoped that the Bronfenbrennerian approach to the issue as presented here provides a useful conceptual framework for designing future validation research towards this end.

References

- Alexander, K. L., & Entwisle, D. R. (1988). Achievement in the first two years of school: Patterns and processes. *Monographs of the Society for Research in Child Development, 53*, 1-157.
- Anderhalter, O. F. & Perney, J. (Eds.) (2004). *School Readiness Test (3rd Edition)*. Scholastic Testing Service: Bensenville, IL.
- Baltes, P. B., Reese, H. W., & Lipsitt, L. P. (1980). Life-span developmental psychology. *Annual Review of Psychology, 31*, 65-110.
- Bentley, A. F. & Dewey, J. (1949/1991). Knowing and the known. In J. A. Boydston (Ed.), *John Dewey: The later works, Vol. 16*. Carbondale, IL: University of Southern Illinois Press.
- Berliner, D. C. (2005). Our impoverished view of educational reform. *Teachers College Record*. Retrieved on November 20, 2006, from <http://www.tcrecord.org/content.asp?contentid=12106>
- Blumer, H. (1969). *Symbolic Interactionism: Perspective and method*. Berkeley, CA: University of California Press.
- Lennox, R., & Bollen, K. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*, 305-314.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review, 111*, 1061-1071.
- Bourdieu, P. (1985). The forms of capital. In J. G. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education*, pp. 241-258. New York, Y: Greenwood.

- Bredekamp, S. (1986). *Developmentally Appropriate Practice*. Washington, DC: National Association for the Education of Young Children.
- Breton, J. J., Bergeron, L., Valla, J. P., Berthiaume, C., Gaudet, N., Lambert, J., et al. (1999). Quebec child mental health survey: prevalence of DSM-III-R mental health disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 40, 375-384.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32, 513-531.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U. (1988). Interacting systems in human development. Research paradigms: Present and future. In N. Bolger, A. Caspi, G. Downey, & M. Moorehouse (Eds.), *Persons in context: Developmental processes* (pp. 25-49). New York, NY: Cambridge University Press.
- Bronfenbrenner, U. (1992). Ecological systems theory. In R. Vasta (Ed.), *Six theories of child development: revised formulations and current issues* (pp. 187-249). London, UK: Jessica Kingsley Publishing.
- Bronfenbrenner, U. (2001). The theory of human development. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (Vol. 10, pp. 6963-6970). New York: Elsevier.
- Bronfenbrenner, U., & Morris, P. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Development, Volume 1, Theoretical Models of Human Development*, (6th ed.),

- 793-828. Hoboken, NJ: John Wiley & Sons.
- Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K., & Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology, 99*, 353–395.
- Caughy, M. O., O’Campo, P. J., & Muntaner, C. (2003). When being alone might be better: Neighborhood poverty, social capital, and child mental health. *Social Science & Medicine, 57*, 227-237.
- CCSD (2004). Data from the Canadian Council on Social Development, retrieved on February 15, 2008, from http://www.ccsd.ca/factsheets/economic_security/poverty/index.htm
- Cohen, N. J., Kiefer, H., & Pape, B. (2005). *Handle with care: Strategies for promoting the mental health of young children in community-based child care*. Toronto, ON: Hincks-Dellcrest. Retrieved on November 20, 2007, from <http://www.hincksdellcrest.org/institute/documents/FinalReportMentalHealthReport11.pdf>
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*, 95-121.
- Cronbach, L. J., & Meehl P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.
- Cuban, L. (1992). Why some reforms last: The case of the Kindergarten. *American Journal of Education, 100*, 166-194.
- Dalton, J. H., Elias, M. J., & Wandersman, A. (2007). *Community psychology: Linking individuals to communities* (2nd ed.). Belmont, CA: Wadsworth Publishing.

- Doherty, G. (1997). *Zero to six: The basis for school readiness* (No. R-97-8E). Ottawa, ON: Human Resources Development Canada, Applied Research Branch, Strategic Policy.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446.
- Elder jr., G. H. (1994). Time, human agency, and social change: Perspectives on the life course. *Social Psychology Quarterly*, *57*, 4-15.
- Elias, M. J., Zins, J. E., Graczyk, P. A., & Weissberg, R. P. (2003). Implementation, sustainability, and scaling-up of social-emotional and academic innovations in public schools. *School Psychology Review*, *32*, 303-319.
- Ellwein, M. C., Walsh, D. J., Eads II, G. M., & Miller, A. (1991). Using readiness tests to route Kindergarten students: The snarled intersection of psychometrics, policy, and practice. *Educational Evaluation and Policy Analysis*, *13*, 159-175.
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *The Sociology of Education*, *80*, 114-138.
- Gardner, H. (1999). *Intelligence Reframed: Multiple Intelligences for the 21st Century*. New York, NY: Basic Books.
- Governor General of Canada (September, 1997). *Speech from the Throne*. Retrieved on November 20, 2007, from <http://www.gg.ca/media/doc.asp?lang=e&DocID=190>
- Graue, M. E. (1995). Connecting visions of authentic assessment to the realities of educational practice. In T. A. Romberg (Ed.), *Reform in school mathematics and authentic assessment*, pp. 260–275. Albany, NY: State University of New York

- Press.
- Guhn, M., Janus, M., & Hertzman, C. (2007). The Early Development Instrument: Translating school readiness assessment into community actions and policy planning. *Early Education & Development, 18*, 369-374.
- Hebb, D. O., Lambert, W. E., & Tucker, G. R. (1971). Language, thought, and experience. *The Modern Language Journal, 55*, 212-222.
- Heymann, J., Hertzman, C., Barer, M. L., & Evans, R. G. (Eds.) (2006). *Healthier societies: From analysis to action*. New York, NY: Oxford University Press.
- Janus, M., Brinkman, S., Duku, E., Hertzman, C., Santos, R., Sayers, M., et al. (2007). *The Early Development Instrument: A population-based measure for communities. A handbook on development, properties, and use*. Hamilton, ON: Offord Centre for Child Studies. Retrieved on June 10, 2008, from http://www.offordcentre.com/readiness/pubs/2007_12_FINAL.EDI.HANDBOOK.pdf
- Janus, M., & Offord, D. R. (2000). Readiness to learn at school. *Canadian Journal of Policy Research, 2*, 71–75.
- 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*, 1–22.
- Jencks, C., & Mayer, S. (1990). The social consequences of growing up in a poor neighbourhood. In L. E. Lynn & M. G. H. McGeary (Eds.), *Inner-city poverty in the United States* (pp. 111–186). Washington, DC: National Academy Press.

- Kagan, S. L., & Rigby, E. (2003). *Policy matters: Setting and measuring benchmarks for state policies. Improving the readiness of children for school: Recommendations for state policy*. Washington, DC: Center for the Study of Social Policy.
- Kane, M. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th edition). Westport, CT: American Council on Education/Praeger.
- Kessler, S. A., & Swadener, B. B. (Eds.) (1992). *Reconceptualizing the early childhood curriculum: Beginning the dialogue*. New York, NY: Teachers College Press.
- Kershaw, P., Irwin, L., Trafford, K., & Hertzman, C. (2005). *The British Columbia atlas of child development* (1st ed., Vol. 40). Vancouver, BC: Human Early Learning Partnership, Western Geographical Press.
- Kohen, D. E., Brooks-Gunn, J., Leventhal, T., & Hertzman, C. (2002). Neighborhood income and physical and social disorder in Canada: Associations with young children's competencies. *Child Development, 73*, 1844-1860.
- Kohen, D. E., Hertzman, C., & Brooks-Gunn, J. (1998). *Neighborhood influences on children's school readiness*. Hull, QB: Human Resources Development Canada.
- La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research, 70*, 443-484.
- LaPointe, V. R., Ford, L., & Zumbo, B. D. (2007). Examining the relationship between neighborhood environment and school readiness for Kindergarten children. *Early Education & Development, 18*, 473-495.
- Lewin, K. (1951). *Field theory in social science*. New York, NY: Harper & Row.

- Lewit, E. M., & Schuurmann Baker, L. (1995) School readiness. *The Future of Children*, 5, 128-139.
- Lloyd, J. E. V., & Hertzman, C. (2009). From Kindergarten readiness to fourth-grade assessment: Longitudinal analysis with linked population data. *Social Science and Medicine*, 68, 111-123.
- Love, J. M., Aber, J. L., & Brooks-Gunn, J. (1994). *Strategies for assessing community progress toward achieving the First National Educational Goal*. Princeton, NJ: Mathematica Policy Research.
- Mead, G. H. (1934). *Mind, Self, and Society from the standpoint of a social behaviorist*, (C.W. Morris, Ed.). Chicago, IL: University of Chicago Press.
- Meisels, S. (1999). Assessing readiness. In R. C. Pianta, & M. J. Cox (Eds.), *The transition to Kindergarten* (pp. 39-66). Baltimore, MD: Paul H. Brookes.
- Messick, S. (1995). Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- NAEYC (1996). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8*. Washington, DC: National Association for the Education of Young Children. Retrieved on November 20, 2007, from <http://www.naeyc.org/about/positions/pdf/PSDAP98.pdf>
- National Education Goals Panel (1995). *The National Education Goals report*. Washington, DC: US Government. Retrieved on November 20, 2007, from http://eric.ed.gov:80/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/41/7e.pdf

- NHSA (2005, September). National Head Start Association mourns death of Head Start founding father Dr. Urie Bronfenbrenner. Alexandria, VA: National Head Start Association. Retrieved on November 20, 2007, from http://www.nhsa.org/press/News_Archived/2005/index_news_092705.htm
- Nosbush, L. (2006). *Community report for children. The Understanding the Early Years final report*. Prince Albert, SK: Saskatchewan Rivers School Division. Retrieved on November 20, 2007, from <http://www.srsd119.ca/uey/>
- Offord, D. R., Boyle, M. H., Szatmari, P., Rae-Grant, N. I., Links, P. S., Cadman, D. T., et al. (1987). Ontario Child Health Study. II. Six-month prevalence of disorder and rates of service utilization. *Archives of General Psychiatry*, 44, 832-836.
- Pelletier, J., & Corter, C. (2005). Design, implementation, and outcomes of a school readiness program for diverse families. *The School Community Journal*, 89-116.
- Pianta, R. C., Cox, M. J., & Snow, K. L. (Eds.) (2007). *School readiness and the transition to Kindergarten in the era of accountability*. Baltimore, MD: Paul H. Brooks.
- Piantadosi, S., Byar, D. P., & Green, S. B. (1988). The ecological fallacy. *American Journal of Epidemiology*, 127, 893-904.
- Piotrkowski, C. S. (2004). A community-based approach to school readiness in Head Start. In E. Zigler & S. Styfco (Eds). *The Head Start debates* (pp. 129-142). New Haven: Yale University Press.
- Putnam, R. D. (2000). *Bowling alone. The collapse and revival of American community*. New York, NY: Simon & Schuster.

- Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical Linear Models (2nd ed.)*. Thousand Oaks, CA: Sage Publications.
- Raudenbush, S. W., & Sampson, R. S. (1999). 'Ecometrics': Toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology*, 29, 1-41.
- Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An ecological perspective on the transition to Kindergarten: A theoretical framework to guide empirical research. *Journal of Applied Developmental Psychology*, 21, 491-511.
- Sampson, R. J., Morenoff, J. D., & Earls, F. (1999). Beyond social capital: spatial dynamics of collective efficacy for children. *American Sociological Review*, 64, 633–660.
- Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing “neighborhood effects”: social processes and new directions in research. *Annual Review of Sociology*, 28, 443-478.
- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*, 277, 918–924.
- Schrag, R. D. A., Styfco, S. J., & Zigler, E. (2004). Familiar concept, new name: Social competence/school readiness as the goal of Head Start. In E. Zigler & S. J. Styfco (Eds.), *The Head Start debates*. Baltimore, MD: Paul H. Brookes.
- Schweinhart, L. J., Barnes, H. V., & Weikart, D. P. (1993). Significant benefits: The High/Slope Perry Preschool Study through age 27. *Monographs of the High-Slope Educational Research Foundation*, No. 10. Ypsilanti, MI: High/Slope Press.

- Shepard, L. A. (1997). Children not ready to learn? The invalidity of school readiness testing. *Psychology in the Schools, 34*, 85-97.
- Smith, G. T. (2005). On construct validity: Issues of method and measurement. *Psychological Assessment, 17*, 396-408.
- Spady, D. W., Schopflocher, D. P., Svenson, L. W., & Thompson, A. H. (2001). Prevalence of mental disorders in children living in Alberta, Canada, as determined from physician billing data. *Archives of Pediatrics & Adolescent Medicine, 155*, 1153 - 1159.
- Statistics Canada (2007). *National Longitudinal Study of Children and Youth (NLSCY)*. Retrieved on Feb 15, 2008, from <http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=4450&lang=en&db=imdb&adm=8&dis=2>
- Taylor, C., Gorard, S., & Fitz, J. (2003). The modifiable area unit problem: Segregation between schools and levels of analysis. *International Journal of Social Research Methodology, 6*, 41-60.
- Tremblay, M. S., & Willms, J. D. (2000). Secular trends in the body mass index of Canadian children. *Canadian Medical Association Journal, 163*, 1429-1433.
- Vygotsky, L. S. (1936/1978). *Mind in Society: The development of higher psychological processes*. Cambridge, MA : Harvard University Press.
- Watson, M., & Ecken, L. (2003). *Learning to trust*. San Francisco, CA: John Wiley & Sons.
- Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press.

Zigler, E., Gilliam, W. S., Jones, S. M. (2006). *A vision for universal preschool education*. New York, NY: Cambridge University Press.

Zigler, E., & Valentine, J. (1979). *Project Head Start: A legacy of the war on poverty*. New York, NY: The Free Press.

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, D.C.: American Psychological Association Press.

3. Does the EDI measure school readiness in the same way across different groups of children?¹⁸

The Early Development Instrument (EDI; Offord & Janus, 1999) is “a teacher-completed measure of children’s school readiness at entry to grade 1 [that] was designed to provide communities¹⁹ with a feasible, acceptable, and psychometrically reliable instrument that [can] be used for whole populations of children to monitor community efforts to improve early years’ outcomes over time” (p.12, Janus & Offord, 2007). In other words, the EDI is a community tool to assess “early years’ outcomes” or “school readiness” at an aggregated community or population level.

In the fields of public health and epidemiology, this concept of measuring and reporting certain characteristics (e.g., health outcomes, incidence of illnesses) of people at a population level is very common. In education, this approach is less common, but has also been used (e.g., when reporting achievement scores, drop-out rates). However, in regard to characteristics of children, and particularly in regard to a holistic view of school readiness as assessed by the EDI, this is a relatively new approach. Traditionally, children’s school readiness has been assessed at an individual level, for the purpose of assigning individualized prevention and intervention strategies to children with perceived needs.

¹⁸ A version of this chapter has been published. Guhn, M., Gadermann, A., & Zumbo, B. D. (2007). Does the EDI measure school readiness in the same way across different groups of children? *Early Education and Development, 18*, 453-472.

¹⁹ Here, solely the term *community* is used, to refer to the concept that is delineated in the British Columbia Atlas of Child Development (cf. Kershaw, Irwin, Trafford, & Hertzman, 2005)—in which it is used synonymously with *neighborhood*.

Thus, the construct of school readiness as defined by the EDI is to be understood quite differently than the traditional notion of school readiness; that is, school readiness is seen as a characteristic of interest at an aggregated group level (e.g., community or school), not at an individual level. Accordingly, interpretations of EDI-data are advised to be solely conducted at such group (i.e., community, school, etc.) level, and the EDI is explicitly not a tool to diagnose (and to thus label) individual children. Despite the fact that EDI data are aggregated and then interpreted at the aggregated level, the data are nonetheless obtained at an individual level, namely via teaching ratings. Thus, the fact that EDI data are interpreted at the aggregate level does not mitigate the need to examine psychometric properties of the EDI in regard to the individual level data, because a systematic bias at the individual level could result in systematic bias at the aggregate level.

This paper investigates item bias, a pivotal aspect of test validity, for the EDI. Investigating item bias is important, because it presents a threat to the validity and fairness of a test or scale (Kristjansson, Aylesworth, McDowell, & Zumbo, 2005; Zumbo, 2007b). Bias at the item level, if large enough, may translate to bias at the domain or scale score level. This is of particular relevance for tests that are (primarily) interpreted at the domain or scale score level, as is the case for the EDI. Accordingly, it is also investigated whether cumulative effects of (potential) bias at the item level lead to bias at the domain score level.

With regard to validity, these issues are of particular importance, because inferences that are made on the basis of domain or scale scores that are biased are not

equally appropriate, meaningful, and useful for different sub-groups of the target population. This highlights that the investigation of test and item bias is targeted at establishing the inferential limits of a test, that is, for which group of respondents the inferences made on the basis of the test scores are valid and for which they are not (Zumbo, 2007a; 2007b). In broad terms, this is a matter of measurement invariance; that is, is the EDI performing in the same manner for each group of examinees (e.g., boys and girls)?

It is important at this point to highlight how the EDI is administered. Kindergarten teachers are asked to rate each of their students on 103 items, which are then separated into five developmental domains: (i) Physical health and well-being, (ii) Social competence, (iii) Emotional maturity, (iv) Language and cognitive development, and (v) Communication skills and general knowledge (cf. Janus & Duku, 2007). Inferences from the EDI are then based, as noted above, at a group level (e.g., a community) on the five domains.

Because the EDI items involve binary and rating (Likert) response formats, ordinal logistic regression is employed in order to examine item bias (Zumbo, 1999). Ordinal logistic regression is a method to statistically identify the presence of *Differential Item Functioning* (DIF). The presence of DIF at the item level is a necessary, but not sufficient condition for item bias (Camilli & Shepard, 1994). That is, if an item is flagged as displaying DIF, it does not necessarily mean that item bias is present. Rather, one has to ascertain whether the statistical presence of DIF is due to item bias or item impact. The following definitions (cf. Zumbo, 1999) illustrate the distinction between these different terms:

Differential Item Functioning. DIF occurs when examinees from different groups have differing probabilities of responding correctly to (or endorsing) an item after matching on the underlying ability that the item is intended to measure. The existence of DIF—a statistical property—indicates the presence of either item impact or item bias, and the distinction between the two cannot be inferred by statistical analysis alone.

Item impact. The presence of DIF indicates item impact when examinees from different groups have differing probabilities of responding correctly to (or endorsing) an item due to *true* differences between the groups in the underlying ability being measured by the item.

Item bias. The presence of DIF indicates item bias when examinees from different groups have differing probabilities of responding correctly to (or endorsing) an item due to differences between the groups in regard to some characteristic of the test item or testing situation that is *not relevant* to the test purpose.

These definitions illustrate that DIF is, as mentioned above, a necessary but not sufficient (statistical) condition for item bias. Therefore, in the case of the statistical presence of DIF, subject matter experts should be consulted to differentiate theoretically and conceptually between item bias—implying that the item is measuring construct-irrelevant differences—and item impact—implying that the item is measuring construct-related differences (Camilli & Shepard, 1994).

It thus needs to be emphasized that a statistical examination of items can solely indicate the presence of DIF; the statistical analysis itself cannot make a distinction between item bias and item impact.

Procedures to identify DIF, and thus potential item bias, are frequently used in the process of developing and adapting educational and psychological tests measures, as well as for the validation of test score inferences. In particular, the analysis of DIF is performed to examine five issues that are foundational for establishing test validity (Zumbo, 2007a): (a) fairness and equity in testing for test participants from different groups, (b) ruling out measurement artifact as potential threat to internal validity, (c) identifying group differences in item responding that—pending further investigation—arise from group differences that are either criterion-relevant or irrelevant, such as differences in ability, differences in cognitive processing, and/or differences in contextual or psycho-social factors, (d) translation and/or adaptation of measures to different languages or cultures, and (e) as part of item response theory and other such latent variable modeling.

This paper primarily examines the first two issues, with a passing nod to the third. In this context, it is important to recall that the EDI is filled out by the Kindergarten teacher, and not by the children themselves. Accordingly, any DIF on the EDI is to be understood as a difference between groups *with respect to the perception and rating of the Kindergarten teachers*. Therefore, for the EDI, the issues of (a) fairness and equity in testing, (b) ruling out measurement artifact as potential threat to internal validity, and (c) identifying group differences in item responding are all to be interpreted in light of the fact that the ratings reflect the perceptions of the Kindergarten teachers.

Research Objective

The EDI is, as is mentioned above, a *community tool* to measure school readiness of groups of children. The implications for research investigating the validity of the EDI is that one needs to examine the decisions and inferences that are made based on EDI data at a group (e.g., community or population) level. According to this purpose, the reporting of EDI results has occurred by grouping children at the community level, as well as at the school district or health district level (Kershaw et al., 2005).

The EDI has been used across diverse communities and school or health districts within Canada (and also in Europe, Australia, and South America; Janus et al., 2007). Given this diversity of communities, it is pivotal to examine DIF, in order to allow for meaningful comparisons across these communities and districts. After all, it is important to find out whether the EDI is measuring school readiness similarly across different groups of children and, likewise, across communities with diverse compositions of groups of children (Zumbo & Gelin, 2005). The analyses address this issue within the context of the Canadian province of British Columbia (BC). BC (equal in geographical size of France, Germany, and the Netherlands combined, and with a population of about 4½ million people) is made up of about 500 communities.

The definition and boundaries of these communities are based on research with, and reports of, the people living in these communities. These communities differ largely in respect to their demographic, cultural, geographic, and socioeconomic characteristics (Kershaw et al., 2005). Accordingly, the question was which criterion for the grouping of Kindergarten children would have significance for the BC context, and should therefore be used for the DIF analyses.

It was decided to focus on three criteria, namely (i) students' gender, (ii) students' English-as-a-Second-Language (ESL) status (i.e., ESL versus non-ESL/native speakers), and (iii) students' Aboriginal status (i.e., Aboriginal versus non-Aboriginal background). These groupings have commonly been used in developmental research. Gender differences in regard to school readiness are of general interest to developmental researchers as well as educators and parents (e.g., Angenent & de Man, 1989; Dauber, Alexander, & Entwisle, 1993; Duncan et al., 2006; Gullo & Burton, 1992; McCoy & Reynolds, 1999). Examining DIF—and thus the presence of item bias or item impact—with regard to gender contributes important information as far as the interpretation of gender differences is concerned.

Likewise, differences in school readiness with regard to ESL status are also of importance to educators and others, particularly in regard to language, reading, and writing acquisition (e.g., Chiappe & Siegel, 1999; Lesaux & Siegel, 2003). In this area, an examination of DIF with respect to ESL status can contribute important interpretative information towards, for example, policies regarding language instruction and educational support for ESL children.

Finally, examining DIF with respect to Aboriginal status is of particular socio-political relevance, because it is associated with issues regarding the cultural relationship between Aboriginal and European immigrant education (see Miller, 1996, for a historical account of schooling and education from an Aboriginal perspective).

Method

Sample

The sample consists of 43,900 Kindergarten children from BC. Data collection occurred during the spring terms of five consecutive school years, 1999-2000 until 2003-04. Of the children, 49% are female, 51% are male. According to the information provided by the Kindergarten teachers on the EDI, 17% of the children are non-native speakers (i.e., ESL), and about 7% are Aboriginal. A comparison of the EDI data set with a data set from the BC Ministry of Education in regard to the designations 'ESL/non-ESL' and 'Aboriginal/Non-Aboriginal' showed that both the groups of 'ESL' and 'Aboriginal' children are slightly under-represented; in other words, on the EDI, Kindergarten teachers did not assign 'ESL' status or 'Aboriginal' status to as many children as are indicated as such by the Ministry data.

Therefore, it was examined whether this under-representation occurred in a systematic way. However, a comparison between the groups (Group 1: children for which the ESL or Aboriginal designation in the EDI data base coincides with the Ministry's designation; Group 2: children that are designated as ESL or Aboriginal solely by the Ministry data²⁰) shows that there are no statistically significant differences in regard to the groups' respective EDI scores. Given this finding, it could be assumed for the further analyses that the EDI results were not systematically influenced (i.e., biased) by an under- or misrepresentation of ESL Kindergarten children due to differences in teachers' and the Ministry's classification criteria for ESL status.

²⁰ In British Columbia, both the EDI and the Ministry of Education data include individual child information. Thus, children that are identified as ESL or Aboriginal in the EDI data base, but not in the Ministry data base, can individually be identified.

Participation in the EDI survey is voluntary, even though it is facilitated and supported by the Ministry of Education. Overall, participation was very high, with representation from all 59 school districts in BC. Of the schools that opted out, a relatively high number of schools were among those that are located on Aboriginal reserves. There are undoubtedly a number of different reasons for Aboriginal schools opting out of the EDI administration. It should be noted though, that Aboriginal communities in BC – indeed, across Canada – have suffered a long history of oppression, and that much physical and psychological pain has been inflicted by the formal public 'education' system and various systems of assessment and student evaluation. This history of oppression and abuse may account for some of the reluctance to participate in what may be seen as an attempt to control, categorize, and discriminate against Aboriginal children. For an account of this history, the reader is referred to Miller (1996).

Measure

All children were rated on the EDI by their Kindergarten teachers. The EDI contains demographic information (e.g., gender, age, first language, Aboriginal background) and 103 binary and Likert-scale items on five developmental domains: Physical health and well-being (13 items), Social competence (26 items), Emotional maturity (30 items), Language and cognitive development (26 items), and Communication skills and general knowledge (8 items).

A sample question²¹ of the EDI, for the Communication skills and general language domain, is: “How would you rate this child’s ability to tell a story?”, with the response options ‘very good/good’, ‘average’, ‘poor/very poor’, and ‘I don’t know’. For data analysis purposes, all responses on binary items are coded as ‘0’ or ‘10’, 3-point Likert-scale items are coded ‘0’, ‘5’, and ‘10’, or 5-point Likert-scale item are coded ‘0’, ‘2.5’, ‘5’, ‘7.5’, and ‘10’. All items contain an additional response option, ‘I don’t know’ (coded as ‘99’), which is not included in the statistical analyses. For every item, ‘10’ designates the highest (i.e., most positive, most developmentally desirable) score.

For every domain, the average score is calculated, ranging from ‘0’ to ‘10’. In addition, the five domain scores are combined into a total EDI score, ranging from ‘0’ to ‘50’. It needs to be noted that solely the domain scores are reported in practice (Kershaw et al., 2005²², Janus et al., 2007), in concert with the recommendations of the authors of the EDI; here, the total score is purely used for methodological research purposes.

DIF Analyses

The following section provides a brief non-technical introduction to DIF analysis using ordinal logistic regression. For a comprehensive, in-depth coverage of the method, the interested reader is referred to Zumbo (1999; 2007b), Shimizu and Zumbo (2005), and Slocum, Gelin, and Zumbo (in press).

There are several ways to examine DIF, and thus measurement and test bias, statistically. For tests that consist of items with binary and ordinal (e.g., Likert-scale)

²¹ The EDI is available (English & French) at: www.offordcentre.com/readiness/EDI_viewonly.html

²² *The British Columbia Atlas of Child Development* is available at: www.earlylearning.ubc.ca

response format, Zumbo (1999) has developed a method that integrates binary and ordinal logistic regression. In this method, as the first step, groups of participants (e.g., boys and girls) are matched on the variable of interest (e.g., the total EDI score as indicator of ‘overall school readiness’). Then, the probability of obtaining a certain score on the item under investigation is calculated for both groups, for each total EDI score level, respectively. Accordingly, the logistic regression model includes variables to represent (i) the groups, (ii) the score for the variable of interest, and (iii) the interaction between the group status and the score of the variable of interest (Shimizu & Zumbo, 2005). For the analysis of DIF, the predictor variables are not entered simultaneously, but successively for the following three models (Zumbo, 1999; Gelin, Carleton, Smith, & Zumbo, 2004):

- Model 1: The conditioning variable (i.e., the total EDI score) is the sole predictor.
- Model 2: The conditioning variable (i.e., the total EDI score) *and the grouping variable* are in the equation.
- Model 3: The conditioning variable (i.e., the total EDI score), the grouping variable, *and the interaction term representing the interaction of the total EDI score and the grouping variable* are in the equation.

These three models correspond to the following three equations, in which ‘ $y^*_{\text{item score}}$ ’ represents the predicted item score; ‘ $\text{TOTAL}_{\text{EDI score}}$ ’ denotes the conditioning variable, the total score of the EDI; ‘GROUP’ refers to the grouping variables gender, ESL, or Aboriginal status; ‘ $\text{TOTAL}_{\text{EDI score}} * \text{GROUP}$ ’ represents the interaction term

between the total EDI score and either gender, or ESL, or Aboriginal status; and 'e' designates the error term.

$$\text{Model 1: } y^*_{\text{item score}} = b_0 + b_1 \text{TOTAL}_{\text{EDI score}} + e$$

$$\text{Model 2: } y^*_{\text{item score}} = b_0 + b_1 \text{TOTAL}_{\text{EDI score}} + b_2 \text{GROUP} + e$$

$$\text{Model 3: } y^*_{\text{item score}} = b_0 + b_1 \text{TOTAL}_{\text{EDI score}} + b_2 \text{GROUP} + b_3 (\text{TOTAL}_{\text{EDI score}} * \text{GROUP}) + e$$

This sequence allows one to calculate how much variance the grouping variable (in Model 2) explains over and above the conditioning (i.e., matching) variable (in Model 1). The difference between Model 1 and Model 2 can then be tested for significance via a χ^2 test, and an effect size can be calculated via a comparison of the R^2 values. Similarly, a comparison of Model 2 and 3 allows one to calculate how much variance the interaction term (in Model 3) explains over and above the effects of the conditioning and grouping variable (in Model 2) (Zumbo, 1999). In other words, this analysis allows one to address the following questions: (i) Is there a significant group difference? If so, what is its effect size?, and (ii) Is there a significant interaction? If so, what is its effects size?

Accordingly, in DIF terminology, *uniform* DIF refers to the group differences (i.e., the main effect, comparing Models 1 and 2), and *non-uniform* DIF refers to the group by total score interaction (i.e., the interaction effect, comparing Models 2 and 3). For the interpretation of effect sizes from ordinal logistic regression DIF analyses, Jodoin and Gierl (2001) have suggested guidelines. According to their criteria, effect sizes of $R^2 < .035$ are considered negligible, those between .035 and .070 moderate, and the ones with $R^2 \geq .070$ large.

Statistical analysis

Ordinal logistic regression DIF methodology was employed for the analyses (Zumbo, 1999). The existence of DIF is examined for each of the 103 items of the EDI, for each of the following group comparisons, respectively: (i) Gender (Girls versus Boys); (ii) ESL designation (ESL versus non-ESL); and (iii) Aboriginal background ('Aboriginal' versus 'non-Aboriginal').

Models 1 through 3, as described above, are fit for each of the 103 items, separately. For every analysis, Kindergarten children are matched based on their total EDI score. Although the total EDI score is not reported in practice (see above), conditioning (i.e., matching) on the total score was done for the following two reasons: (i) a factor analysis of the items indicates that there is one dominant factor (suggesting that the total score is a proxy for a child's 'overall school readiness'), and (ii) when matching on domain scores (with different scale lengths), each item has a different and, potentially, relatively large contribution to the matching criterion (e.g., for the Communication skills and general knowledge domain, consisting of 8 items, each item would contribute an eighth to the matching score).

In a second step, for those items flagged with DIF—using the criteria for statistical significance as well as the effect size criteria proposed by Jodoin and Gierl (2001); see above—the analyses investigated whether DIF items have an effect at the domain level. In other words, the analyses examined whether DIF of a single or multiple items on one domain resulted in Differential Functioning at the domain level. This is of

particular relevance, because EDI scores are solely reported at the domain scale level, and hence one would want to see the cumulative effect of the item level properties on the domain score. The examination is done graphically, as statistical significance tests and effect size estimations for the analysis of differential domain-level functioning are yet to be developed.

In a third step, after the matching on the total EDI score, the probabilities for obtaining a certain domain score, for each of the groups being compared, and for the entire range of the matching score are calculated by adding up the predicted item scores of the domain score under investigation. These domain score probabilities can then be presented graphically in a curve that is the domain-level equivalent of an item response function from Item Response Theory. The total is then divided by the number of items on the domain scale for ease of interpretation (the reported EDI domain scores are also average scores, likewise ranging from 0 to 10.). The predicted average domain scores are then plotted for the respective group comparisons to visually represent the differential domain functioning. In essence, in the language of psychometrics, the item characteristic curves are 'translated' into domain level characteristic curves, which are then plotted and compared on the same graph.

The last step in the analysis was to conceptually examine whether those items flagged with DIF were indicative of item bias or of item impact. Accordingly, subject matter experts were consulted to scrutinize whether the findings coincided with the research literature (suggesting item impact) or whether the findings were more likely to be consequences of the measurement process (suggesting item bias).

Results

DIF Grouping Variables

In addition to the theoretical, educational, and socio-cultural reasons for selecting the grouping criteria discussed above (gender, ESL, and Aboriginal status), it is noteworthy that the EDI results for each of these three comparisons show statistically significant differences of substantial effect size. In Table 3.1, these differences are presented for each of the three group comparisons and for each of the five EDI domains individually. The differences are expressed in the raw score metric—the actual differences between the groups’ respective average domain scores, on a 10-point scale—and as standardized effect sizes (Cohen’s *d*). In the table, Cohen’s (1992) general guidelines for interpreting effect sizes are included.

Table 3.1 Mean Group Differences (on a 10-point Scale) and Effect Sizes (Cohen’s *d*)¹ for Group Comparisons.

<i>EDI domains</i>	<u>Gender comparison</u> (Means higher for girls)	<u>ESL comparison</u> (Means higher for native speakers)	<u>Aboriginal comparison</u> (Means higher for non-Aboriginal children)
Physical Health and Well-Being	.32 (<i>d</i> = .30)	.14 (<i>d</i> = .13)	.58 (<i>d</i> = .50)
Social Competence	.75 (<i>d</i> = .43)	.46 (<i>d</i> = .25)	.66 (<i>d</i> = .36)
Emotional Maturity	.77 (<i>d</i> = .52)	.32 (<i>d</i> = .21)	.51 (<i>d</i> = .31)
Language and Cognitive Development	.56 (<i>d</i> = .27)	.80 (<i>d</i> = .40)	.86 (<i>d</i> = .41)
Communication skills and General Knowledge	.56 (<i>d</i> = .29)	2.15 (<i>d</i> = 1.1)	.68 (<i>d</i> = .33)
Total EDI score ²	2.96 (<i>d</i> = .43)	3.86 (<i>d</i> = .56)	3.27 (<i>d</i> = .45)

¹ Effect sizes (Cohen’s *d*) of .2, .5, and .8 are considered as small, medium, and large, respectively (Cohen, 1992).

² Sum of domain scores; 50-point scale

DIF Analyses

ESL. For the ESL versus non-ESL comparison, seven items display (uniform) DIF. (No item displays non-uniform DIF.) Five of these items with DIF belong to the Communication skills and general knowledge domain (with a total of eight items), the other two to the Language and cognitive development domain (with a total of 26 items) (see Table 3.2). In Table 3.2, the three items of the Communication skills and general knowledge domain that display DIF of negligible effect size are also included (in parentheses, printed in italics).

Table 3.2 Effect Sizes for EDI Items with Uniform DIF in the ESL versus non-ESL comparison

<i>Communication skills and general knowledge domain (8 items)</i>	Effect size ³
How would you rate this child's ¹ ...	(ΔR^2)
(1) ...ability to use language effectively in English?	.091
(2) ...ability to listen in English?	.041
(3) ... ability to tell a story?	.067
(4) ... <i>ability to take part in imaginative play?</i>	[.020]
(5) ... ability to communicate own needs in a way understandable to adults and peers?	.035
(6) ... ability to understand on first try what is being said to him/her?	.035
(7) ... <i>ability to articulate clearly, without sound substitutions?</i>	[.028]
Would you say that this child ² ...	
(8) ... <i>answers questions showing knowledge about the world?</i>	[.033]
<i>Language and cognitive development domain (26 items)</i>	
Would you say that this child ² ...	
(I) ... is showing awareness of rhyming words?	.048
(II) ... understands simple time concepts (e.g., today, summer, bedtime)?	.036

¹Response options for these items are on a 3-point Likert scale: 'very good/good' (10); 'average' (5) 'poor/very poor' (0); and 'don't know'

²Response options for these items are binary: 'yes' (10); 'no' (0); and 'don't know'

³Effect sizes of $R^2 < .035$ are considered negligible, those between .035 and .070 moderate, and ones $\geq .070$ large (Jodoin & Gierl, 2001). Negligible effect sizes are put in parentheses.

In Figure 3.1, one item displaying DIF is represented graphically as an illustrative example. This item is the first item on the Communication skills and general knowledge domain, and the one with the largest DIF effect size ($\Delta R^2 = .091$; $p < .001$). An examination of the group differences at the domain score level (for the Communication skills and general knowledge domain) shows that ESL children, on average, receive lower scores (2-point difference on a 10-point scale) than their non-ESL counterparts when matched at the same level of school readiness (i.e., on the total EDI score). Corresponding to the term Differential Item Functioning, it seemed appropriate to refer to this difference as *Differential Domain Functioning*. Figure 3.2 graphically represents these group differences at the domain level (i.e., the Differential Domain Functioning). For the Language and cognitive development domain, the two items with DIF do not have an effect at the domain score level.

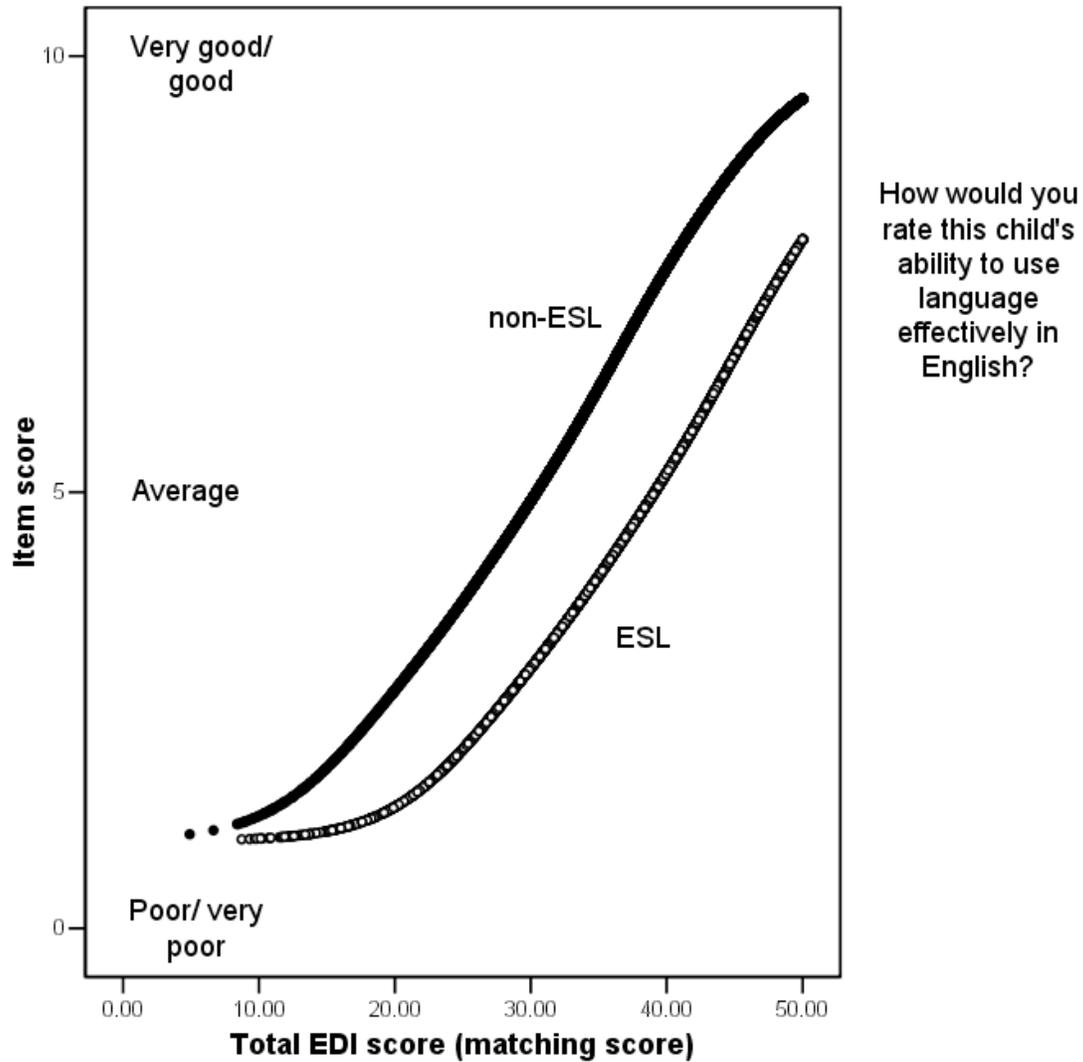


Figure 3.1 Item 'Ability to use English effectively', on the Communication skills and general knowledge domain, displaying DIF in ESL comparison

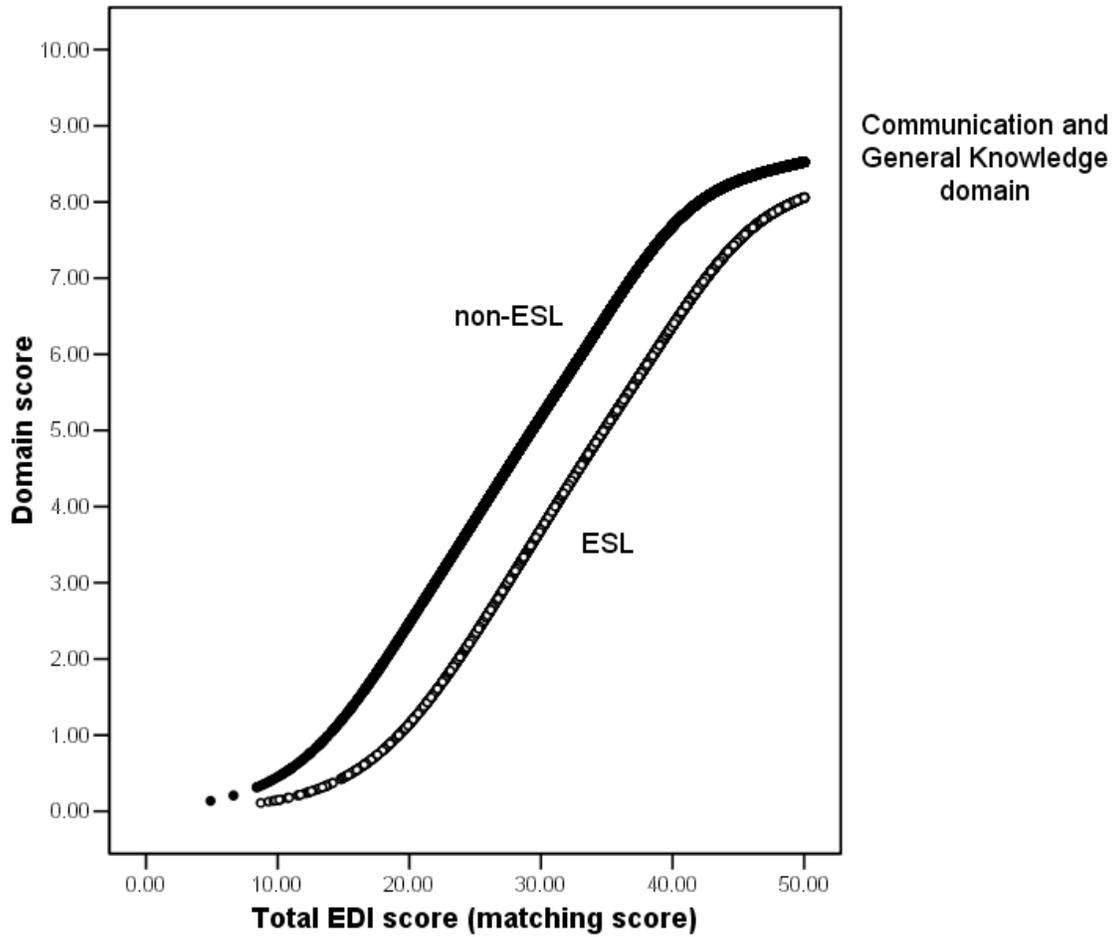


Figure 3.2 Differential Domain Functioning (Communication skills and general knowledge domain) due to (cumulative) item-level DIF

Gender. In the analysis for gender, one item shows (uniform) DIF, namely an item from the domain Emotional maturity, “Would you say this child gets into physical fights?”, with 5-point Likert-scale response format: ‘Often’ or ‘very true’ (0), ‘Sometimes’ or ‘somewhat true’ (5), ‘Never’ or ‘not true’ (10), and ‘Don’t know’. (No item displays non-uniform DIF.)

The effect size for the DIF of this item is $\Delta R^2 = .053$ ($p < .0001$). In Figure 3.3, the item score probabilities for boys and girls depending on their total EDI score are plotted. The graph illustrates that boys, on average, and at every total EDI score level (the matching criterion), have a higher probability of obtaining a teacher rating designating them as more²³ physically aggressive than girls. (Figure 3.3 is thus equivalent to Figure 3.1 in that both depict DIF at the item level.)

The graph in Figure 3.4 presents the probabilities for obtaining a certain score on the domain Emotional Maturity, for boys and girls, at every level of the matching score (i.e., total EDI score). As can be seen, the two plots are nearly identical, showing that the DIF on that one item has no effect at the domain score level. In other words, the graphic examination suggests that there is no substantial Differential Domain Functioning, despite the fact that one item on this scale displays DIF. (Figure 3.4 is equivalent to Figure 3.2 in that both depict the effect—or lack of effect—of the item DIF at the domain level.)

²³ Please note the score coding for this item in the graph of Figure 3.3. A low score indicates ‘more physically aggressive’, because the item is reverse coded.

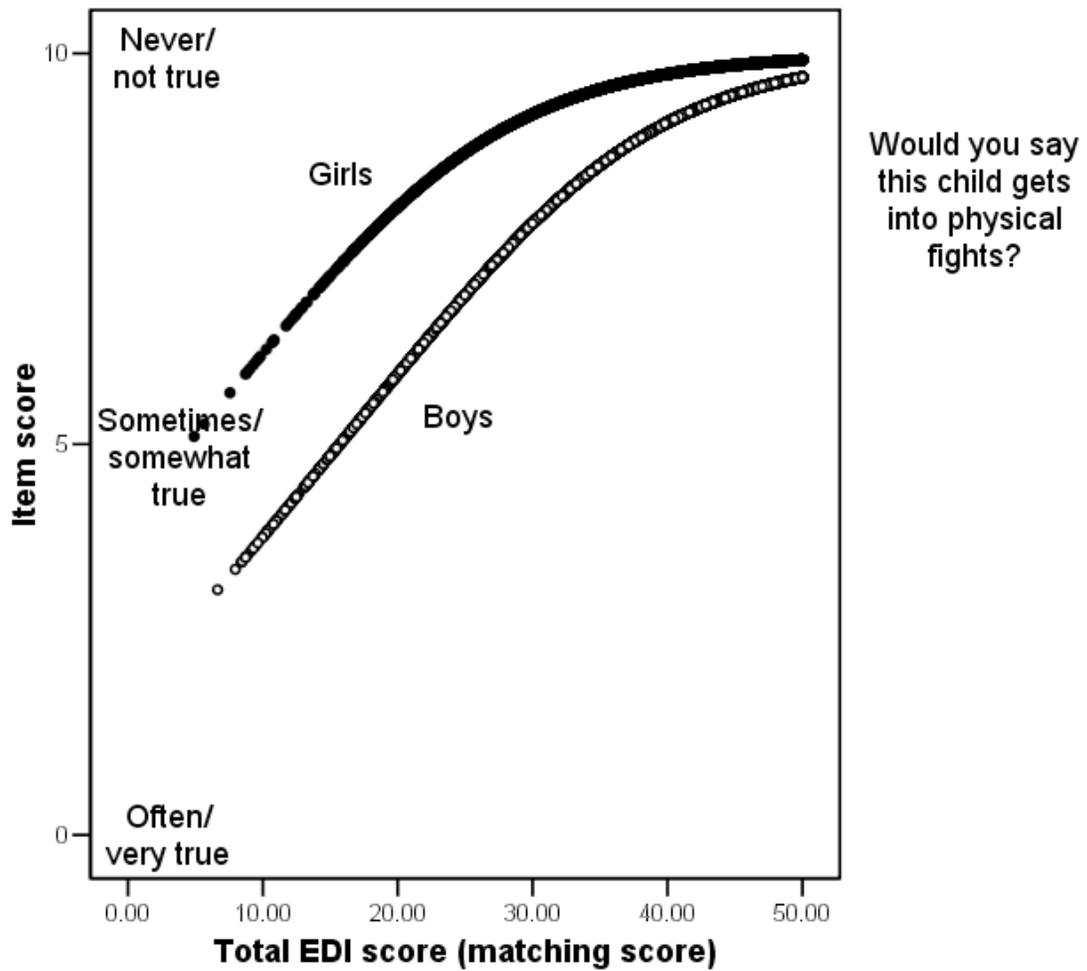


Figure 3.3 Item 'Gets into physical fights', on the Emotional maturity domain, displaying DIF in the gender comparison

Note: This item is reverse coded, so that low scores indicate higher physical aggression.

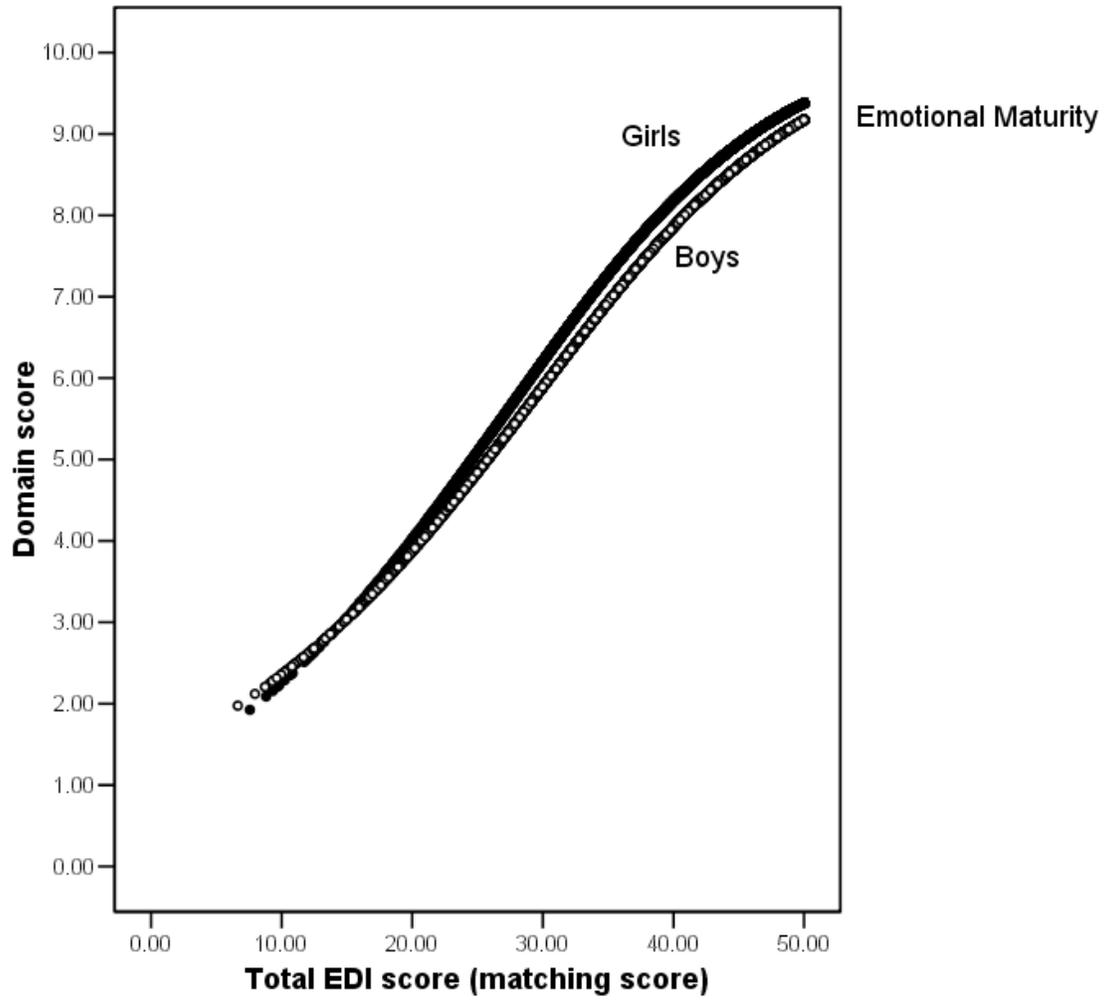


Figure 3.4 No Differential Domain Functioning (on Emotional maturity domain), despite item-level DIF

Aboriginal background. In the analysis comparing children designated as Aboriginal with non-Aboriginal, no item shows DIF.

Discussion

The DIF analysis identified several items that display uniform DIF of substantial effect size. In one case, DIF at the item level results in DIF at the domain level—or, to be more exact, in Differential Domain Functioning. In regard to fairness and measurement bias, the results support the validity of the EDI, and suggest that the EDI is measuring school readiness similarly across groups of Kindergarten children, grouped according to gender, ESL status, or Aboriginal background. In other words, the Kindergarten teacher's ratings of the children on the EDI are not biased by their perceptions of children's gender, ESL, or Aboriginal status. This general conclusion is based on arguments that suggest that all cases of DIF on the EDI are cases of item impact, meaning that group differences on these items reflect actual group differences on the underlying ability or skill that is being measured rather than construct-irrelevant variance. The following section discusses the results and some pertaining arguments in detail, and also addresses a number of questions raised by the results.

ESL status group comparison. Most of the items that display DIF are identified in the comparison between ESL and non-ESL children. In this comparison, seven items display DIF. Five of these items are on the Communication skills and general knowledge domain, and the other two on the Language and cognitive development domain. On the Language and cognitive development domain, the presence of two items (out of 26) with DIF does not have an effect on the average score for this domain. However, for the Communication skills and general knowledge domain, the (cumulative) presence of DIF

on five out of eight items does add up to the point that it clearly affects the domain level score²⁴.

The size of this effect at the domain score level is quite substantial as the following points illustrate: EDI results are, as mentioned, reported at the domain score level, for each community or district. For the five EDI domains, the ranges of the average scores—for the 59 school districts, and on the 10-point scale, are as follows: 8.08-9.08 (Physical health and well-being), 7.38-8.96 (Social competence), 7.35-8.71 (Emotional maturity), 7.44-9.02 (Language and cognitive development), and 6.29-8.36 (Communication skills and general knowledge). As the results indicate, the group difference between ESL and non-ESL children on this scale is about two points. Numerous communities in BC have more than 50% of their children with ESL status, and for such communities, the average domain score for Communication skills and general knowledge is going to be, on average, about one point below the one for communities without ESL children.

What does this mean in practical terms? For a district to drop by one point on the average score of the Communication skills and general knowledge domain is equivalent to dropping from the top quintile to the lowest quintile²⁵, and the same is true at the community level. Commonly, the relative ranking of districts or communities derived from the EDI average scores, as well as an associated percentage of vulnerable children within a district or community has been used as an argument to either back up funding

²⁴ Due to the absence of a statistical test, one cannot refer to this difference as *statistically* significant, even though the size implies *practical* significance.

²⁵ In the BC Atlas of Child Development (Kershaw et al., 2005), quintiles are used for the report of results at the community and district levels.

requests (in the case of perceived need; i.e., relatively low average scores) or to praise community initiatives or political action (in the case of relatively high scores).

Considering such usage of EDI scores, the question is what the implications of these findings are. Does the same Communication skills and general knowledge score have the same practical implications for a community with a high proportion of ESL children as opposed to one with a low proportion? And—taking this thought one step further—could separate group norms for ESL and non-ESL children provide an answer to this question? At least for one reason, providing ESL-specific norms seems to be ill-advised, as it could invite false inferences: It might convey the misleading perception that a level of Communication skills and general knowledge, which is considered insufficient for native English speaking children, is perceived as ‘normal’ for ESL children—with the implication that ‘normal’ might translate into the equivalent of meaning ‘ok’ or ‘acceptable’. From a societal and educational point of view, however, the goal ought to be that (almost) all children reach a sufficient level of school readiness, and thus communication skills, so that they can thrive in school. The challenge therefore is how communities and schools can jointly provide support for families and their children with relatively low English communication skills. By definition, a large proportion of these children comes from an ESL background, because the ESL designation is not assigned to children who are non-native speakers, but only to those children that (i) are non-native speakers and (ii) are deemed in need of targeted ESL language support in school.

An additional argument for advising against norms for ESL groups is that there is a wide variation among different sub-groups of the ESL population. Rather than masking this variation by providing overall ESL norms, it seems more appropriate to conduct

further investigations of specific ESL sub-groups in order to identify language-specific strengths and needs, which can then be addressed via educational support. At this point, it should also be noted that the ESL variable, as reported on the EDI, is a variable that is binary (ESL or not-ESL), even though children's English communication competence is distributed along a continuum, or even multiple continua (e.g., children might communicate well in some situations, but not well in others). Accordingly, it needs to be emphasized that there is some variation with regard to the communication competence (as rated by their teachers) among ESL and non-ESL children, which is not captured by the dichotomous nature of the ESL status variable. In addition, teachers might not be consistent in their assignment or reporting practice of ESL status. Finally, it needs to be noted that studies examining ESL in relation to school success have identified that the socioeconomic status of the children's families and communities has a strong relation to the children's academic achievement (for the BC context, see Toohey & Derwing, 2006). Analyses linking EDI scores to SES at the community level support this claim (Kershaw et al., 2005). In regard to the ESL comparison, it can thus be concluded that the DIF analysis identified group differences that, in fact, are to be expected on those items that refer to English communication skills. After all, that distinction is the main criterion for designating children as ESL. Accordingly, the displayed DIF most probably is item impact, and not item bias, as it refers to actually occurring differences between the groups on the characteristic that is measured.

Gender group comparison. For the gender comparison, one item with (uniform) DIF was identified. This item belongs to the EDI domain Emotional maturity and is

related to physical aggression (i.e.: “Would you say this child gets into physical fights?”). On this item, boys have a higher probability of obtaining a higher (i.e., more physically aggressive, because the item is reverse-coded) score than girls, after matching boys and girls on their total EDI school readiness score. In other words, boys with the same overall school readiness as girls are, on average, perceived and rated as more physically aggressive than girls by their Kindergarten teachers. This finding coincides with numerous child development studies, which suggest that boys, on average, tend to be more physically aggressive (e.g., Hyde, 1984; Alink et al., 2006). Therefore, it can be assumed that this finding also represents a case of item impact, and not item bias, as the statistical gender DIF can be assumed to reflect actual group differences with respect to the characteristic that is being measured (Emotional maturity, as a domain of school readiness). It is important to add that the DIF identified for this item has no effect at the domain score level. This can be attributed to the fact that all other items do not display DIF of substantial effect size, and because the influence of one item on a scale with 30 items is relatively small. It can thus be concluded that, in regard to gender, the reporting of EDI scores, which is solely done at the domain score level, is unaffected by the presence of DIF in one item.

Aboriginal background group comparison. In regard to the comparison between Aboriginal and non-Aboriginal children, the analyses did not identify any DIF, implying that the EDI is not affected by measurement bias in regard to children’s Aboriginal status. In other words, children’s Aboriginal status does not seem to systematically bias Kindergarten teacher’s ratings. This being said, it must, however, be emphasized that the

EDI results may not—despite the large sample size—be representative of the diversity of Aboriginal children and their communities, since numerous on-reserve schools (with a high number/proportion of Aboriginal children) opted out from participating in the EDI assessment²⁶.

Current and future EDI-related research has been and will have to continue to be conducted in order to validate the EDI in an ongoing and context-dependent manner. The studies in this special issue provide an illustrative overview of how research projects in combination address different aspects of validity. One question raised by the findings pertains to the group differences that are, for each of the three groupings, consistent across all five developmental domains of the EDI, and, in some cases, these group differences are of a large effect size.

The interested reader is referred to a number of studies that have examined related issues in the BC context (e.g., Ministry of Education of BC, 2006; Toohey & Derwing, 2006), the Canadian context (e.g., Bonneau & Lauzon, 2006; Statistics Canada & Bowlby, 2006), or in similar US contexts (e.g., Duncan et al., 2006). Further EDI-related studies examining these group differences specifically can hopefully provide further information to meaningfully interpret these differences. In addition, current and future EDI-related research has been and will have to continue to be conducted in order to validate the EDI comprehensively, in an ongoing and context-dependent manner. In fact,

²⁶ Currently, the Human Early Learning Partnership at the University of British Columbia, as the organization that coordinates the EDI project in British Columbia, is collaborating with numerous stakeholders towards developing an early childhood education tool that more clearly integrates Aboriginal values and their cultural diversity.

the studies in this special issue provide an illustrative overview of how research projects in combination address different aspects of validity.

The presented DIF analysis of the EDI data from Kindergarten children in BC presents results that are foundational for the establishment of the EDI's validity. Due to representativeness of this sample (nearing census dimensions) in regard to ethnic diversity, demographics, all socioeconomic status strata, community contexts, etc., the results may be assumed to also be generalizable to other North-American jurisdictions that share some of the same characteristics (e.g., high degree of ethnic diversity). In regard to Kindergarten teacher's perceptions and ratings of children, the results suggest that the EDI is fair and unbiased in regard to gender, ESL, and Aboriginal status. The identified cases of DIF can compellingly be related to research findings, and be referred to as item impact, accordingly. These findings provide useful information for the interpretation of other EDI-related research, as measurement bias can, to a certain degree, be ruled out as a confounding issue.

References

- Alink, L. R. A., Mesman, J., van Zeijl, J., Stolk, M. N., Juffer, F., Koot, H. M., et al. (2006). The Early Childhood Aggression Curve: Development of Physical Aggression in 10- to 50-Month-Old Children. *Child Development, 77*, 954-966.
- Angenent, H., & de Man, A. F. (1989). Intelligence, gender, social maturity, and school readiness in Dutch first-graders. *Social Behavior and Personality, 17*, 205-209.
- Bonneau, C. E., & Lauzon, J. (2006). First Nations learners and extracurricular activities: Barriers and bridges to participation. *Educational Insights, 10*(1).
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousands Oaks, CA: Sage Publications.
- Chiappe, P., & Siegel, L. S. (1999). Phonological awareness and reading acquisition in English- and Punjabi-speaking Canadian children. *Journal of Educational Psychology, 91*, 20-28.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.
- Dauber, S. L., Alexander, K. L., & Entwisle, D. R. (1993). Characteristics of retainees and early precursors of retention in grade: Who is held back? *Merrill-Palmer Quarterly, 39*, 326-343.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2006, November). *School readiness and later achievement*. Working paper, Institute for Policy Research, Northwestern University, Evanston, Illinois.
- Gelin, M. N., Carleton, B. C., Smith, M. A., & Zumbo, B. D. (2004). The dimensionality and gender differential item functioning of the Mini Asthma Quality of Life Questionnaire (MINIAQLQ). *Social Indicators Research, 68*, 91-105.

- Gullo, D. F., & Burton, C. B. (1992). Age of entry, preschool experience, and sex as antecedents of academic readiness in Kindergarten. *Early Childhood Research Quarterly, 7*, 175-186.
- Hyde, J. S. (1984). How large are gender differences in aggression? A developmental meta-analysis. *Developmental Psychology, 20*, 722-736.
- 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.*
- Janus, M., Brinkman, S., Duku, E., Hertzman, C., Santos, R., Sayers, M., & Schroeder, J. (2007). *The Early Development Instrument: A Population-based measure for communities. A Handbook on development, properties, and use.* Hamilton, Ontario: Offord Centre for Child Studies.
- Janus, M., & Duku, E. (2007). The school entry gap: Socioeconomic, family, and health factors associated with children's school readiness to learn. *Early Education and Development, 18*, 375-404.
- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type I error rate and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education, 14*, 329-349.
- Kershaw, P., Irwin, L., Trafford, K., & Hertzman, C. (2005). *The British Columbia Atlas of Child Development* (1st edition). Victoria, BC: Human Early Learning Partnership and Western Geographical Press.

- Kristjansson, E., Aylesworth, R., McDowell, I., & Zumbo, B. D. (2005). A comparison of four methods for detecting differential item functioning in ordered response items. *Educational and Psychological Measurement, 65*, 933-953.
- Lesaux, N., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology, 39*, 1005-1019.
- McCoy, A. R., & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology, 37*, 273-298.
- Miller, J. R. (1996). *Shingwauk's Vision: A History of Native Residential Schools*. Toronto: University of Toronto Press.
- Ministry of Education, British Columbia (2006, September). *Foundation Skills Assessment. 2001/02-2005/06*. Retrieved on November 10, 2006, from www.bced.gov.bc.ca/reporting/achieve/fsa-bas.php
- Offord, D., & Janus, M. (1999). *Early Development Instrument. A Population-based Measure for Communities, (2004/05 version)*. Offord Centre for Child Studies. McMaster University, Hamilton, Ontario. Retrieved on November 20, 2006, from www.offordcentre.com/readiness/EDI_viewonly.html
- Shimizu, Y., & Zumbo, B. D. (2005). A logistic regression for differential item functioning primer. *Japan Language Testing Association Journal, 7*, 110-124.
- Slocum, S. L., Gelin, M. N., & Zumbo, B. D. (in press). Statistical and graphical modeling to investigate differential item functioning for rating scale Likert item formats. In B. D. Zumbo (Ed.), *Developments in the theories and applications of measurement, evaluation, and research methodology across the disciplines*, Volume 1. Vancouver, BC: Edgeworth Laboratory, UBC.

- Statistics Canada, & Bowlby, G. (2006). *Provincial Drop-out rates - Trends and Consequences*. Retrieved on November 10, 2006, from www.statcan.ca/english/freepub/81-004-XIE/2005004/drop.htm
- Toohey, K., & Derwing, T. M. (2006, July). Hidden Losses: How Demographics can Encourage Incorrect Assumptions about ESL High School Students' Success. *Working Paper Series: Research on Immigration and Integration in the Metropolis*. Vancouver, BC: Simon Fraser University.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.
- Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. In C. R. Rao and Sinharay (Eds.), *Handbook of statistics, Vol. 27: Psychometrics*. Elsevier Science B. V.: The Netherlands.
- Zumbo, B. D. (2007). Three generations of differential item functioning (DIF) analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, 4, 223-233.
- Zumbo, B. D., & Gelin, M. N. (2005). A matter of test bias in educational policy research: Bringing the context into picture by investigating sociological/community moderated (or mediated) test and item bias. *Journal of Educational Research & Policy Studies*, 5, 1-23.

4. Language Background, Gender, and Socioeconomic Status Interactions Related to Kindergarten Children's Development²⁷

This paper examines the relationships between gender, language group (English, Punjabi, and Cantonese), and socioeconomic background with respect to the developmental status of Kindergarten-aged children for a cohort of 40,772 children in British Columbia, Canada. Data on the children's development in this analysis stem from Kindergarten teacher ratings on the Early Development Instrument (EDI; Janus & Offord, 2007), for the domains of physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge, and data on the socioeconomic status of their family from block-level tax filer data on median income. The focus of the descriptive and statistical analyses is on the interaction effects of children's gender, English-as-a-Second-Language (ESL) status, and first language background (English, Punjabi, and Cantonese) with socioeconomic status. In other words, the research examines whether the relationships between gender or language group status and developmental outcomes are consistent across the socioeconomic spectrum.

The purpose of the paper is twofold. First, the research aims to contribute to the empirical early childhood development research literature by examining a representative, population-based sample with regard to a number of developmental outcomes that

²⁷ A version of this chapter will be submitted for publication. Guhn, M., Gadermann, A., Hertzman, C., & Zumbo, B. D. Language background, gender, and socioeconomic status interactions related to Kindergarten children's development.

holistically reflect children's school readiness, and by examining whether there are any interaction effects between children's gender or language background and their families' socioeconomic background (as approximated by block-level income data) with regard to those developmental outcomes. Second, the research is conducted to also contribute theoretically to the child development literature, by exploring whether the previously proposed theories of *relative functionalism* (Sue & Okazaki, 1990) and *double jeopardy* (Willms, 2003) can also account for the findings from this study.

The overall conceptual approach in this paper is largely informed by propositions drawn from Bronfenbrenner's bioecological theory of human development (1979; 1992; Bronfenbrenner & Morris, 2006), and an illustration of these propositions is interwoven throughout the empirical, theoretical, and methodological argumentation.

Background

There are extensive literatures on gender differences, cultural differences, (family and neighborhood) context effects, and socioeconomic status with regard to children's development. Given the scope of the literatures, the focus is on reviewing a small number of exemplary studies that have simultaneously examined gender, language or culture group status, and socioeconomic status factors with respect to young children's developmental outcomes for representative samples. Apart from the empirical findings, the theoretical frameworks that have been presented as explanations for the studies' findings are also reviewed.

It has been widely documented that developmental trajectories in a number of domains are substantially affected by what happens during the first years of life (Keating & Hertzman, 1999; McCain & Mustard, 1999). Also, it is widely recognized that children's transition from preschool to Kindergarten to school is a developmentally critical period with regard to children's social, emotional, and cognitive adjustment to school (Alexander & Entwisle, 1988).

Apart from a wide interest in understanding children's transition to school in general, there has also been an interest in whether developmental patterns during the transition are systematically different for girls and boys, for children from different socioeconomic backgrounds, and for different cultural groups (Entwisle, Alexander, & Olson, 2007; Rauh, Lamb Parker, Garfinkel, Perry, & Andrews, 2003). In numerous studies, the primary research questions have been asked with respect to the main effects. For example, "Do girls do better than boys with respect to early academic achievement?"

(see Buchman, DiPrete, & McDaniel, 2008, for a review), or “Do children from disadvantaged socioeconomic backgrounds fare worse with respect to health outcomes and school adjustment than children from high socioeconomic backgrounds?” (e.g., Kohen, Hertzman, & Brooks-Gunn, 1998²⁸), or “Are there differences between different ethno-cultural groups with respect to their developmental outcomes in the early years?” (see, for example, a recent special issue on ‘racial and ethnic gaps’ in school readiness; Rouse, Brooks-Gunn, & McLanahan, 2005).

One of the main findings, and one that has been fairly consistent, is that children from disadvantaged socioeconomic and ethno-culturally segregated backgrounds (i.e., poor minorities) tend to do worse on a range of developmental outcomes (e.g., health, school achievement) than children from high socioeconomic, and cultural mainstream backgrounds (e.g., Berliner, 2005; Jencks & Mayer, 1990; Wilson, 1987).

However, studies that examined not only main effects, but also interactions between gender, cultural group, and socioeconomic status factors have found results that are of great theoretical interest and that may have important implications for educational practices. In the following, a small number of particularly relevant studies are summarized in order to illustrate examples for each of the following: the interaction between gender and socioeconomic status, the interaction between immigrant/ESL and socioeconomic status, and the interaction between ethno-cultural background and socioeconomic status, with regard to children’s developmental outcomes.

²⁸ Kohen, Brooks-Gunn, Leventhal, and Hertzman (2002) examined the interactions between neighborhood and family level socioeconomic status; however, the study did not examine interactions between socioeconomic status and demographic child or family characteristics, such as gender, or ethno-cultural background.

Interaction between gender and socioeconomic status. Drawing from a city-wide panel study—the Beginning School Study from Baltimore—Entwisle et al. (2007) report on gender and socioeconomic background interactions with regard to reading outcomes in the first years of school. In summary, the findings of the study indicate that boys’ and girls’ reading scores are at similar levels at grade 1, and that a gender gap (in favor of girls) emerges until grade 5, which is driven by significant differences between the poorest boys and girls—whereas boys and girls at the high socioeconomic spectrum still show similar reading skills. (Of methodological interest is that socioeconomic disadvantage was defined by meal subsidy status, and thus represented by a dichotomous variable; i.e., yes, no). The findings from this study deserve particular mention, because the longitudinal analysis controls for parental expectations, children behavior, and for the gender specific patterns of retention.

As stated by Entwisle et al. (2007), the finding that there is no gender difference at the grade 1 level does not correspond to nationally representative studies, as they, in contrast, indicate that literacy skills of girls are already higher than those of boys in Kindergarten (see Buchmann et al., 2008). At the same time, the study by Entwisle et al. appears to be the only study that explored the interaction between gender and socioeconomic status with respect to literacy during the first years of school. Also, the study takes a rigorous empirical and theoretical approach in interpreting the emerging gender gap by referring to gender- and socioeconomic specific parental expectations with regard to school, as well as to teacher bias rooted in gender-specific behavioral differences. What is particularly noteworthy is that Entwisle et al. illustrate how a number of child and context variables and repeatedly occurring processes (e.g., gender-

specific treatment of children by teachers and parents), over time, interact and jointly contribute to the emerging gender gap at the low socioeconomic spectrum.

Interaction between immigrant/ESL and socioeconomic status. Willms (2003) analyzed data from the *Understanding the Early Years* (Nosbush, 2006) project conducted in a number of Canadian communities. In his multilevel analyses, he found that across all children, children from communities with high socioeconomic status are faring better on a vocabulary test than children from low socioeconomic communities. As mentioned above, this socioeconomic gradient has been found in a number of contexts, and with regard to a wide range of developmental outcomes in health and education. However, it has also been found that such socioeconomic gradients can substantially differ in their steepness (cf. Willms, 1999).

In this context, it is of interest that children from a recent immigrant background—a characteristic that was used as an indicator of lower social status—not only obtain lower vocabulary scores than their non-immigrant peers living in the same communities, but that immigrant children also had a steeper socioeconomic gradient than non-immigrant children. In other words, the difference between the two groups was more pronounced at the low socioeconomic spectrum than at the high socioeconomic spectrum. It must be noted that the interaction effect did not reach statistical significance in the study, and that socioeconomic status was operationalized (i.e., statistically modeled) as a community-level variable, as family level socioeconomic status data were not available.

These limitations notwithstanding, the importance of the findings is that Willms' study is one of the few attempts to identify interaction effects between socioeconomic

status and ethno-cultural group status with respect to children's developmental outcomes at an early age. Even more important is that the examinations of the interactions are clearly derived from theoretical hypotheses. In fact, in the paper, Willms delineates a number of different theories on socioeconomic gradients and community differences with regard to child development, and translates their assumptions into concrete empirical questions and the respective statistical multilevel models. In this particular case, Willms proposes the theory of *double jeopardy*, which states that the effect of lower social status and the disadvantageous effect of low socioeconomic status have not only additive, but also an additional multiplicative (negative) effect on the developmental outcome.

Interaction between ethno-cultural background and socioeconomic status. In their review paper on the educational achievement of Asian-American students, Sue and Okazaki (1990) analyze empirical studies and discuss theoretical explanations with regard to the finding that Asian-American students are often highly successful in school, especially in mathematics. The main argument of the paper is that this effect cannot be explained by theories that suggest a single main effect, such as those that solely invoke the notion of genetic differences or cultural differences as the causal factor behind these group differences. Rather, Sue and Okazaki propose that the academic success of Asian-American students is the result of an interaction between cultural values and the particular current conditions in a given societal context.

The significance of the paper is that it takes a rigorous empirically based and theoretical approach to explain findings pertaining to the same general phenomena, namely that specific cultural (minority) groups have repeatedly had different

developmental outcomes and/or trajectories than the majority group (see Sue & Okazaki, 1990, for details and examples). Of particular theoretical relevance is that these differences in developmental outcomes and trajectories are, however, not consistent, but differ from context to context, and from (historical) time to time. Accordingly, rather than theoretically attributing differences between cultural groups to general main effect differences between those groups, Sue and Okazaki's so-called theory of *relative functionalism* thus concludes that different development patterns need to be explained in light of interactions between group and context characteristics.

Theoretical considerations from the bioecological theory of human development

The studies reviewed above all have in common that they empirically and theoretically pursued an approach that highlights how the interaction of child, process, and context variables, over time, exert effects on children's developmental outcomes. This approach is congruent with one of the main propositions of Bronfenbrenner's bioecological theory of human development (Bronfenbrenner & Morris, 2006). Specifically, Bronfenbrenner's states that in "ecological research, the principal main effects are likely to be interactions" (p. 802). Bronfenbrenner's theory elaborately illustrates why this is, in fact, a generally applicable principle in human development research, and proposes specific theoretical and methodological guidelines for how to conduct research in line with this principle.

In Bronfenbrenner's view, arguably the most appropriate type of research model is the so-called process-person-context-time model. The basic notion of this model is that for any developmental outcome, it is critical to simultaneously examine the

developmentally most relevant process variables, person characteristics, and context characteristics, over time, in order to examine how these multiple factors interact with regard to the developmental outcome under question.

Bronfenbrenner's framework, which obtained prominence in the 1970s (Bronfenbrenner, 1979), and which has been substantially refined and elaborated since (see Bronfenbrenner, 1992; Bronfenbrenner & Morris, 2006), is frequently cited in the human development literature, particularly in studies that examine human development in its ecological context (e.g., Brooks-Gunn, Duncan, Klebanov, & Sealand, 1993; Leventhal & Brooks-Gunn, 2000) as opposed to, for example, the experimental laboratory. Interestingly, however, even though these studies thus acknowledge that person and context variables (at different levels of the ecology) collectively affect developmental outcomes, they do not necessarily examine theoretically interesting interactions, even if the data would allow them to do so (cf. Bronfenbrenner & Morris, 2006, for an example).

In light of these considerations, the next section provides a description of the Early Development Instrument project of the Human Early Learning Partnership, based on which the analyses presented in the following were conducted. First, the overall purpose of the project is illustrated, then the specific research questions are presented, and finally, it is highlighted in what way the approach of the study has been guided by Bronfenbrenner's bioecological theory of human development.

The current study: A description of the EDI project

The Human Early Learning Partnership at the University of British Columbia is an interdisciplinary research consortium that brings together researchers from numerous disciplines in order to study child development. The overall mission of this consortium is “to create new knowledge to help children thrive” (see www.earlylearning.ubc.ca). One of the major initiatives with regard to collecting data on children’s development is the EDI project (see Guhn, Janus, & Hertzman, 2007; Kershaw, Irwin, Trafford, & Hertzman, 2005; Lloyd & Hertzman, 2009). For this project, HELP has collected province-wide child development data since 1999, in 3-year cycles. To date, one pilot cycle (1999 to 2001), and two full province-wide cycles (2002-2004, and 2005-2007) have been completed.

Currently, the data on Kindergarten children’s development is linked to academic achievement data, at an individual level, from grades 4 and 7, and preparations to individually (and anonymously) link the data to children’s health files as well as to their families’ socioeconomic background data are under way. Eventually, the goal is to (anonymously) track trajectories of child development from the early years throughout school, in a number of developmental domains. The purpose of monitoring children’s development at the population level is to analyze, interpret, and communicate the data to stakeholders in the community, who can use the data to inform and complement their respective discussions, decisions, practices, and policies to support children, families, and their communities and to increase children’s well-being.

In their overall approach, the Human Early Learning Partnership very explicitly and exclusively pursues a community-oriented approach. This approach involves close

collaborations with community stakeholders, schools, child and family service organizations, policy makers, and so forth. In line with the community-oriented approach is the dissemination strategy of the child development data, which represents a major component of the collective work. That is, all child development data are solely communicated to the public at an aggregate neighborhood level. To this end, the Human Early Learning Partnership has engaged with groups of community residents across the province, and has, via focus groups, defined neighborhood boundaries that reflect the collective lived experiences of what geographical areas constitute meaningful entities (i.e., neighborhoods) in the views of the residents. This approach is therefore distinctly different from approaches that define neighborhood boundaries according to administrative and legislative boundaries (as defined, for example, by census areas).

Previous findings from the EDI project

Three consistently reported findings are that (i) boys, on average, obtain significantly lower EDI ratings from teachers than girls on all five developmental domains (Guhn, Gadermann, & Zumbo, 2007; Janus & Offord, 2007; LaPointe, Ford, & Zumbo, 2007), that (ii) children designated with an ESL status receive, on average, lower teacher ratings on all of the EDI's five developmental domains (Guhn et al., 2007; Janus & Offord, 2007; Kershaw et al., 2005), and that (iii) children from neighborhoods with low average socioeconomic status receive, on average, lower EDI scores on all of the five developmental domains than children from neighborhoods with high average socioeconomic status (Kershaw et al., 2005; LaPointe et al., 2007).

Deriving the hypotheses

In the bioecological theory of human development (Bronfenbrenner & Morris, 2006), Bronfenbrenner delineates what he coined *science in discovery mode*. In essence, science in discovery mode is a concept that refers to the notion that research on human development, in the absence of theories that can capture human development in its complexity, must pursue an approach that very cogently integrates empirical examinations with theoretical and methodological considerations. This is particularly necessary, because human development is influenced by numerous interactive factors, at different levels of the ecology, which exert their effect over long periods of time, and in ways that may differ from one developmental outcome to the next. Given this interactive nature of the influential forces on development, Bronfenbrenner repeatedly emphasizes the importance of explicitly examining interactions among developmentally relevant factors. In addition, Bronfenbrenner urges that the interactions to be examined are not only derived from theory, but should, in fact, be aimed at refining and expanding theory.

When one applies this general proposition—that is, of focusing on interactions—to the EDI research project, it becomes critical to determine which interaction effects are of particular theoretical and practical interest with regard to children's development. As the literature review suggested, one emerging key question is to what extent one can generalize findings about the association between socioeconomic status and children's development across groups and across contexts. Low socioeconomic status, particularly in its extreme form (i.e., poverty), has repeatedly been found to be associated with undesirable developmental outcomes among children (cf., Berliner, 2005; Jencks & Mayer, 1990; Wilson, 1987). However, the research literature also suggests that the

association between socioeconomic status and developmental outcomes cannot be attributed to a direct, causal relationship, but must rather be attributed to the fact that socioeconomic status is positively correlated with developmentally relevant factors (e.g., parenting style, nutrition, peer interaction, role expectations, access to resources). In fact, a number of studies corroborate this notion, as they suggest that policy interventions that simply raise families' socioeconomic status (e.g., by providing financial resources; e.g., Leventhal, Fauth, & Brooks-Gunn, 2005) have no or much less of a beneficial effect on children's development in comparison to interventions that systematically target developmentally beneficial processes directly (e.g., by providing comprehensive early childhood programs; e.g., Schweinhart & Weikart, 1997). Thus, there are, in principle, three general explanations for the correlations between socioeconomic status and developmental outcomes: First, socioeconomic status may be correlated with factors that are (causally) associated with developmental outcomes, and may, according to the size of the correlation, be considered a proxy measure for those developmentally relevant factors. The accuracy and validity of using socioeconomic status as a proxy for developmentally relevant processes is thus affected by measurement error, differences in operationalization of the socioeconomic status construct, and the developmental outcome under question. Second, socioeconomic status, or its developmental correlates, may interact with other variables with regard to developmental outcomes. According to Bronfenbrenner (1979), for example, there are some indications that boys are "being more affected by environmental contrasts than girls" (p. 181), and it might similarly be the case that girls and boys have different sensitivities with regard to differences in

socioeconomic status (or the associated developmentally relevant factors). Third, it might be the case that a combination of both phenomena is at play.

This argument is in line with the propositions of the bioecological theory of human development. In the theory, the developmentally relevant processes are coined *proximal processes*, and are defined as the repeatedly occurring and increasingly complex interactions between a person and its environment that are of particular significance for a particular developmental outcome. For example, parenting patterns (e.g., supportive versus unsupportive parenting) occurring between children and their parents or caretakers over time have significant effects on children's social and emotional competences. In general, such proximal processes are considered the 'engines' of development. A research study that omits proximal processes that are particularly relevant for the developmental outcome under question must therefore be especially cognizant of the possible misinterpretations of its findings. This point is critical in this context, as the data that are used for the analyses presented here do not contain any direct measures of any proximal processes, such as parenting or teaching.

The absence of proximal process measures notwithstanding, the theories and empirical findings reviewed above allow one to formulate specific hypotheses with regard to the data set and research design. As mentioned above, previous analyses of the EDI data show that girls do better than boys on all five domains of the EDI. Based on the findings by Entwisle et al. (2007) that the 'handicap of being poor and male' is associated with an evolving gender gap at the low end of the socioeconomic spectrum over the first few years of school—which corresponds to the theoretical notion of a double jeopardy of being male and poor—the hypothesis was that there is an interaction effect between

gender and socioeconomic status, in addition to the gender and socioeconomic status main effects. The interaction was theorized to be such that the gap between girls and boys is more pronounced at the low socioeconomic spectrum than at the high socioeconomic spectrum. The rationale for the hypothesis is that the two factors associated with lower EDI scores (male gender and low socioeconomic status) are assumed to be multiplicative (as well as additive). Entwisle et al. did not find a significant interaction at the Kindergarten age with regard to reading skills—but report that a trend was present. Entwisle et al. argue that gender-specific teacher expectations, over time, lead to the increasing gender gap at the low socioeconomic spectrum.

Given that the EDI scores are obtained via teacher ratings, and given that the ratings are obtained after teachers have gotten to know the children for about half a year (the ratings are performed in the middle of the Kindergarten year), it was hypothesized that the ‘handicap of being poor and male’ may already be reflected in the EDI scores.

With regard to children’s first language status and their language group belonging, the hypothesis was similar. As said, ESL children obtain, on average, lower teacher ratings on all five developmental domains assessed by the EDI. As found by Willms (2003), and, again, as proposed by the theory of double jeopardy, the hypothesis is that there is an interaction between ESL status and socioeconomic background. This interaction is, likewise, hypothesized to be such that the gap between ESL and non-ESL children is particularly pronounced at the low end of the socioeconomic spectrum.

In addition, based on the theory of relative functionalism (Sue & Okazaki, 1990), it was hypothesized that the main effect and interaction effects show different patterns for the different language groups (English, Punjabi, and Cantonese). Parallel to findings from

the US context about academic achievement of Asian Americans, it was hypothesized that the Cantonese group obtains relatively high scores on EDI domains that tap basic academic competencies, but relatively low ones on EDI domains related to (English) communication skills. In other words, it was proposed that there is an interaction between language background and developmental outcomes. For the Punjabi group, there is anecdotal evidence and preliminary empirical indications that their EDI scores are relatively lower than those of the English and Cantonese groups. However, there seems to be a lack of representative empirical or theoretical work in this area of research for the Canadian context. Therefore, the hypothesis regarding the Punjabi group remains at this vague level, and the pertaining analyses must thus be considered exploratory.

Finally, it was hypothesized that the interaction between language group and developmental outcomes has no effect on the language group by socioeconomic status interaction. In other words, it is expected that the socioeconomic gradient for ESL children is always steeper than for non-ESL children, for all developmental domains. That is, even if Cantonese children obtain higher scores on domains tapping basic academic skills, but lower scores on the communication domain than the English group, their gradients along the socioeconomic spectrum are, in both cases, hypothesized to be steeper than the gradients of the English group. Likewise, it was hypothesized that the socioeconomic gradients of the Punjabi group are consistently steeper than those of the English group.

Method

Sample

The sample consists of 40,772 Kindergarten children from 1384 schools across the entire province of British Columbia, Canada. Kindergarten in British Columbia is obligatory for all children as part of the K-12 school system, and takes place in the elementary schools (Kindergarten to grade 7). Children typically enter Kindergarten in September of the calendar year in which they become 5 years old. Data on the EDI are obtained via teacher ratings in February. For logistical reasons, data were collected over a three-year period (2002 to 2004); however, each school in the province was only sampled once during the period. The mean age of children at the time of data collection was 5 years, 8 months.

Gender. In the sample, 45% were girls, and 47% boys. (For 8%, gender information is missing.)

ESL status. About 16% of the children are designated with an ESL status. (For 2%, data on the ESL variable was missing.) ESL is solely conferred to children who are considered being in need of English language support. Therefore, not all children who learned a first language different than English are considered ESL, and a few native English speakers are designated ESL, if their English communication skills are low, because it allows schools to dedicate extra resources to support these children. It must be noted that the variable for ESL status is dichotomous (ESL versus non-ESL), even though it reflects an entire continuum (or context-dependent continua) of English communication skills. In addition, the variable might be affected by inconsistencies in the way teachers report ESL status on the EDI.

Language groups. For the language group analyses, three groups were chosen: children speaking English as their first language (and who are not considered ESL; $n = 32,194$), and the two largest language groups among the children designated as ESL, Punjabi ($n = 1,765$), and Cantonese²⁹ ($n = 1,293$). (There are a few hundred children, who speak Punjabi or Cantonese as their first language, but who are not designated as ESL; these children were therefore not included in the ESL/language group comparisons.)

Measures

The Early Development Instrument. Data on children's developmental status in Kindergarten, as well as their gender, ESL status, and first language were obtained via teacher ratings on the Early Development Instrument (EDI; Janus & Offord, 2000; 2007; Janus et al., 2007). The EDI is a rating scale, which is administered to all children in the classroom by Kindergarten teachers in February of the Kindergarten year (which starts in September). Teachers rate children's developmental status as reflected in their school readiness on five developmental domains: physical health and well-being (13 items), social competence (26 items), emotional maturity (28 items), language and cognitive development (26 items), communication skills and general knowledge (8 items). All items have 2- or 3-point Likert scale response formats (i.e., *yes, no, don't know; very true, sometimes or somewhat true, never or not true, don't know*). All items are recoded into scores ranging from 0-10 ('don't know' responses are not scored), and then domain scores are calculated as the average score. (High scores always reflect positive

²⁹ The overall sample also included 604 children with Mandarin as their first language, as well as several hundred children who speak Hindi, Urdu, or other languages spoken in India. It should be emphasized that the intent was to compare patterns of results according to the language background of the children, and not according to their (or their parents') countries of origin (e.g., China, India).

outcomes.) Accordingly, each child receives a score from 0-10 for each of the five domains.

Socioeconomic status. Socioeconomic status (SES) data were obtained from 2001 tax filer data (Statistics Canada). In the analyses, the variable ‘median equivalized disposable income’ at the 6-digit Canadian postal code level is used as an approximation of a child’s family’s socioeconomic status³⁰. The variable represents the disposable income per person within a given postal code (cf. Ebert, 1999). It is calculated from family level income information, but due to privacy laws, Statistics Canada provides the data at an aggregated postal code level. The size of a postal code ranges from a block in densely populated cities to larger areas in sparsely populated rural areas, and typically consists of 200 to 300 households. A representative Canadian study (Mustard, Derksen, Berthelot, & Wolfson, 1999) has shown that the 6-digit postal code income information serves as a reasonable proxy for family level income, as the study compared statistical analyses that used (i) the actual family level income or (ii) the aggregated, 6-digit postal code average income as predictor for the same outcome variables, and found comparable (patterns of) results. Based on this finding, previous population-level studies in the Canadian context have employed the same methodology (e.g., Oliver, Dunn, Kohen, & Hertzman, 2007). In the multi-level analyses in this study, the ‘median equivalized disposable income’ at the 6-digit Canadian postal code level is therefore used as the level 1 (child-level) proxy for the socioeconomic status of a child’s family, according to the

³⁰ In the literature, SES is commonly measured by parental income, education, or occupational status (or any combination thereof). Please see the discussion for some pertaining theoretical comments.

postal code of the child's residence at the time of data collection, which is recorded in the EDI data files.

Analysis

The relationships between children's gender, their ESL and language group status, and their socioeconomic background with respect to the developmental outcomes measured by the EDI were assessed via multilevel modeling techniques, using SPSS software (MIXED; version 15.0). A multilevel approach was chosen, because of the nested data structure (that is, children in the sample are grouped in schools and rated by their Kindergarten teachers). The variable 'school id' was used as a grouping variable in this analysis ($n_{\text{School}} = 1232$). (In the majority of elementary schools included in the sample, there are one (53%) or two (34%) Kindergarten teachers.)

*Gender*socioeconomic background interactions.* Due to the different distributions of children from different language groups across the socioeconomic spectrum³¹, only children from the largest language group (English) were chosen for this analysis. In order to analyze the interaction effects between children's gender and their socioeconomic background, two multilevel models were fitted. These two models were run for each of the EDI domains, with the children's domain scores as the dependent (outcome) variable.

First, an unconditional model was fitted, with a random intercept, but with no predictors at either the individual child level (Level 1) or the school level (Level 2). This

³¹ The mean income for the ESL children's families (about 19,500 Canadian dollar) is almost one standard deviation lower than the mean income for the English language group children's families (about 24,000 Canadian dollar).

analysis made it possible to calculate the intraclass correlations, in order to test whether schools significantly varied with respect to their children's EDI scores.

Second, a conditional model with the main effect predictors of gender and socioeconomic background as well as the interaction effect term gender*socioeconomic background (Level 1) was fitted. The gender variable was included as a fixed effect, and socioeconomic status as random effect. Both the gender and socioeconomic status variables were centered, so that 0 represented girls and 1 represented boys for the gender variable, and so that 0 represented the overall mean for the socioeconomic status (i.e., income) variable. The full, conditional final model was compared to the unconditional model via a χ^2 -test in order to see whether the inclusion of the predictors significantly improved model fit.

*Language group*socioeconomic background interactions.* For this analysis, two sets of models were run, each using the English language group as the reference group. That is, the English group was compared to the Punjabi group and to the Cantonese group, respectively. As for the gender*socioeconomic analysis, two multilevel models were run. First, an unconditional model, with a random intercept, but no predictors at Level 1 or 2. Second, a conditional model, including the children's language group status (centered, so that 0 represented the English language group, and 1 the Punjabi and Cantonese groups, respectively) and socioeconomic background (also centered, as above) as main effects, and the interaction effect term language group*socioeconomic status. Like the gender variable, the language group variable was modeled as fixed effect, and socioeconomic status again as random effect. As above, EDI domain scores were used as

dependent variables. Table 4.1 reports descriptive statistics for the dependent variables (i.e., the five domains of the EDI), and for the Level 1 predictor variable ‘disposable income/socioeconomic status’ for the respective groups that are analyzed (girls versus boys; non-ESL children speaking English as their first language, and ESL children speaking Punjabi- or Cantonese as their first language).

Table 4.1 *Descriptive statistics (means, standard deviations) of the EDI scores and socioeconomic status, for girls and boys of the non-ESL group, and the ESL (Punjabi-, Cantonese-speaking) and non-ESL groups*

Outcome	Girls (Non-ESL; n=14,487)	Boys (Non-ESL; n=15,131)	English- speaking children (n=32,194)	Punjabi- speaking children (n=1,765)	Cantonese- speaking children (n=1,293)
Physical health & wellbeing	9.1 (1.3)	8.7 (1.5)	8.9 (1.5)	8.9 (1.4)	9.0 (1.4)
Social competence	8.7 (1.6)	8.0 (2.0)	8.3 (1.8)	7.9 (1.9)	8.1 (1.9)
Emotional maturity	8.4 (1.3)	7.6 (1.7)	8.0 (1.6)	7.9 (1.5)	7.8 (1.5)
Language & cognitive development	8.6 (1.7)	8.0 (2.0)	8.2 (1.9)	7.2 (2.3)	8.3 (1.8)
Communication skills & general knowledge	8.7 (2.0)	8.1 (2.4)	8.4 (2.3)	5.8 (2.9)	5.8 (2.8)
Predictor (Level 1)					
SES/	24,100	24,200	24,300	18,400	19,800
Disposable income ¹	(7,400)	(7,400)	(7,400)	(4,800)	(5,900)

¹ The variable SES (socioeconomic status)/disposable income is derived from census data on median equivalized disposable income aggregated at the 6-digit postal code level (see text for details).

Results

*Gender*socioeconomic background interactions*

According to the unconditional models that were fitted for each of the five developmental domains of the EDI, the intraclass correlations ranged from .07 to .12. Following recommendations in the literature (Cohen, Cohen, West, & Aiken, 2003), intraclass correlations of this size justify a multilevel approach. (Ordinary least square multiple regressions assume an intraclass correlation of zero.)

Data screening and assumption checking. EDI scores on all five domains are highly negatively skewed (skewness ranges from -1.0 to -1.7), as most children receive very high (i.e., favorable) ratings (average scores range from 8.0 to 8.9, on a scale from 0 to 10, for the five domains). Skewness typically lowers statistical power, but due to the large sample size, lack of power is not a concern.

A scatterplot of the relationship between socioeconomic status and EDI scores clearly indicated heteroscedasticity. That is, EDI scores were spread out across the entire score range at the low socioeconomic spectrum, but were continuously restricted to a narrower and higher score range towards the high end of the socioeconomic spectrum³². Attempts to alleviate this violation of assumptions via transformations were conducted according to the standard recommendations in the literature (Cohen et al., 2003). However, due to the fact that the EDI score distribution is truncated at the maximum score (i.e., 10)—meaning that a high proportion of children receive a score of 10—the transformations did not yield satisfactory results. Therefore, it was decided to proceed with the multilevel analysis with the original data, and correct for the heteroscedasticity

³² The heteroscedasticity can be seen in Figure 4.7.

according to a procedure that is described below. First, however, the results from the multilevel analyses are presented in Table 4.2.

Table 4.2
Summary of multilevel modeling results (coefficients, standard errors, and significant levels) for the gender and language group analyses

	Physical health & wellbeing	Social competence	Emotional maturity	Language & cognitive development	Communication skills & general knowledge
Gender ¹	-.40 (.02)***	-.80 (.02)***	-.81 (.02)***	-.59 (.02)***	-.68 (.03)***
SES ²	.31 (.03)***	.24 (.02)***	.17 (.02)***	.27 (.02)***	.35 (.03)***
Gender X SES ³	.03 (.02)	.08 (.03)**	.05 (.02)*	.10 (.03)***	.09 (.03)*

¹ The gender variable is coded 0 for girls and 1 for boys.

² The coefficients and standard errors for SES (socioeconomic status) and the gender X SES interaction pertain to Can\$10,000 increments.

³ Interactions effects in the gender analyses that were statistically significant in the multilevel analyses were not statistically significant in follow-up analyses that controlled for the heteroscedasticity of the data (see text for details).

* $p < .05$; ** $p < .01$; *** $p < .001$

In all analyses, the inclusion of the main effect and interaction terms significantly improved model fit in comparison to the unconditional model (with all χ^2 values, with $df = 3$, being statistically significant at $p < .0001$).

According to recommendations provided in Hox (1995), the results for the social competence domain³³ are visualized in a graph that plots the EDI scores as predicted by the multilevel model for girls and boys (Figure 4.1).

³³ The results are similar for the other four EDI domains.

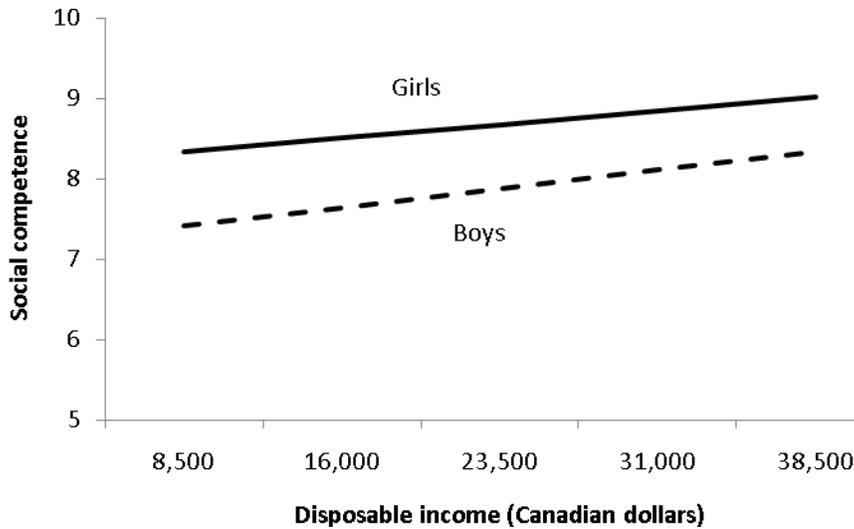


Figure 4.1 Gender by socioeconomic status interactions (Social competence domain)

As can be seen, the regression line representing the girls is higher than that of the boys, across the entire socioeconomic spectrum. Also, the graph suggests that the gap between girls and boys widens towards the lower end of the socioeconomic spectrum—a finding that is reflected in the significant interaction term. However, as indicated above, this finding cannot be interpreted in a straightforward manner, because the data are heteroscedastic. Therefore, a consequent analysis for each of the five developmental domains was conducted.

In this analysis, the effect size (Cohen's *d*) for the difference between girls and boys at every income quintile were calculated, using the pooled variance of the EDI scores for each income quintile, respectively. This was done based on the following rationale: The EDI scores have a greater variability at the low socioeconomic spectrum than at the high end (this is also seen in Figure 4.7). As the calculation of Cohen's *d* as an effect size estimator takes score variability into account (by means of dividing the group

difference by the pooled variance of the compared groups), separate effect size calculations for each income quintile, with the respective variances, made it possible to examine whether the significant interaction term resulted from the heteroscedasticity. That is, if the effect sizes are similar (i.e., not statistically significant different from each other) at every income quintile, the significant interaction term in the multilevel model may be considered spurious. Alternatively, if the effect sizes were consistently higher for the low socioeconomic quintiles than for the high ones, the multilevel results would be corroborated.

In order to test whether the effect sizes at the lowest income quintile were significantly larger than at the highest income quintile, the effect sizes (Cohen's d) were converted into correlations, then transformed into z -scores via the Fisher transformation, and the difference of z -scores was then tested for statistical significance (Cohen et al., 2003). This procedure produced the following results. For the physical health and well-being domain, the interaction term was not significant, and this finding was corroborated by the fact that the gender difference showed similar effect size magnitudes at all income quintiles (Cohen's d ranging unsystematically from .26 to .31). For the social competence domain, the effect size at the lowest income quintile was $d_{\text{lowSES}} = .50$, and at the highest income quintile $d_{\text{highSES}} = .45$. The difference between these two effect sizes was not significant ($p = .09$). Similarly, the difference in effect sizes for the other domains (emotional maturity: $d_{\text{lowSES}} = .56$, $d_{\text{highSES}} = .53$, $p = .22$; language and cognitive development: $d_{\text{lowSES}} = .34$, $d_{\text{highSES}} = .31$, $p = .19$; communication skills and general knowledge $d_{\text{lowSES}} = .33$, $d_{\text{highSES}} = .31$, $p = .33$) were also not statistically significant.

*Language group*socioeconomic background interactions*

According to the unconditional models that were fitted for each of the five developmental domains of the EDI, the intraclass correlations ranged from .09 to .12.

Table 4.3 presents the multilevel results for the analyses that compared the non-ESL children with English as their first language with (i) ESL children that speak Punjabi or (ii) Cantonese as their first language.

Table 4.3
Summary of multilevel modeling results (coefficients, standard errors, and significant levels) for the gender and language group analyses

	Physical health & wellbeing	Social competence	Emotional maturity	Language & cognitive development	Communication skills & general knowledge
Punjabi ¹	-.54 (.16)**	-.37 (.21)	-.46 (.18)**	-.11 (.20)	1.5 (.25)***
SES ²	.07 (.08)	-.04 (.10)	-.07 (.09)	-.16 (.11)	-.13 (.12)
Punjabi X SES	.22 (.08)**	.32 (.10)**	.27 (.09)**	.45 (.10)***	.52 (.12)***
Cantonese ⁴	-.33 (.15)*	-.07 (.20)	-.05 (.16)	-.48 (.21)*	2.5 (.24)***
SES	.21 (.08)**	.20 (.09)*	.11 (.09)	.16 (.10)	.38 (.11)***
Cantonese X SES	.03 (.07)	.08 (.09)	.04 (.08)	.17 (.10)	.00 (.11)

¹ The language group variable is coded 0 for Punjabi- and Cantonese-speaking children (in the respective analyses) and 1 for English-speaking children.

² The coefficients and standard errors for SES (socioeconomic status) and the gender X SES interaction pertain to Can\$10,000 increments.

* $p < .05$; ** $p < .01$; *** $p < .001$

With regard to model fit, the inclusion of the main effect and interaction terms significantly improved model fit in comparison to the unconditional model for all analyses (with the χ^2 , with $df = 3$, all being statistically significant at $p < .0001$).

The results are visualized in graphs that plot the EDI domain scores as predicted by the multilevel model for the three groups (English, Punjabi, and Cantonese; Figures 4.2 to 4.6). In these figures, the regression lines for the Punjabi and Cantonese groups are restricted to the CAN\$8,500 to 23,500 disposable income bracket, as most children fall into this income range, whereas the English group is distributed (normally) over a wider range of the income spectrum. This point is illustrated in the three graphs of Figure 4.7, in which the EDI scores of the respective three groups are plotted against the socioeconomic status variable (i.e., disposable income).

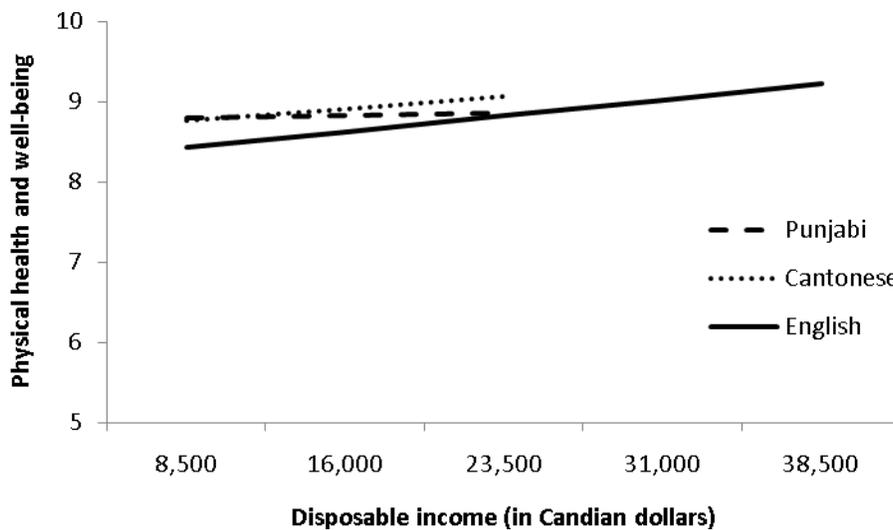


Figure 4.2 Language group by socioeconomic status interactions (Physical health and well-being domain)

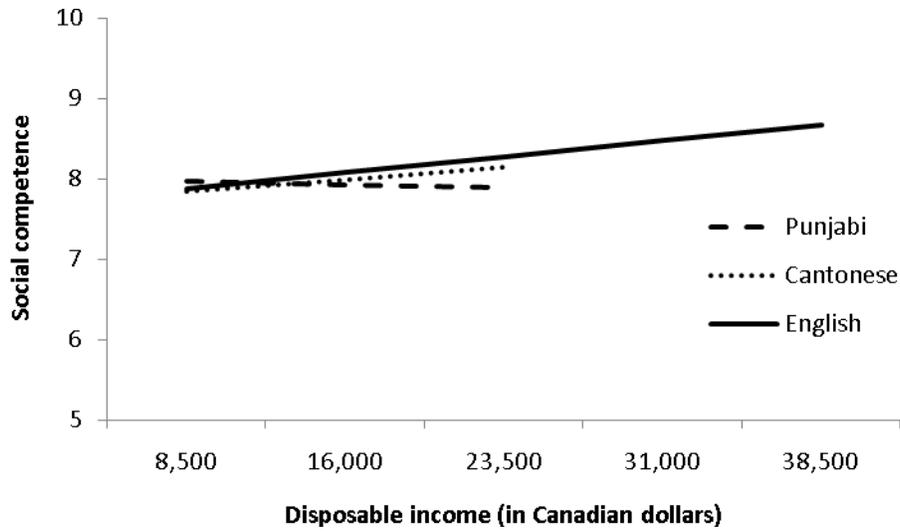


Figure 4.3 Language group by socioeconomic status interactions (Social competence domain)

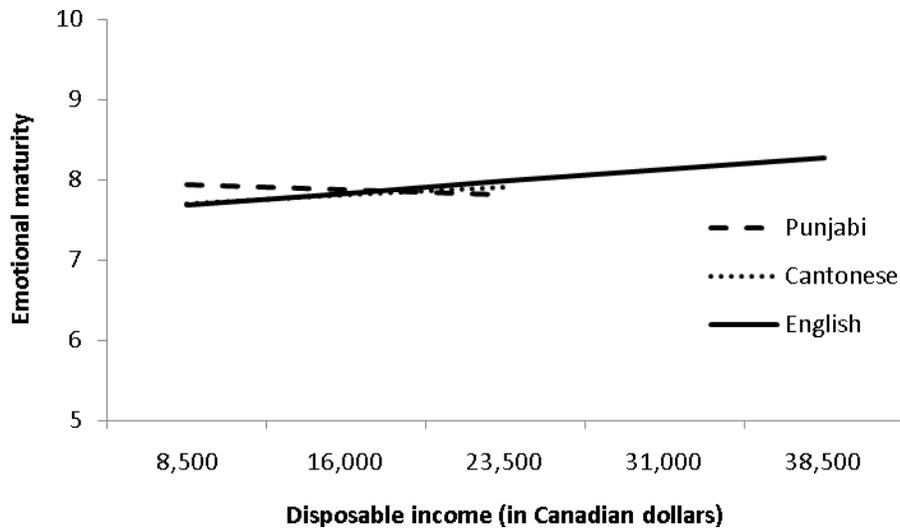


Figure 4.4 Language group by socioeconomic status interactions (Emotional maturity domain)

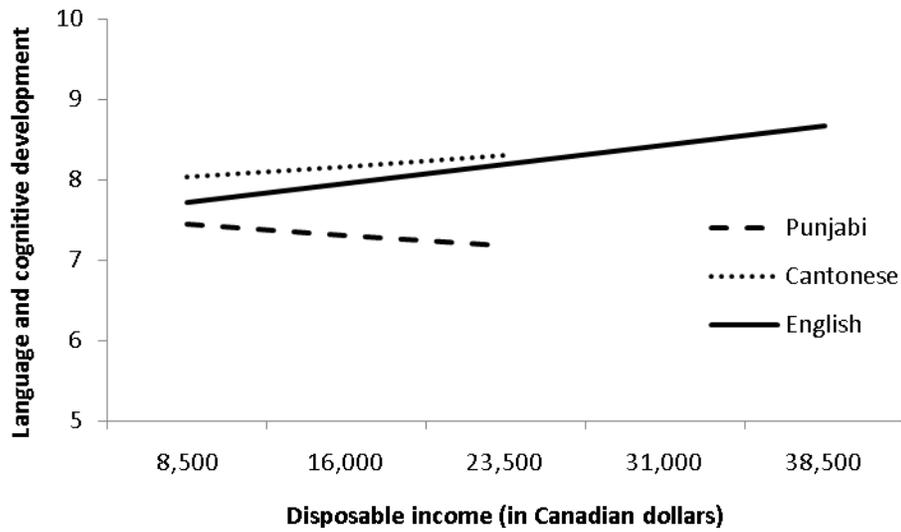


Figure 4.5 Language group by socioeconomic status interactions (Language and cognitive development domain)

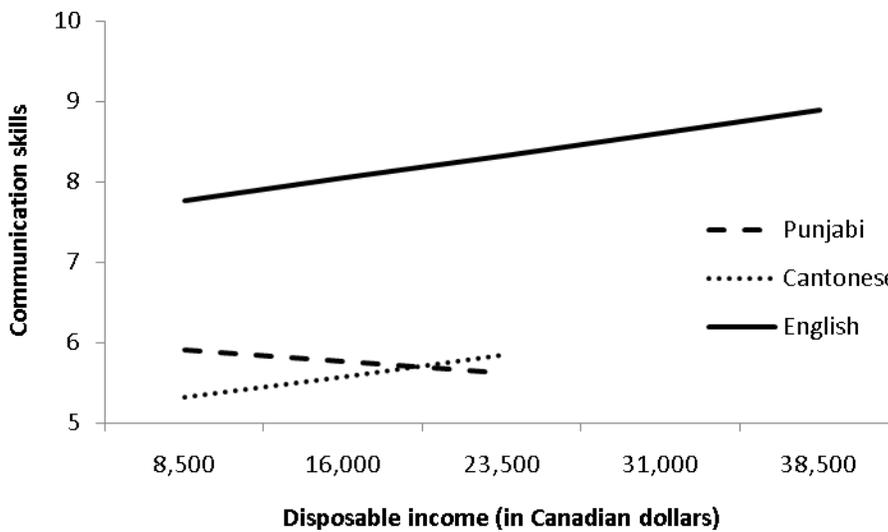
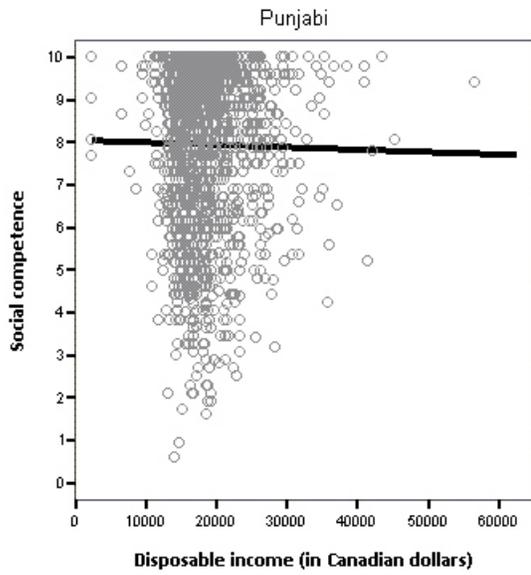
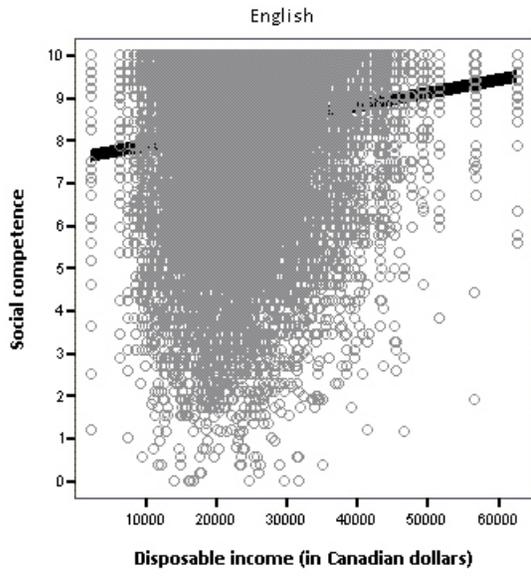


Figure 4.6 Language group by socioeconomic status interactions (Communication skills and general knowledge domain)



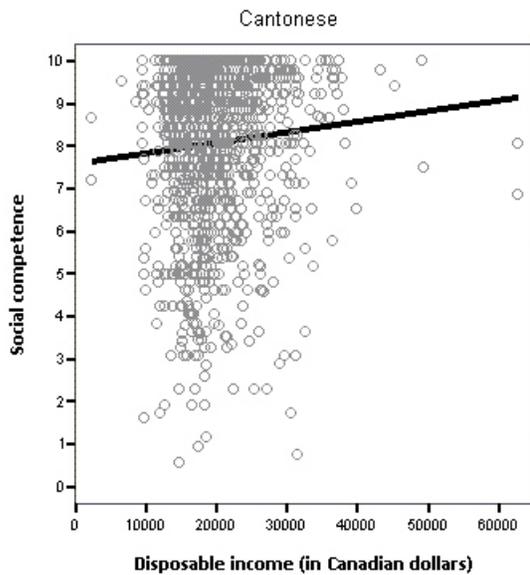


Figure 4.7 Scatterplots for the relationship between developmental outcome and socioeconomic status for the English, Punjabi, and Cantonese groups (Social competence domain)

Figures 4.2 to 4.6 show that the gradients for the three groups differ distinctly depending on the developmental domain under question. Particularly noteworthy is that the Punjabi and Cantonese children seem—according to the models—to be receiving similar or higher ratings than the English group at the low end of the socioeconomic spectrum in all domains except for the communication skills and general knowledge domain. Also, it can be seen that the effect of socioeconomic status is pronounced for the English and the Cantonese groups, whereas the Punjabi group shows a relatively flat socioeconomic gradient on all five domains. Finally, it is interesting to see that the Punjabi and Cantonese children receive approximately equal ratings on all domains, except for the language and cognitive development domain, where the Cantonese group receives the highest ratings.

Discussion

This section first discusses the gender analyses and then the language group analyses. Thereafter, the strengths and limitations of the study are discussed. At the end, recommendations are made for future analyses that examine the consistency of these findings across time and space, as well as for examinations that explore whether the gender effects are consistent for the different language groups.

The gender analyses showed a consistent, statistically significant difference between boys and girls on all five developmental domains (in favor of girls) across the socioeconomic spectrum. This finding is contrary to the hypothesis that ‘the handicap of being poor and male’ (cf. Entwisle et al., 2007) is (already) reflected in teacher ratings of developmental outcomes at a Kindergarten age. It should be noted that the empirical findings do show a trend that is in line with the hypothesis in that the difference between girls and boys is more pronounced at the lower socioeconomic spectrum, for the emotional maturity, language and cognitive development, and, most clearly, for the social competence domain—however, these interactions were not statistically significant in the analyses (after taking the heteroscedasticity of the data into account). The interpretation of this finding is complicated by a number of points. First of all, school readiness data, covering five developmental domains, have not been available at a population level for the Canadian context in previous research; therefore, the results cannot be compared to equally representative samples. Secondly, analyses of data from the representative National Longitudinal Study of Children and Youth (for the Canadian context) and large-scale early development and school readiness from the US (e.g., Early Childhood Longitudinal Study) have not examined interactions between gender and socioeconomic

status. It will therefore be of interest to see whether such analyses provide similar results to the ones reported in this study. Third, it must be noted that different studies have used different methodologies (e.g., tests versus, parent reports, teacher observations) to measure developmental outcomes, such as school readiness, and constructs such as socioeconomic status, which makes comparisons across studies more complicated. Fourth, gender differences have been examined with regard to different developmental outcomes (e.g., social competence, reading, math, grades, intelligence tests), at different age groups (preschool, K-12), in different contexts (e.g., inner city versus residential suburbs), and at different (historical) times (for a review on gender differences in education, see Buchmann et al., 2008). With regard to the purpose of this study, as well as the proposed *double jeopardy* hypothesis (Willms, 2003) regarding the ‘handicap of being poor and male’ (Entwisle et al., 2007), it is recommended that follow-up studies conduct longitudinal analyses, to examine whether gender differences remain constant or change in pattern over time, and whether the trajectories are moderated by socioeconomic status.

The language group analyses showed a number of interesting findings. The hypothesis that both ESL groups (Punjabi and Cantonese) show steeper socioeconomic gradients with regard to the five developmental outcomes assessed with the EDI was not supported. To the contrary, both groups showed flatter gradients, and for the Punjabi group, this difference in steepness of the regression line was significant for all five developmental domains. That is, for the Punjabi group, the socioeconomic gradients were almost completely flat. The theory of double jeopardy, according to which the language barrier and socioeconomic disadvantage have multiplicative disadvantageous effects on

their developmental outcomes (cf. Willms, 2003) is thus not supported by the data from this study. As mentioned with regard to the gender analyses, the interpretation of these findings is similarly complicated by a number of points. As mentioned, population-level data on children's school readiness have not previously been available for the Canadian context. With regard to studying the developmental outcomes of children from different language backgrounds, this is of particular relevance, because population-level data (or large-scale representative samples) are necessary to obtain sufficiently large sample sizes for individual language groups to warrant multilevel statistical analyses. As a result, the presented findings cannot be directly compared to previous research findings in this area. Finally, the study by Willms (2003), which provided evidence in support of the hypothesis of *double jeopardy*, used different developmental outcomes (scores on a vocabulary test) and neighborhood level socioeconomic status, rather than family level socioeconomic status for the analyses. It is therefore recommended that future research conducts studies with regard to a number of different developmental domains in longitudinal designs, in order to investigate in what way differences between language groups are related to age, or to context characteristics, in what way these are specific to particular outcomes, or to language-related cultural differences, and—finally—to what extent interactions of these factors are related to developmental patterns of the different language groups.

The previously reported finding that English-speaking children receive, on average, higher ratings than ESL children is, it seems, primarily associated with the differences in socioeconomic status between the groups. After all, in these analyses, which take language background and socioeconomic status into account simultaneously,

the main effects of language background were only significant in favor of the English-speaking group for the communication skills and general knowledge domain. On all four other developmental domains, the Punjabi- and Cantonese-speaking children receive similar or higher teacher ratings at comparable levels of socioeconomic status. Finally, the hypothesis that the Cantonese-speaking children would receive higher ratings in the domain that most closely resembles basic academic competences (i.e., the language and cognitive development domain) is supported by the empirical findings. The same pattern was not found for the Punjabi-speaking children. In general, this finding is in line with the theory of relative functionalism (cf. Sue & Okazaki, 1990).

This research presents a critically important step toward a better understanding of the interactive effects between children's gender, cultural background, and socioeconomic context. Future analyses of the (expanding) database will be able to build on the presented analyses, to explore higher order interactions, such as gender*language group*socioeconomic status interactions. Soon, such analyses can be conducted over time, as developmental outcome data at grade 4 and grade 7 for the same cohorts of children will become available. Eventually, it will thus be possible to implement a systematic analytical research design that comes closer to Bronfenbrenner's propositions outlined with respect to the so-called process-person-context-time model (Bronfenbrenner & Morris, 2006). In particular, such research would allow one to explore the interactions between *person* characteristics (gender, language background), *context* variables (socioeconomic status at the family level and at the neighborhood or aggregated group level), over *time* (Kindergarten to Grade 4, to Grade 7), on a number of

developmental outcomes (five developmental domains in Kindergarten, and academic achievement in reading and numeracy in Grades 4 and 7).

Such a design would still be missing any measures of proximal processes—which brings one back to a consideration that was briefly anticipated in the introduction of this paper. According to the bioecological theory of human development, proximal processes are the engines of development. In the EDI database, there are, as mentioned above, no direct measures of proximal processes (such as parenting or teaching). At the same time, a variable such as language group, which was used here, may turn out to be a good proxy indicator of developmentally relevant processes (for example, at a cultural macrosystem level, representing certain cultural norms, expectations, and habits that characterize regularly occurring processes).

The difficulty is, of course, to accurately interpret such proxy variables. As mentioned above, the findings show, for example, a flat socioeconomic gradient for the Punjabi group. A large body of literature that has examined socioeconomic status with regard to developmental outcomes has found that socioeconomic status is correlated with a wide range of developmental outcomes. However, this does not allow one to conclude that socioeconomic status has a direct, causal effect on developmental outcomes. After all, socioeconomic status is also correlated to a number of developmentally relevant processes, which, in turn, have a direct, causal effect on developmental outcomes (e.g., Berliner, 2005; Jencks & Mayer, 1990; Wilson, 1987). In the absence of data on developmentally relevant proximal processes, socioeconomic status will thus statistically explain variance in the outcome variable to the degree that socioeconomic status, correlates with (unmeasured) proximal processes, and to the degree that those processes

are related to the (measured) developmental outcomes (cf. Cohen et al., 2003). In other words, in the absence of data on processes, socioeconomic status can solely be interpreted as a proxy for proximal processes. More practically speaking, from a developmental perspective, it is more important *what to do* with socioeconomic status than to *have* (high) socioeconomic status. From this argument follows that if the correlation between socioeconomic status and developmentally relevant processes is different for different subgroups, such difference will be reflected in differences of the correlations between socioeconomic status and developmental outcome variables for those subgroups. That is, the reliability of using socioeconomic status as a proxy might be different for different samples or sub-samples. In the case of the Punjabi group in the British Columbia context, one might speculate that the developmentally critical social processes occurring within the Punjabi communities are not captured by an income variable in the same way they are captured for the English- and/or Cantonese-speaking speaking groups. Anecdotal evidence suggests that social processes among Punjabi-speaking families are more closely related to the degree to which there is a Punjabi-speaking social network in the community. In fact, such an effect was found by Rauh et al. (2003) for Hispanic children in New York: Among Hispanic children, those who lived in communities with a relatively high proportion of other Hispanic families had relatively better developmental outcomes than their peers in communities with a smaller proportion of Hispanic families. (Incidentally, this effect was more pronounced for boys than for girls.) For the Punjabi case in the British Columbia context, it may very well be that Punjabi communities with highly functional social networks are found in low socioeconomic status neighborhoods as often as, or even more frequently than in high

socioeconomic status neighborhoods. The data indicate that the distribution of Punjabi children is, on average, denser in a number of low socioeconomic status neighborhoods than it is in most high socioeconomic neighborhoods.

With regard to the language group analyses, it needs to be asked whether the developmental constructs that are measured (e.g., social competence) are universally valid. Also, it needs to be asked whether the developmental domains are valued in similar or different ways. Guhn, Gadermann, and Zumbo (2007) examined whether there is bias between ESL and non-ESL children, with a differential item functioning analysis. Such an analysis could also be conducted to compare individual language groups with each other. In addition, further research with families from the different language groups could explore whether the questions on the EDI comparably reflect the overall construct of 'school readiness' and the constructs represented by the five developmental domains of the EDI (e.g., social competence, emotional maturity) for the different language groups. Previous research (from the US context) has indicated that the construct of school readiness can substantially differ from one (geographic, cultural) community to the next (e.g., Graue, 1993), as well as from one (professional, disciplinary) community to the next (e.g., Scott-Little, Kagan, & Frelow, 2006). According to the notion that, first and foremost, school should be ready for all children (Graue, 1993), such knowledge of culturally differently constructed meanings of school readiness for the BC context might be invaluable for schools in order to respond to culture- and context-specific strengths and needs of children and their families.

The presented study contains a number of strengths and limitations. First of all, the study can draw from the first Canadian population-level database on children's school readiness, measured in five developmental domains, which makes it unique. This strength, however, inevitably came with some trade-offs. In comparison to other large-scale early childhood and school readiness studies, such as the National Longitudinal Study of Children and Youth, the current study solely used data from one measure (the EDI), from one method (teacher rating), and at one point in time (Kindergarten), whereas the NLSCY, and comparable studies, typically use multiple measures and methods (e.g., tests and observations), from multiple informants (e.g., parents and teachers), over multiple time points from a multi-year long period (e.g., from birth to high school). It is therefore recommended that future research make use of the differences between findings from the EDI project and other research projects by using them as a starting point for further validation research as well as a starting point for in depth examinations that explore the factors that can explain such differences.

Likewise, it needs to be reiterated that an interpretation of the results in this study is complicated by the fact that several of the variables used are proxy variables, and refer to broad constructs (e.g. socioeconomic status, school readiness). With regard to the construct of socioeconomic status, for example, using the 6-digit postal code median equivalized disposable income as a proxy for family income might have led to small distortions of the coefficients in the multilevel models; that is, potentially the relationship between socioeconomic status (i.e., the income variable) and the EDI outcome variables in the multilevel models is slightly attenuated (cf. Mustard et al., 1999). Similarly, even though ESL status is measured by dichotomous variables, it cannot be assumed that ESL

status should be conceptualized as a dichotomous construct. Therefore, it is particularly critical to be careful with regard to interpretation of the findings and with regard to any generalizations that require certain implicit or explicit assumptions hold.

Under the condition that one is being mindful of these limitations, the presented analyses contribute important empirical insights, based on a representative, population-level data base, to developmental questions regarding children's early developmental outcomes as reflected in their school readiness, as related to their gender, language background, and socioeconomic status. It is hoped that these findings will, eventually, in combination with further complementary analyses, contribute to a body of knowledge that can inform practitioners and policy makers—thus addressing a basic premise of the bioecological theory of human development. In fact, the major strength of the population health approach to assessing children's school readiness lies in the fact that it provides a big picture snapshot of all children's development, which can be used to ask and refine questions with regard to similarities or differences in the patterns of results for different groups of children.

Questions regarding gender differences, and differences between different cultural groups, and the (in)consistency of such differences across time and across contexts remains a topic of high importance in a society, which is increasingly characterized by changing gender roles and a high level of migration. Therefore, obtaining a clear grasp of how these factors are associated with developmental trajectories of children will contribute critical information to decision making pertaining to practices and policies that affect children and families in the communities in which they live. Toward this end, a sequence of analyses building on the presented findings that is theoretically and

methodologically guided by propositions of the bioecological theory of human development, with a focus on uncovering the proximal processes that underlie children's developmental trajectories, promises to be a fruitful route.

References

- Alexander, K. L., & Entwisle, D. R. (1988). Achievement in the first two years of school: Patterns and processes. *Monographs of the Society for Research in Child Development, 53*, 2.
- Berliner, D. C. (2005). Our impoverished view of educational reform. *Teachers College Record*. Retrieved on November 10, 2007, from <http://www.tcrecord.org/content.asp?contentid=12106>
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U. (1992). Ecological systems theory. In R. Vasta (Ed.), *Six theories of child development: revised formulations and current issues* (pp. 187-249). London, UK: Jessica Kingsley Publishing.
- Bronfenbrenner, U., & Morris, P. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Development, Volume 1, Theoretical Models of Human Development*, (6th ed.), 793-828. Hoboken, NJ: John Wiley & Sons.
- Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K., & Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology, 99*, 353–395.
- Buchman, C., DiPrete, T. A., McDaniel, A. (2008). Gender inequalities in education. *Annual Review of Sociology, 34*, 319-337.

- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ebert, U. (1999). Using equivalent income of equivalent adults to rank income distributions. *Social Choice and Welfare*, *16*, 233-258.
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *The Sociology of Education*, *80*, 114-138.
- Graue, M. E. (1993). *Ready for what? Constructing meanings of readiness for Kindergarten*. Albany, NY: State University of New York Press.
- Guhn, M., Janus, M., & Hertzman, C. (Eds.) (2007). The Early Development Instrument: Translating school readiness assessment into community actions and policy planning [Special Issue]. *Early Education & Development*, *18*, 369-570.
- Hox, J. J. (1995). *Applied multilevel analysis*. Amsterdam, NL: TT-Publikaties.
- Janus, M., Brinkman, S., Duku, E., Hertzman, C., Santos, R., Sayers, M., et al. (2007). *The Early Development Instrument: A population-based measure for communities. A handbook on development, properties, and use*. Hamilton, ON: Offord Centre for Child Studies. Retrieved on Nov 20, 2008, from http://www.offordcentre.com/readiness/pubs/2007_12_FINAL.EDI.HANDBOOK.pdf
- Janus, M., & Offord, D. R. (2000). Readiness to learn at school. *Canadian Journal of Policy Research*, *2*, 71-75.

- 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*, 1–22.
- Jencks, C., & Mayer, S. (1990). The social consequences of growing up in a poor neighbourhood. In L. E. Lynn & M. G. H. McGahey (Eds.), *Inner-city poverty in the United States* (pp. 111–186). Washington, DC: National Academy Press.
- Keating, D. P., & Hertzman, C. (Eds.) (1999). *Developmental Health and Wealth of Nations*. New York, NY: Guilford.
- Kershaw, P., Irwin, L., Trafford, K., & Hertzman, C. (2005). *The British Columbia atlas of child development* (1st ed., Vol. 40). Vancouver, BC: Human Early Learning Partnership, Western Geographical Press.
- Kohen, D. E., Brooks-Gunn, J., Leventhal, T., & Hertzman, C. (2002). Neighborhood income and physical and social disorder in Canada: Associations with young children's competencies. *Child Development, 73*, 1844-1860.
- Kohen, D. E., Hertzman, C., & Brooks-Gunn, J. (1998). *Neighborhood influences on children's school readiness*. Hull, QB: Human Resources Development Canada.
- Leventhal, R., & Brooks-Gunn, J. (2000). The Neighborhoods They Live in: The Effects of Neighborhood Residence on Child and Adolescent Outcomes. *Psychological Bulletin, 126*, 309-337.
- Lloyd, J. E. V., & Hertzman, C. (2009). From Kindergarten readiness to fourth-grade assessment: Longitudinal analysis with linked population data. *Social Science & Medicine, 68*, 111-123.

- Leventhal, T., Fauth, R. C., & Brooks-Gunn, J. (2005). Neighborhood poverty and public policy: A 5-year follow-up of children's educational outcomes in the New York City Moving to Opportunity demonstration. *Developmental Psychology, 41*, 933-952.
- McCain, M. N., & Mustard, J. F. (1999). *Early Years Study: Reversing the Real Brain Drain*. Toronto, ON: Publications Ontario.
- Mustard, C., Derksen, S., Berthelot, J.-M., & Wolfson, M. (1999). Assessing ecological proxies for household income: a comparison of household income and neighborhood level income measures in the study of population health status. *Health and Place, 5*, 157-171.
- Nosbush, L. (2006). *Community report for children. The Understanding the Early Years final report*. Prince Albert, SK: Saskatchewan Rivers School Division. Retrieved on November 20, 2007, from <http://www.srsd119.ca/uey/>
- Oliver, L., Dunn, J. R., Kohen, D., & Hertzman, C. (2007). Do neighborhoods influence the readiness to learn of Kindergarten children in Vancouver? A multilevel analysis of neighborhood effects. *Environment and Planning, 39*, 848-868.
- Rauh, V. A., Lamb Parker, F., Garfinkel, R. J., Perry, J., & Andrews, H. F. (2003). Biological, social, and community influences on third-grade reading levels of minority Head Start children: A multilevel approach. *Journal of Community Psychology, 31*, 255-278.
- Rouse, C., Brooks-Gunn, J., & McLanahan, S. (Eds.) (2005). School readiness: Closing racial and ethnic gaps [Special issue]. *Future of Children, 15*, 5-196.

- Schweinhart, L. J., & Weikart, D. P. (1997). The High/Scope Preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly, 12*, 117-143.
- Scott-Little, C., Kagan, S. L., & Frelow, V. S. (2006). Conceptualizations of readiness and the content of early learning standards: The intersection of policy and research? *Early Childhood Research Quarterly, 21*, 153-173.
- Sue, S., & Okazaki, S. (1990). Asian-American educational achievements: A phenomenon in search of an explanation. *American Psychologist, 45*, 913-920.
- Willms, J. D. (1999). Quality and inequality in children's literacy: The effects of families, schools, and communities. In D. P. Keating & C. Hertzman (Eds.), *Developmental health and the wealth of nations* (pp. 72-93). New York, NY: The Guilford Press.
- Willms, J. D. (2003). *Ten hypotheses about socioeconomic gradients and community differences in children's developmental outcomes*. Gatineau, QB: Human Resources Development Canada; ARB report, SP-560-01-03E.
- Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press.

5. Concluding chapter

In this dissertation, a theoretical framework, integrating Bronfenbrenner's bioecological theory of human development, theories from the school readiness and neighborhood effects literatures, and considerations from validity theory, was proposed for the EDI project. Then, two empirical studies based on the EDI project, and pertaining to gender and language background differences in developmental outcomes were presented. In this concluding chapter, the following points will be addressed. First, the connections between the theoretical framework proposed in chapter 2 and the two EDI-based empirical studies presented in chapters 3 and 4 are outlined. Second, the strengths and limitations of the presented research are discussed, and a number of questions that have evolved from the research are delineated. Third, a number of working hypotheses and research design recommendations for future analyses are illustrated. Finally, the contribution of the presented research to the research literature is discussed.

Connections between the Presented Theoretical Framework and Empirical Studies

Developmental research conducted with the purpose to provide knowledge that can inform developmental and educational practice is particularly concerned about external validity or generalizability questions. In other words, a key concern about research findings is commonly about whether the findings can validly be applied to other (sub)populations, contexts, and/or times than the one studied. There are numerous examples of how research findings have been generalized to other populations and other contexts, only to find out that these generalizations were based on untenable assumptions.

For example, intervention programs that work in one context have been found to be ineffective or even counterproductive in others³⁴ (see Smith, Pepler, & Rigby, 2004, for an example). The point is that for the type of developmental research that is conducted within the EDI project, it is necessary to continuously check the assumptions that underlie one's inferences and generalizations.

Validity researcher Kane (2006) has metaphorically likened this procedure of checking one's assumptions to a "chain of inferences", and the references to Messick (1995), in chapter 2, have, accordingly, emphasized the importance of explicitly stating and checking one's assumptions in conducting and applying developmental and educational research. In essence, this argument connects the proposed theoretical framework to the two empirical studies: The integrated theoretical framework explicates the assumptions that underlie constructs such as school readiness, neighborhood effects, constructs pertaining to person characteristics, such as gender and first language background, and constructs referring to context variables, such as socioeconomic status. In addition, it provides principles according to which these constructs can be viewed in relation to each other. The two empirical studies then provide two examples of how certain (implicit or explicit) assumptions related to the EDI project can be tested.

The first study examined one aspect of the (implicit) assumption that the EDI measures the school readiness construct similarly for different groups of the Kindergarten population—as EDI scores, or vulnerability rates, are being used to refer to gender

³⁴ It must be noted that such differences in program effectiveness can often be attributed to a number of different factors (e.g., target group differences; differences in implementation; different levels of external support).

differences, ESL versus non-ESL differences, and differences between children with or without Aboriginal background.

In a similar vein, the second empirical study examines whether gender differences as well as ESL versus non-ESL differences and differences between different language groups (English, Punjabi, Cantonese) are consistent across the socioeconomic spectrum. In other words, both empirical studies have in common that they theoretically question and empirically test whether statements about school readiness, based on EDI scores, can be generalized according to, for example, gender and socioeconomic status, or whether such statements need to be qualified with references to moderating variables.

Strengths and Limitations of the Presented Research

The main strengths of the research presented here are conceptual, as the research illustrates the value of integrating theoretical work and then applying it to population-level data, employing state-of-the-art statistical methods. By doing so, it is illustrated, based on the EDI project, how developmental *science in the discovery mode* benefits from coherently interconnecting theoretical interpretations, methodological considerations, and empirical examinations in an ongoing, alternating manner. In addition, the research shows how this process can help to address some long-standing questions in the area of school readiness research and related areas. In particular, throughout the theoretical discussion, the importance of *proximal processes* with regard to development is repeatedly highlighted, and a theoretical link to *developmentally*

appropriate practices (Bredekamp, 1986) is established³⁵. For the empirical studies, as mentioned, the data did not provide measures of proximal processes. The lack of data for proximal processes of relevance for the developmental outcomes studied is a limitation of the current analyses, and is a yet missing piece in the EDI project BC. At the same time, the analyses provide critical information for future research, especially with regard to research that aims to obtain, analyze, and interpret proximal process data in regard to the same developmental outcomes. In fact, the presented analyses provide indications for where to look for differences with regard to such developmentally relevant proximal processes between subgroups of the population. This issue will be discussed in the following with regard to unpacking the meaning of variables such as gender, language group, and socioeconomic status, and addressing the implications for future research.

Gender. In an abundance of developmental studies, the variable of gender is used as a criterion to differentiate between groups. Routinely, the variable is defined by a person's biological gender. At the same time, gender has socially constructed meaning, as far as gender roles, expectations, opportunities for girls and boys, women and men are concerned, and this meaning might be particular to a given society or (sub)culture and a given time. This situation leads to some interpretive challenges and potential misunderstandings, because the criterion for creating gender groups may be universally the same, whereas the developmental relevance of gender might be different from context to context, from time to time, and from person to person. The implications of this

³⁵ Rimm-Kaufman and Pianta (2000), who applied Bronfenbrenner's bioecological theory to the transition to Kindergarten, review empirical findings on what type of proximal processes are positively related to children's developmental trajectories during that transition period.

ambiguity of the gender construct are discussed below, in conjunction with the implications derived from the discussion of the language group and socioeconomic status variables.

Language group. Like gender, the constructs of ethnicity, race, language group, country of origin, religion, Aboriginal background, and so forth—which may all be considered variants of a theme—have been used widely as criteria for dividing people into groups in developmental research (or other areas, such as health research). Like using the gender variable, this procedure of dividing people into groups implies a certain degree of homogeneity within and heterogeneity between those groups with respect to the phenomenon that is studied—because overall group differences are then often attributed to and generalized across the grouping variable. However, what applies to the case of gender also applies here. A variable such as language background may be constructed according to an objective, universal criteria, but the developmental relevance of such a variable may, again, be quite different from context to context, from time to time, and even from one person to the next within the same context.

Socioeconomic status. The same that was said with regard gender and language group (and its variants) can be said with regard to socioeconomic status. If defined according to income, education, or professional status (or any combination thereof), socioeconomic status can be readily used as a criterion for separating groups. The critical question with regard to developmental research, though, pertains to the question of what the developmental relevance of the socioeconomic status construct is.

The preceding arguments are not presented to dismiss the use of gender, language group, or socioeconomic criteria as research variables. To the contrary, such analyses can provide invaluable insights, if interpreted with adequate cautions: First, it is imperative to not automatically assume homogeneity (in terms of the developmental outcome) within groups that are simply created according to easily available and quantifiable criteria (such as gender, language background, and socioeconomic status). Rather, it is necessary to explore such groups with regard to homogeneity or heterogeneity, and to clearly state the theoretical assumptions for why the criteria should be related to developmentally relevant processes. Second, it is critical to examine to what extent such variables (which may or may not be of developmental significance) are correlated (or confounded) with other variables that are of developmental relevance. For example, socioeconomic status might be related systematically to developmentally relevant parental expectations and parent-child interactions. Accordingly, it is necessary to theoretically and empirically examine the relationship between developmental outcomes and variables such as gender, language background, and socioeconomic status in terms of the underlying proximal processes that are the actual factors causally related to the developmental outcomes.

These issues have, to some extent, been addressed in the proposed theoretical framework, and also in the presented empirical analyses. For example, it was examined whether gender differences are similar across the socioeconomic spectrum; a stronger interaction between the two variables would have suggested that being a girl versus being a boy has different meanings—or, is associated with different processes—depending on socioeconomic status. This being said, it is important to note that most of the variation of the EDI scores occurs within the gender groups, not between them.

The language group analyses across the socioeconomic spectrum have shown that overall group differences in the average scores of Punjabi-, Cantonese-, and English-speaking children disappear once socioeconomic status is taken into account. In other words, at comparable levels of socioeconomic status, all groups seem to be doing equally well (except on the communication skills and general knowledge domain)—the Punjabi and Cantonese groups are simply distributed along the lower range of the socioeconomic spectrum than the English group. At the same time, the language group comparisons show that the socioeconomic status gradient is not the same for the three different groups, as the gradient is pronounced for the Cantonese and English group and flat for the Punjabi group. This implies that socioeconomic status does not capture developmentally relevant processes for the three groups in the same way. Given that the patterns of these relationships were slightly different for the different developmental domains—for example, Cantonese-speaking children receive the highest ratings in the language and cognitive development domain and the lowest scores on the communication skills and general knowledge domain—it can be hypothesized that there are language group-related cultural differences that are developmental outcome specific, and potentially moderated by factors associated with socioeconomic status.

Working Hypotheses for Future Research

Of the five hypotheses that were proposed for the multi-level analyses of gender as well as language group interactions with socioeconomic status, only one was supported empirically; namely, that the Cantonese group obtained relatively high scores in the language and cognitive development domain, and relatively low scores in the

communication skills and general knowledge domain. The data did not confirm the hypotheses suggested by the theory of double jeopardy, saying that the ‘handicap of being male and poor’, or, equivalently, the double barrier of being ESL and being poor results in an interaction between these variables, to the extent that the gap between boys and girls or ESL and non-ESL children widens at the low end of the socioeconomic spectrum. In order to interpret the findings, it is necessary to revisit the theoretical arguments, and to also explore potentially confounding methodological issues.

The theory of functional relativism (Sue & Okazaki, 1990) is, like Bronfenbrenner’s bioecological theory of human development (Bronfenbrenner & Morris, 2006), a theory that emphasizes the developmental relevance of interactions between personal and contextual characteristics. It states that human beings develop their relative strengths in relationship to perceived situational opportunities. Following this idea, it could be hypothesized that gender differences as well as language group differences are, at least to some extent, driven by the combination of children’s personal profiles of (actual or perceived) strengths and weaknesses and the (actual or perceived) opportunities and expectations presented to them by their environment. In part, such personal profiles of (relative) strengths and weaknesses as well as opportunities and expectations may be distributed along gender and language group lines, due to societally and/or culturally shared norms. On the other hand, the development of individuals will be affected by the potency of individual strengths and environmental opportunities, even if they are deviating from the societal and/or cultural expectations. The argument thus is that future studies that examine children’s EDI profiles with respect to information

collected on children's expectations and motivations and/or gender-specific or language group/culture-specific expectations and motivations would significantly contribute to a better understanding of the findings.

Finally, it is hypothesized, in line with Entwisle, Alexander, and Olson's (2007) finding that a gender gap with regard to reading skills widens over time at the low end of the socioeconomic spectrum, that the effects of relative functionalism become more pronounced over time. In other words, it is anticipated that the joint effects of a person's relative strengths and weaknesses and the environmental expectations and opportunities on development become more noticeable the longer these factors have shaped a person's development.

Future Directions

In this section, a research design is sketched out that targets the questions and hypotheses that have evolved from the research in this dissertation. In essence, this research design follows the recommendations that are made with regard to the process-person-context-time model presented in the theoretical chapter. This approach not only has theoretical advantages, but also helps to conceptually, logistically, and methodologically manage such a research program.

In the EDI project, data are available on children's personal characteristics, on context characteristics, and on developmental outcomes in a number of different domains, and at different points of time in their lives. In order to understand, for example, the emergence of developmental gender outcomes in certain domains, it will be very interesting to look at the interaction between gender and time with regard to the different

developmental outcomes. As described above, the gender gap might be moderated by the socioeconomic characteristics of the children's environment. In other words, it will be necessary to examine the two-way (or second-order) interactions between gender, time, and socioeconomic status with regard to developmental outcomes. Finally, as also described above, such two-way interactions may look different in form and magnitude for different developmental outcomes—a phenomenon that, statistically, would be equivalent to a three-way interaction.

Accurately interpreting higher-order interactions is complex, conceptually and statistically. Therefore, an approach is recommended in which each lower-order interaction is examined by itself, before higher-order interactions are addressed. As the gender*socioeconomic status analyses have shown, interactions that were statistically significant in the multilevel analyses were not significant in consequent analyses that took into account that assumptions of the multilevel analyses were violated. Given the nature of the construct studied, the violation of statistical assumptions, such as non-normality or heterogeneity of variances, can be expected to be the rule rather than the exception.

The last recommendation refers to the collection of proximal process data, or proxies thereof. The reality is that the collection of such data is exponentially more resource-intensive than collecting more easily obtainable person and context characteristics. This being said, there are efforts under way that seek to collect such data via, for example, tracking of children's participation in certain programs and activities, student self-reports, parent surveys, neighborhood residence surveys, and so forth. With an eye toward efficiency, it may be desirable to focus on data that refer to proximal

processes that are not only developmentally significant, but that are also referring to processes that can be affected and changed—after all, the eventual goal of applied developmental research is to translate research evidence into actions that can prevent developmentally detrimental processes, and promote developmentally beneficial ones. This notion is elegantly captured in the words of Dana Brynelsen, when she, in expressing her approach to her work with children, families, and communities, asks: *‘What differences can we make that make a difference?’*

A note of caution

Before addressing the contributions of the research presented in this dissertation, a few words of caution are due. It has already been mentioned that proximal processes are considered the ‘engines of development’, and there is apt research evidence to corroborate this claim. In fact, in Rimm-Kaufman and Pianta’s (2000) discussion of an ecological perspective on the transition to Kindergarten, which similarly draws from Bronfenbrenner’s theory, their review of the literature clearly identifies a number of such proximal processes. Evidently, it cannot realistically be expected that population-based research and, for that matter, EDI research fully captures the richness of such proximal processes. Also, the theoretical framework that has been presented here was not intended to convey the notion that research designs that do not adhere to the process-person-context-person model are less insightful or less valid than those that do fit that model: Quite to the contrary, research that painstakingly examines the minutia of proximal processes with regard to developmental outcomes is, of course, indispensable—regardless of whether it manages to simultaneously examine person or context

characteristics. So, the purpose of this theoretical integration has much rather been to convey the notion that any research approach to school readiness—or any other developmental construct, for that matter—can thrive on a heightened awareness of its limitations—or of ‘knowing what one does not know’. The theoretical integration presented here has thus also delineated the limitations of the EDI and the EDI project—and it is hoped that by doing so, it can motivate and facilitate a process in which the EDI project engages in and partners with theoretical and empirical research that provides those missing pieces en route to creating research knowledge that can be applied to help children thrive.

Contribution of the Dissertation to the Research Literature

The dissertation was introduced by a brief reflection on the current high visibility of the school readiness construct in the arenas of education, developmental and educational research, and politics. As indicated in the previous section, one of the fundamental questions that is asked in regard to school readiness (or, more generally, with regard to fostering young children’s education and development during the early years) has been, ‘What differences can we make to make a difference?’

The research presented in this dissertation has addressed a number of points that are relevant for this question as far as the research arena is concerned. Specifically, the presented research has shown that the school readiness construct has been approached from different disciplines and their respective disciplinary lenses. This has led to multiple interpretations, meanings, and connotations of the school readiness construct. The theoretical work presented here claims that developmental research can greatly benefit

from bringing these different disciplinary literatures, theories, research tools, epistemologies, and ontologies together.

In addition, it has been proposed that Bronfenbrenner's bioecological theory of human development provides specific principles and recommendations for how to organize such an integration conceptually, logistically, and methodologically. The theoretical and empirical work in this dissertation has been aimed at contributing toward the goal of coherently integrating a number of different disciplinary theoretical and methodological approaches to school readiness specifically, and early children's development in general. In addition, it has been shown how such an integration can meaningfully refine research questions, inform research designs and methodologies, guide validation strategies, lead to more differentiated interpretation of research findings, and, ultimately, contribute to the translation of research findings into practice that help children thrive.

Proposing that Bronfenbrenner's bioecological theory is of scientific utility is certainly not original. However, Bronfenbrenner himself refers to it as a continuously developing theory, and urges that its propositions and principles are critically applied anew whenever one is confronted with a research question that presents itself in a new light. This dissertation has undertaken first steps in this direction as far as the EDI project and developmental research questions related to it are concerned. It now remains to be seen whether the theory is as practical as it is intended to be. In other words, it is hoped that the dissertation contributes to knowledge that helps 'to make the difference we can make for the wellbeing of our children'.

References

- Bredenkamp, S. (1986). *Developmentally Appropriate Practice*. Washington, DC: National Association for the Education of Young Children.
- Bronfenbrenner, U., & Morris, P. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Development, Volume 1, Theoretical Models of Human Development*, (6th ed.), 793-828. Hoboken, NJ: John Wiley & Sons.
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *The Sociology of Education*, 80, 114-138.
- Kane, M. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th edition). Westport, CT: American Council on Education/Praeger.
- Messick, S. (1995). Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An ecological perspective on the transition to Kindergarten: A theoretical framework to guide empirical research. *Journal of Applied Developmental Psychology*, 21, 491-511.
- Smith, P. K., Pepler, D., & Rigby, K. (Eds.) (2004). *Bullying in schools: How successful can interventions be?* New York, NY: Cambridge University Press.
- Sue, S., & Okazaki, S. (1990). Asian-American educational achievements: A phenomenon in search of an explanation. *American Psychologist*, 45, 913-920.