Modeling Parental Involvement

A Psychometric Framework

Srinivasanan Bala

Modeling Parental Involvement: A Psychometric Framework

Modeling Parental Involvement: A Psychometric Framework

Edited by
Srinivasanan Bala

Published by Vidya Books,
305, Ajit Bhawan,
21 Ansari Road,
Daryaganj, Delhi 110002

Edited by Srinivasanan Bala
ISBN: 978-93-5431-889-4

Table of Contents

1 Introduction

Although the role of parental involvement in student achievement (and reading literacy in particular) is widely acknowledged, research findings regarding its effect differ considerably (Fan and Chen 2001; Hoover-Dempsey et al. 2001; Mattingly et al. 2002; Jeynes 2005; Patall et al. 2008). Owing to the large variation in the methods used to measure student achievement and parental involvement in these studies, it is difficult to establish whether the inconsistency in the results is caused by differences between educational systems and cultures, or by the method applied and the instruments used. Empirical research is required into measurement of student achievement and indicators of the parental involvement, and how comparisons can be made between educational systems (countries), to find out to what extent and under which conditions, parental involvement influences student achievement. In-depth analyses of large-scale international comparative data, such as that contained in the Progress in International Reading and Literacy Study (PIRLS) undertaken by the International Association for the Evaluation of Educational Achievement (IEA), may provide valuable additions to the research into parental involvement.

The main purpose of this research was to develop a psychometric framework for assessing the relation between parental involvement and reading literacy, using the PIRLS-2011 data as a case study. The framework included country specific differences, both at the item and scale level, to gain insight into cultural differences in the parental involvement construct and its relation to student achievement in reading literacy. One of the possible limitations of international large-scale assessment studies such as PIRLS is the extent to which the data of different countries can be usefully compared. Despite the high quality demands for the translation of the instruments and the conditions of administration in each participating country, cultural differences could influence the international validity of the indicators measured. PIRLS offers international scales based on partial credit item response theory (IRT) models, which can improve the quality of secondary analyses of data. However, the differences in the Cronbach's alphas of scales reported in the PIRLS-2011 international report suggest that the meaning of a scale and the interpretation of items within a scale could differ between countries (Martin and Mullis 2012).

The main research objective of this study was to assess the extent to which parental involvement is related to the achievement of primary school students in reading literacy, taking into account student background characteristics, differences between schools and countries, and cultural differences in the parental involvement construct. More specifically, the study aimed to respond to the following research questions:

(1) *Which dimensions of parental involvement can be discerned and to what extent is there empirical evidence that these dimensions are related to student attainment?*

To answer this first question, we reviewed the recent research literature on parental involvement. The main goal was to identify dimensions of parental involvement and to gain further insight into the extent to which the different dimensions were potentially related to student achievement in reading literacy and other achievement-related outcomes.

(2) *To what extent are there any cultural differences (differences between countries) in the components that measure dimensions of parental involvement?*

To address this question, five extracted item sets using IRT models were studied for item-by-country interactions in item parameters, indicating cultural differential item functioning (CDIF). The five scales were subsequently modeled using random and country-specific item parameters for the most extreme interactions. A bi-factor IRT model was applied, where the country-specific covariance matrix gives an indication of the extent to which the scale loads on the intended latent variable and the extent to which the responses are country-specific.

(3) *To what extent are the different dimensions of parental involvement related to student achievement in reading literacy, taking into account student background characteristics and differences between countries?*

A multilevel analysis of the PIRLS-2011 datasets from 41 countries explored the relationship between parental involvement and student reading literacy. For all 41 countries, data from students, schools and parents were available. The analyses used a three-level model (consisting of a student, school and country level). The third question was addressed by exploring the three-level model with a random intercept with fixed effects for the predictors and by exploring a random model for each component showing a meaningful relation with achievement in the random intercept model. This random model explores the extent to which the association between the dimensions of parental involvement and student reading literacy differ between countries. Subsequently, a model without a correction for country differences in the parental components was compared with models with different types of corrections for country differences in the parental components.

In Chap. 2, we review the current literature on parental involvement and its perceived influence on student achievement. Chapter 3 covers the measurement of parental involvement in the PIRLS background questionnaires. In the first part of Chap. 4, we present a psychometric framework to identify and model CDIF in

multiple ways, including a non-standard application of the bi-factor model. We discuss the results of all approaches for each of the five parental involvement components in the second part of the chapter, providing insights into the extent to which they are affected by CDIF. The last step in the psychometric framework was to relate the parental involvement components to reading literacy, and compare the outcomes for the different methods used to model CDIF. In Chap. 5, we present the method and the multilevel analyses. Finally, in Chap. 6, we evaluate the meaning of the results both for measuring parental involvement in an international context, and in discussions of the importance of parental involvement for student attainment.

References

Fan, X. T., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13*(1), 1–22.

Hoover-Dempsey, K. V., Battiato, A. C., Walker, J. M. T., Reed, R. P., DeJong, J. M., & Jones, K. P. (2001). Parental involvement in homework. *Educational Psychologist, 36*(3), 195–209. doi:10.1207/S15326985EP3603_5.

Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban Education, 40*(3), 237–269. doi:10.1177/0042085905274540.

Martin, M. O., & Mullis, I. V. S. (Eds.). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA, USA: TIMSS & PIRLS International Study Center, Boston College.

Mattingly, D. J., Prislin, R., McKenzie, T. L., Rodriguez, J. L., & Kayzar, B. (2002). Evaluating evaluations: The case of parent involvement programs. *Review of Educational Research, 72*(4), 549–576. doi:10.3102/00346543072004549.

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). Parent involvement in homework: A research synthesis. *Review of Educational Research, 78*(4), 1039–1101. doi:10.3102/0034654308325185.

Literature Review

2.1 Introduction

Parental involvement, that is "parents' interactions with schools and with their children to promote academic success" (Hill and Taylor 2004, p. 1491), is an umbrella term that includes a variety of behaviors and activities of parents directly or indirectly related to the education of their children. Therefore, the review of the literature had three main goals. First, to investigate the underlying (sub)dimensions of the parental involvement concept in the research literature and theories about how these dimensions of parental involvement are related to educational outcomes. Second, based on empirical studies, to explore which of these dimensions of parental involvement show the most potential to be related to student attainment, and with reading literacy in particular. The third aim of the review was to identify any indication of cultural differences between countries in how parental involvement is perceived and its relation with student attainment.

A literature search was conducted during the spring of 2015 using the social research databases ERIC, Web of Science, Scopus and PsycINFO. The search keys included (combinations of) parental involvement, parents, parental support, parent-child relationship, parent-school partnership, home environment, student achievement, student performance, reading performance, and reading literacy.

In reviewing the literature, we did not aim to present an exhaustive overview of all published studies about this subject. To better explore the relationship between dimensions of parental involvement and student attainment (the second goal of this review), we targeted our search by applying a number of exclusion and inclusion criteria. Non-empirical studies and empirical studies that were published before 2000 were excluded. However, to obtain comprehensive knowledge of the (sub) dimensions underlying parental involvement and incorporate theories about how these dimensions of parental involvement are related to educational outcomes, we included theory-based publications and some older publications.

Because a systematic quality analysis of the published studies was not part of this review, only studies published in peer-reviewed journals were included, to ensure a certain methodological rigor. As PIRLS investigates the reading literacy of grade 4 students, and research has indicated that parental involvement and its effects change as children become older (Domina 2005; Mattingly et al. 2002), only studies focusing on primary school students were included. Studies focusing on special groups (e.g., students from minority groups, dyslexic students, and urban students) were also excluded, as this study does not focus on these special groups and comparisons between them across countries are difficult. Furthermore, we only included studies that took the effects of variables related to socioeconomic status (SES) at family or school level into account. Finally, although this part of the review focused mainly on studies investigating the relationship between parental involvement and students' reading literacy, studies using other achievement-related outcomes (such as math achievement, grade retention, and motivation) were also included.

Because of the abundance of studies on this subject, several meta-studies (meta-reviews or meta-analyses) have been published. We used the same exclusion and inclusion criteria for these studies. However, in this case, studies focusing (also) on secondary school students were included, since most of the available meta-studies included these students in their review.

For the second part of the review, 22 articles were summarized, including nine meta-studies. We present these studies within a general framework for parental involvement (Sect. 2.4), based on the information derived from the primary literature review (see Sect. 2.3).

2.2 Theories About the Importance of Parental Involvement for Student Attainment

As parental involvement is one of the most malleable factors of the student's home situation, it has become a relevant subject for schools and educational policy. Encouraged by studies showing a positive relationship between parental involvement and student attainment, educational policy makers in the USA, and in western European countries such as the UK and the Netherlands, have initiated and stimulated parent participation programs over the last 15 years (Mattingly et al. 2002; Driessen et al. 2005; Bakker et al. 2007; Patall et al. 2008; Selwyn et al. 2011). Schools are increasingly using digital technologies to support and promote the involvement of parents, by providing laptop computers, online intranets or learning platforms (Selwyn et al. 2011). This encouragement is sometimes less voluntary; in the UK, some schools present parents mandatory school-home agreements about their involvement in homework and expected behavior, such as attending parent meetings (Selwyn et al. 2011).

The underlying assumption of these initiatives is that parental involvement and educational outcomes are not only correlated, but that parental involvement actually influences educational outcomes (Mattingly et al. 2002). In the studies we reviewed, we found theoretical discussions concerning the influence of parental involvement on student achievement that provided conditions under which students may profit from this involvement. Hoover-Dempsey and Sandler (1997) developed a model summarizing the factors influencing parental involvement at five levels:

(1) the decision to become involved (e.g., school invitations, parent role construction);
(2) the choice of type of involvement (e.g., parents' skills, knowledge, and available time);
(3) how involvement influences school outcomes (e.g., modeling, reinforcement, instruction);
(4) tempering or mediating variables (e.g., fit between parents' involvement actions and school expectations); and
(5) student outcomes (e.g., achievement, self-efficacy).

How parental involvement influences school outcomes, namely the third level in Hoover-Dempsey and Sandler's (1997) model, was the main focus of McNeal (1999). McNeal (1999) argued that the involvement of parents in their child's education influences student outcomes through three mechanisms. The first mechanism, socializing, refers to home-based involvement, such as supervising homework, by which parents emphasize the importance of schooling. Generating social control through school-based involvement is the second mechanism; school-based involvement offers parents the opportunity to develop relationships with teachers and other parents, and in discussing their child's behavior, to learn from them. The last mechanism is having access to insider information by communicating with the school. By this mechanism, parents will be, for example, earlier and better informed about the available solutions in instances of learning or behavioral problems. McNeal's theory specifies different outcomes for the three mechanisms: socialization and social control affect the attitude, motivation and behavior of the student, while having access to insider information has a direct effect on both cognitive and behavioral student outcomes (McNeal 1999).

Generating social control and having access to insider information can also be regarded as part of a mechanism called increasing social capital (Hill and Taylor 2004). By being engaged in school-related activities, parents learn about schools' expectations of student behavior and homework and learn how to help with their child's homework and learning at home. They also learn from other parents about available and suitable options to improve their child's learning. At the same time, teachers learn about the parental expectations of their students.

Grolnick and Ryan's (1989) theory on "parenting styles" focused on the effects of home-based parental involvement. According to this theory, three dimensions of parenting style are important for the learning behavior and achievement of students. The first is support for autonomy, by which children are encouraged in independent problem solving and participation in decision making. The second style, direct

parental involvement, refers to the extent to which parents are interested in, knowledgeable about, and taking an active part in their child's life. The third parenting style that may have a positive influence on student attainment is provision of structure (i.e., providing clear and consistent guidelines and rules) with regard to homework or other school-related aspects. Cooper et al. (2000) applied this theory in their study on parental involvement in homework and showed (1) a positive relationship between support for autonomy and student achievement, and (2) a negative association between direct parental involvement and student achievement. These findings are explained by the student's performance at school; parents of low-performing students become more proactively involved with their child's learning, while parents of students doing well at school, are more likely to support the autonomy of their child.

With regard to the role of parental support for homework, it is widely assumed that when parents help their child with homework there are positive effects on student learning and academic achievement; the student will study more efficiently, effectively and with greater focus (Hoover-Dempsey et al. 2001; Patall et al. 2008). In the long term, parental involvement may have a positive effect on student's ability to engage in adaptive self-regulation by promoting the development of learning strategies such as goal-setting, planning, time management, and attentiveness (Patall et al. 2008). Hoover-Dempsey et al. (2001) described three reasons why helping with homework may have these positive effects: (1) modeling, (2) reinforcement, and (3) instruction (see also their general model for parental involvement, mentioned previously). First, while helping their child with homework, parents can serve as salient models. This is based on the idea that children learn through observation. Parents are influential role models because they possess skills and abilities that children value highly. Because there are no direct consequences of the child's performance at home (in contrast to school), home provides a safe environment where the parent becomes an even more powerful role model. The second reason is reinforcement; by providing positive consequences in response to the child's homework behaviors, the child is stimulated to demonstrate similar skills and behaviors again. Parents may even have an advantage over teachers because they have better insight into which reinforcement strategies are the most effective for their child. Finally, helping with homework may have a positive effect on student achievement because parents tend to use the learning strategy "guided or collaborative learning" (Hoover-Dempsey et al. 2001). This includes directing the child to the task at hand, simplifying the task, providing extra explanation, or relating the task to familiar contexts.

However, some scholars argue that parental involvement with homework may also have negative effects (Patall et al. 2008; Dumont et al. 2012). They point out that the involvement may lead to tensions between parents and children, caused by parental frustrations about the child not performing as expected, or by the frustrations of the child who perceives their parents as too controlling. For lower-achieving children, or parents with unrealistically high expectations, these tensions may have a negative impact on the child's self-esteem and performance at school. Helping with homework might also interfere with learning if parents are not

sufficiently equipped to help, if their instruction is very different from that of the school teacher, or if parents are overly involved, for example, completing assignments themselves (Cooper et al. 2000).

2.3 Measuring Parental Involvement

The literature review reveals that parental involvement is a multifaceted construct that includes a variety of parental behaviors and practices. It is important to view and measure parental involvement as a multidimensional concept, because research has indicated that some types of parental involvement may exert more influence on student performance (Jeynes 2005). Most studies tend to focus on parental involvement at home, at school or at both places (Powell et al. 2012). Bakker et al. (2007) discerned three types of parental involvement: (1) parents' behavior towards and activities with their child at home (home-based involvement), (2) parents' participation in school (school-based involvement), and (3) communication between parents and school (home-school communication). An additional distinction sometimes used is school-initiated parental involvement versus parent-initiated involvement (Driessen et al. 2005).

The most commonly used framework for parental involvement was developed by Epstein at the beginning of the 1990s (Epstein 1992; Manz et al. 2004). This framework refers mainly to school-initiated parental involvement, but is also characterized as a holistic approach to student learning: parents or families, schools and communities should work together as a partnership to create a positive learning environment for children (Mattingly et al. 2002). Epstein (1992) defined six types of parental involvement that can be influenced by the school; these also contain the three types of parental involvement identified by Bakker et al. (2007). Epstein (1992) stated that to increase the involvement of parents, schools and teachers should:

(1) assist parents in child rearing skills (home-based involvement);
(2) communicate with parents regularly (home-school communication);
(3) involve parents in school volunteer opportunities (school-based involvement);
(4) involve parents in home-based learning (home-based involvement);
(5) involve parents in school-based decision making (home-school communication); and
(6) involve parents in school-community collaborations (school-based involvement).

The studies reviewed indicate that home-based involvement is mostly measured from the perspective of parents and sometimes from the perspective of students. With regard to communication between parents and schools, and school-based involvement, this can also be measured from the perspective of parents, and the perspective of school principals or teachers. Based on the literature review, we developed a general framework to categorize the studies reviewed (Table 2.1).

Table 2.1 General framework of parental involvement

Parent perspective	Dimension 1
	Home-based involvement: e.g., helping with homework, (early) literary activities, parent-child discussion of child's schooling, parenting style, parental monitoring and rule-setting, ensuring school readiness
	Dimension 2
	School-based involvement and home-school communication: e.g., attending parent meetings, participating in school activities, parent-teacher interaction
Student perspective	Dimension 3
	Home-based involvement: e.g., getting help with homework, parental monitoring and rule-setting, parent-child discussions about school
School perspective	Dimension 4
	School-based involvement and home-school communication: e.g., providing newsletters, asking parents to help at school, individual teacher-parents meetings

In this framework, the different types of involvement and the perspectives by which they are perceived (i.e., parent, student, or school) are combined, resulting in four dimensions. We used this framework to categorize the meta-studies and empirical studies addressing the relationship between the involvement of parents and student attainment.

2.4 The Relation Between Parental Involvement and Student Attainment

As already mentioned, the goal of the literature review was not to present an exhaustive overview of all the available literature on this subject, but to gain insight into the dimensions of parental involvement that show the greatest potential to be related to student attainment, and to examine the role of cultural differences between countries. We identified large variation in outcomes and in the mechanisms used to conduct these studies (Tables 2.2, 2.3, 2.4, 2.5, 2.6 and 2.7). Most studies focused on more than one dimension; we recorded the main characteristics of the meta-studies (Tables 2.2 and 2.3) and individual studies (Tables 2.4, 2.5, 2.6 and 2.7) for each dimension (whenever available).

The outcomes of the meta-studies and the individual studies in this review indicate that parental involvement is generally positively correlated to or has positive effects on student attainment. This is in agreement with the outcomes of a meta-synthesis of nine meta-analyses by Wilder (2013). With regard to the individual dimensions of parental involvement, the results are less definitive. For example, the association between parental involvement with homework and student achievement is positive in some studies, but non-existent or negative in others (Hoover-Dempsey et al. 2001; Patall et al. 2008). Based on a review of literature,

Table 2.2 Overview of meta-studies examining the relation between (dimensions of) parental involvement and student achievement, from the parent perspective

Dimension	Authors	Description	Dependent variable	Main conclusion
Home-based involvement	Castro et al. (2015)	Meta-analysis on parental involvement in kindergarten, primary and secondary education	Achievement in various subjects	In general positive effects: high academic aspirations, communication about school and stimulation of reading habits show strongest positive relation with student achievement. Supervision of homework is unrelated to achievement
	Erion (2006)	Meta-analysis on the effects of parent tutoring in primary and secondary education	Achievement in reading, language and math	Parent tutoring is effective for improving achievement
	Fan and Chen (2001)	Meta-analysis of experimental studies on the effects of parental involvement on achievement in primary and secondary education	Achievement in various subjects	In general positive effects: parental academic aspirations show strongest positive relation with student achievement. Parental monitoring of homework shows a negative relation, direct aid with homework a positive relation. The correlation is lower if the achievement is measured in specific areas, such as math or reading
	Hill and Tyson (2009)	Meta-analysis on parental involvement in	Achievement in various subjects	In general positive effects, academic socialization

(continued)

Table 2.2 (continued)

Dimension	Authors	Description	Dependent variable	Main conclusion
		secondary education		showed the strongest positive relation with achievement. For homework involvement, the results are inconclusive
	Hoover-Dempsey et al. (2001)	Meta-review, parent involvement in homework in primary and secondary education	Achievement in various subjects	Inconclusive, positive and negative relations with or effects on achievement
	Patall et al. (2008)	Meta-analysis, parent involvement in homework in kindergarten, primary education and secondary education	Achievement in various subjects	In general small correlations, often not significant and they vary with students' age. Setting rules about when and where homework should be done has the strongest positive relation with achievement
	Wilder (2013)	Meta-synthesis of nine meta-analyses on parental involvement	Achievement in various subjects	In general positive effects. Parental expectations (beliefs and attitudes towards school teachers and subjects) showed the strongest relationship. No positive relation for help with homework
School-based involvement and home-school communication	Castro et al. (2015)	Meta-analysis on parental involvement in kindergarten, primary and secondary education	Achievement in various subjects	Involvement in school activities is not related to achievement

(continued)

Table 2.2 (continued)

Dimension	Authors	Description	Dependent variable	Main conclusion
	Hill and Tyson (2009)	Meta-analysis on parental involvement in secondary education	Achievement in various subjects	In general positive effects
	Wilder (2013)	Meta-synthesis of nine meta-analyses on parental involvement	Achievement in various subjects	In general positive effects

McNeal Jr (2012) not only concluded that such inconsistencies have become a "standard" in the parental involvement literature, but also that these inconsistent findings cut across grade levels, measures of student attainment and time periods.

One of the explanations for the mixed results is the complexity of the parental involvement concept (Fan and Chen 2001; Hill and Tyson 2009; Castro et al. 2015). Both the meta-studies and our literature review indicated that measurement of parental involvement dimensions differed considerably among studies. According to Fan and Chen (2001), it is the multidimensional nature of parental involvement that has led to a lack of agreement about definitions and measurement inconsistencies, making it difficult to compare findings across studies. For example, "helping with homework" includes providing space and materials (such as a

Table 2.3 Overview of meta-studies examining the relation between (dimensions of) parental involvement and student achievement, from the school perspective

Dimension	Authors	Description	Dependent variable	Main conclusion
School-based involvement and home-school communication	Mattingly et al. (2002)	Meta-review of evaluations of parental involvement programs	Achievement in various subjects	Inconclusive
	Sénéchal and Young (2008)	Meta-analysis of (quasi) experimental studies on encouragement and training programs for parent-child reading activities and reading ability in kindergarten and primary school	Achievement in reading literacy	Positive effects found for parents helping children to read

Table 2.4 Overview of research literature examining the relation between home-based parental involvement (dimension 1) and student achievement from the parent perspective

Authors	Description	Country	Dependent variable	Main conclusion
Bakker et al. (2007)	Survey, parents' and teachers' perception of parental involvement at home and at school	Netherlands	Achievement in different subjects	Parents' perception is positively related to reading achievement
Barnard (2004)	Survey, parents' and teachers' perception of parental involvement at home and at school	USA	School dropout and highest grade completed at age 20	Parental involvement at home is not related to educational attainment
Driessen et al. (2005)	Survey, school principals' perception of parental involvement at school and policy towards parental involvement, parents' perception of parental involvement at home	Netherlands	Achievement in math and language, school-supportive home climate, well-being and self-confidence	A negative direct effect of help with homework on students' achievement in language and math
Kloosterman et al. (2011)	Survey, parents' perception of parental involvement at home and at school, teachers' perception of parental involvement at school	Netherlands	Achievement in math and language	Parents reading activities are most relevant for child's language achievement in later grades of primary school, parental reading instruction is beneficial for language achievement at the start as well as during primary school
Galindo and Sheldon (2012)	Survey, parents' perception of parental	USA	Gains in math and reading in kindergarten	Activities at home showed no association with

(continued)

Table 2.4 (continued)

Authors	Description	Country	Dependent variable	Main conclusion
	involvement at home and at school			math and reading gains
Lau et al. (2011)	Survey, parents' perception parental involvement at home and at school	China	Achievement in tests for readiness for school including reading literacy at kindergarten and entering primary school level	Language and cognitive home-based activities are positively related to reading literacy
Myrberg and Rosen (2009)	Survey, parents' perception of parental involvement at home (early reading activities)	Sweden	PIRLS-2001 reading literacy test, parents perception of early reading abilities	Early reading activities affect early reading abilities. Educational level of the parents has an effect on reading literacy, via books at home and early reading activities
Powell et al. (2012)	Survey, changes in parents involvement in activities at home and at school from pre-K, kindergarten and grade 1	USA	Achievement in language, reading literacy and math	The degree of within-family change in several dimensions of parents' home-based involvement is related to math achievement but not to reading and literacy skills in grade 1
Senechal and LeFevre (2002)	Survey, parents perception of parental involvement at home (early literacy activities)	Canada	Achievement in language and reading literacy at grade 1 and 3	Storybook reading positively related to children's receptive language development, teaching reading by parents positively

(continued)

Table 2.4 (continued)

Authors	Description	Country	Dependent variable	Main conclusion
				related to children (early) literacy skills, parental involvement not related to phonological awareness skills
Stylianides and Stylianides (2011)	Survey, parent-child interactions (including early literacy activities)	USA	Achievement in different subjects in kindergarten	Low parent-child interaction relates negatively with achievement (incl. reading)
Xu et al. (2010)	Survey, parents perception parental involvement at home and at school	USA	Self-regulated learning perceived by teachers, achievement in reading literacy	TV-rules and help with homework negative relation with reading achievement. Involvement in school, parental educational expectations, engaging children in their homework and encouraging children's' extracurricular participation positive relation with reading achievement

computer), monitoring whether homework is completed, having rules about when and where homework is done, and responding to questions and providing direct homework instruction (Hoover-Dempsey et al. 2001). In the meta-study of Hoover-Dempsey et al. (2001), it was clear that some of these dimensions were more effective than others; this may account for the mixed results.

Not only is the measurement of dimensions of parental involvement often very different among studies, the explanation or justification for the measurements used is also frequently absent. Although many studies use Epstein's (1992) framework as a starting point, most scholars in the field have developed their own indicators for parental involvement, and do not seem interested in developing or using existing

Table 2.5 Overview of research literature examining the relation between school-based involvement and home-school communication (dimension 2) and student achievement from the parent perspective

Authors	Description	Country	Dependent variable	Main results
Bakker et al. (2007)	Survey, parents' and teachers' perception of parental involvement at home and at school	Netherlands	Achievement in different subjects	Contact with the school is negatively related to achievement in general, there is no relation for reading achievement
Barnard (2004)	Survey, parents' and teachers' perception of parental involvement at home and at school	USA	School dropout and highest grade completed at age 20	Parent involvement at school is not correlated to educational attainment
Domina (2005)	Survey, mothers' perception of parental involvement at school and students perception of parental involvement at home	USA	Cognitive (reading and math) and behavioral development	A positive effect for preventing behavioral problems was found, but no effect of school-based involvement on achievement
Galindo and Sheldon (2012)	Survey, parents' perception of parental involvement at home and at school	USA	Gains in math and reading in kindergarten	Activities at school show positive association with math and reading gains
Kloosterman et al. (2011)	Survey, parents' perception of parental involvement at home and at school, teachers' perception of parental involvement at school	Netherlands	Achievement in math and language	School-based involvement positively affects students' language and math achievement at start of primary school, but the impact diminishes in subsequent primary school years

(continued)

Table 2.5 (continued)

Authors	Description	Country	Dependent variable	Main results
Lau et al. (2011)	Survey, parents' perception parental involvement at home and at school	China	Achievement in school readiness tests including reading literacy at kindergarten and entering primary school level	Communication with the school is positively related to reading literacy
Okpala et al. (2001)	Survey, parents' volunteering hours at school	USA	Math achievement grade 3–12	Parental volunteering hours at school is not related to math achievement
Powell et al. (2012)	Survey, changes in parents' involvement in activities at home and at school from pre-K, kindergarten and grade 1	USA	Achievement in language reading literacy and math	The degree of within-family change in several dimensions of parents' school-based involvement is related to math achievement but not to reading and literacy skills in grade 1

empirically-tested measures and scales. This is an important limitation in conducting meta-studies, and makes it complicated to draw general conclusions about the relationship between dimensions of parental involvement and student attainment.

However, the measurement of parental involvement is not the only reason why parental involvement is a complex concept. Although it seems logical to assume that parental support has positive effects on student achievement, it may also be reasonable to assume that low-achieving children need and receive more support and interference from their parents, resulting in a negative influence (Castro et al. 2015). This phenomenon is called the "reactive hypothesis," where parents of students with learning or behavioral problems react by intensifying their involvement in their children's education (McNeal Jr 2012). McNeal Jr (2012) examined the relationship between achievement and parental involvement from grade 8 to grade 12 and concluded that there was no empirical evidence to support the reactive hypothesis, which seems to be largely championed by scholars who may be unwilling to believe that parental involvement could (also) have negative effects (McNeal Jr 2012). Conversely, some studies do seem to provide support for the

Table 2.6 Overview of the research literature on the relation between home-based parental involvement (dimension 3) and student achievement, student perspective

Authors	Description	Country	Dependent variable	Main results
Domina (2005)	Survey, mothers' perception of parental involvement at primary school and students perception of parental involvement at home	USA	Cognitive (reading and math) and behavioral development	No effect found for home-based involvement on achievement
Dumont et al. (2012)	Survey, students' perception of parental involvement with homework	Germany	Achievement in math and reading, academic self-concept and homework self-efficacy	Positive effects found for perceived parental competence to help with homework and parental support on student outcomes, negative effect of perceived homework conflict on achievement

reactive hypothesis. When the association between parental involvement and achievement is controlled for the initial ability, intelligence or the SES of the student, the association is more likely to be positive (Wilder 2013; Castro et al. 2015).

A study by Domina (2005) found that the initial positive effects of parents' involvement in school and at home on the reading and math achievement of primary school students, changed to non-significant or even negative effects when controlled for student background characteristics. This indicates that the effects of student background characteristics on the relation between parental involvement and student outcomes are also unclear. While most studies show a positive relationship between SES and parental involvement (Cooper et al. 2000; Fan and Chen 2001; Mattingly et al. 2002), a study of Dutch parents of primary school students (Stoep et al. 2002) revealed that home-based involvement was higher among lower SES parents than among high SES parents. Yet another Dutch study found no relation between the educational level of the parents of primary school children and involvement in their child's education (Bakker et al. 2007).

Whether the reactive hypothesis should be rejected or not, the lack of consensus confirms the difficulties introduced in proposing that parental involvement influences student attainment. One of the meta-studies focused specifically on the effects of parental involvement programs offered by schools (Mattingly et al. 2002). The results were disappointing; according to Mattingly et al. (2002), studies evaluating the effectiveness of these programs did not provide convincing evidence that these

Table 2.7 Overview of the research literature on the relation between school-based involvement and home-school communication (dimension 4) and student achievement, school perspective

Authors	Description	Country	Dependent variable	Main results
Bakker et al. (2007)	Survey, parents' and teachers' perception of parental involvement at home and at school	Netherlands	Achievement in different subjects	Teacher perception of parental involvement is positively related to reading achievement
Barnard (2004)	Survey, parents' and teachers' perception of parental involvement at home and at school	USA	School dropout and highest grade completed at age 20	Teacher perception of parental involvement is positively related to educational attainment
Driessen et al. (2005)	Survey, primary school principals' perception of parental involvement at school and policy towards parental involvement, parents' perception of parental involvement at home	Netherlands	Achievement in language and math, school-supportive home climate, well-being and self-confidence	No direct effect on student outcomes found for school activities to involve parents

programs had any positive effects on student achievement or other achievement-related outcomes. Furthermore, most of these studies were correlational or had other design limitations, which made it difficult to determine whether there was a causal effect.

Another limitation that may have contributed to the inconsistencies in outcomes apparent in this field of research was mentioned by Hoover-Dempsey et al. (2005). In their review of parental motivations for involvement, they recognized that much research on parental involvement has relied on single-source reports. Parents' perception of their school-based involvement can be different from the school's perception of parental involvement. Students and parents may also differ in their perception of home-based involvement. Using parent, as well as student and school information, enables triangulation of essential perspectives on involvement, and thereby allows a more precise determination of parental involvement and its influence on student outcomes.

In our literature review, we also focused on whether cultural differences in parental involvement could also explain some of the differing research results. However, most of the studies reviewed were conducted in the USA, and none of the

meta-studies explicitly incorporated this factor. A study comparing the definition of parental involvement for European-American parents versus that of immigrant-Chinese parents in the USA revealed that European-American parents were more involved in school-based activities while the immigrant-Chinese parents focused more on systematic teaching of their children at home (Huntsinger and Jose 2009). Another study compared the home-based and school-based parental involvement of Jewish and Arab parents in Israel, using the parental involvement model of Hoover-Dempsey and Sandler (1995) for both populations. This study showed that Arab parents were more involved with their child's education than Jewish parents. However, the intensity of the invitation of the child to be involved ("your child has asked you to …") was higher among Jewish parents. Although these two examples did not compare different countries, they do suggest that is very likely that cultural differences in the perception of parental involvement exist. For international comparative studies in education, such as PIRLS, the Programme for International Student Assessment (PISA), or the Trends in International Mathematics and Science Study (TIMSS), this could mean that different parental perceptions of what is important for the education of their child can also have consequences for how survey questions about parental involvement are interpreted.

The next chapters report the results of two strains of analysis exploring the possible cultural differences in perceptions of parental involvement among the PIRLS countries: first by looking at the country-item interactions in scales measuring components of parental involvement, then by the variation in the relation between different parental involvement components and student achievement across countries.

References

Bakker, J., Denessen, E., & Brus-Laeven, M. (2007). Socio-economic background, parental involvement and teacher perceptions of these in relation to pupil achievement. *Educational Studies, 33*(2), 177–192. doi:10.1080/03055690601068345.

Barnard, W. M. (2004). Parent involvement in elementary school and educational attainment. *Children and Youth Services Review, 26*(1), 39–62. doi:10.1016/j.childyouth.2003.11.002.

Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro-Asencio, E., & Gaviria, J. L. (2015). Parental involvement on student academic achievement: A meta-analysis. *Educational Research Review, 14*, 33–46. doi:10.1016/j.edurev.2015.01.002.

Cooper, H., Lindsay, J. J., & Nye, B. (2000). Homework in the home: How student, family, and parenting-style differences relate to the homework process. *Contemporary Educational Psychology, 25*(4), 464–487.

Domina, T. (2005). Leveling the home advantage: Assessing the effectiveness of parental involvement in elementary school. *Sociology of Education, 78*(3), 233–249.

Driessen, G., Smit, F., & Sleegers, P. (2005). Parental involvement and educational achievement. *British Educational Research Journal, 31*(4), 509–532. doi:10.1080/01411920500148713.

Dumont, H., Trautwein, U., Lüdtke, O., Neumann, M., Niggli, A., & Schnyder, I. (2012). Does parental homework involvement mediate the relationship between family background and educational outcomes? *Contemporary Educational Psychology, 37*(1), 55–69. doi:10.1016/j.cedpsych.2011.09.004.

Epstein, J. L. (1992). School and family partnerships. In M. Aikan (Ed.), *Encyclopedia of educational research* (6th ed., pp. 1139–1151). New York: MacMillan.

Erion, J. (2006). Parent tutoring: A meta-analysis. *Education and Treatment of Children, 29*(1), 79–106.

Fan, X. T., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13*(1), 1–22.

Galindo, C., & Sheldon, S. B. (2012). School and home connections and children's kindergarten achievement gains: The mediating role of family involvement. *Early Childhood Research Quarterly, 27*(1), 90–103. doi:10.1016/j.ecresq.2011.05.004.

Grolnick, W. S., & Ryan, R. M. (1989). Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology, 81*(2), 143–154. doi:10.1037/0022-0663.81.2.143.

Hill, N. E., & Taylor, L. C. (2004). Parental school involvement and children's academic achievement pragmatics and issues. *Current Directions in Psychological Science, 13*(4), 161–164. doi:10.1111/j.0963-7214.2004.00298.x.

Hill, N. E., & Tyson, D. F. (2009). Parental involvement in middle school: A meta-analytic assessment of the strategies that promote achievement. *Developmental Psychology, 45*(3), 740–763. doi:10.1037/A0015362.

Hoover-Dempsey, K. V., & Sandler, H. M. (1995). Parental involvement in children's education: Why does it make a difference? *Teachers College Record, 97*(2), 310–331.

Hoover-Dempsey, K. V., & Sandler, H. M. (1997). Why do parents become involved in their children's education? *Review of Educational Research, 67*(1), 3–42.

Hoover-Dempsey, K. V., Battiato, A. C., Walker, J. M. T., Reed, R. P., DeJong, J. M., & Jones, K. P. (2001). Parental involvement in homework. *Educational Psychologist, 36*(3), 195–209. doi:10.1207/S15326985EP3603_5.

Hoover-Dempsey, K. V., Walker, J. M. T., Sandler, H. M., Whetsel, D., Green, C. L., Wilkins, A. S., & Closson, K. (2005). Why do parents become involved? Research findings and implications. *The Elementary School Journal, 106*(2), 105–130. doi:10.1086/499194.

Huntsinger, C. S., & Jose, P. E. (2009). Parental involvement in children's schooling: Different meanings in different cultures. *Early Childhood Research Quarterly, 24*(4), 398–410. doi:10.1016/j.ecresq.2009.07.006.

Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban Education, 40*(3), 237–269. doi:10.1177/0042085905274540.

Kloosterman, R., Notten, N., Tolsma, J., & Kraaykamp, G. (2011). The effects of parental reading socialization and early school involvement on children's academic performance: A panel study of primary school pupils in the Netherlands. *European Sociological Review, 27*(3), 291–306. doi:10.1093/Esr/Jcq007.

Lau, E. Y. H., Li, H., & Rao, N. (2011). Parental involvement and children's readiness for school in China. *Educational Research, 53*(1), 95–113. doi:10.1080/00131881.2011.552243.

Manz, P. H., Fantuzzo, J. W., & Power, T. J. (2004). Multidimensional assessment of family involvement among urban elementary students. *Journal of School Psychology, 42*(6), 461–475. doi:10.1016/j.jsp.2004.08.002.

Mattingly, D. J., Prislin, R., McKenzie, T. L., Rodriguez, J. L., & Kayzar, B. (2002). Evaluating evaluations: The case of parent involvement programs. *Review of Educational Research, 72*(4), 549–576. doi:10.3102/00346543072004549.

McNeal, R. B. (1999). Parental involvement as social capital: Differential effectiveness on science achievement, truancy, and dropping out. *Social Forces, 78*(1), 117–144.

McNeal, R. B, Jr. (2012). Checking in or checking out? Investigating the parent involvement reactive hypothesis. *Journal of Educational Research, 105*(2), 79–89. doi:10.1080/00220671.2010.519410.

Myrberg, E., & Rosen, M. (2009). Direct and indirect effects of parents' education on reading achievement among third graders in Sweden. *British Journal of Educational Psychology, 79*, 695–711. doi:10.1348/000709909x453031.

Okpala, C. O., Okpala, A. O., & Smith, F. E. (2001). Parental involvement, instructional expenditures, family socioeconomic attributes, and student achievement. *Journal of Educational Research, 95*(2), 110–115.

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). Parent involvement in homework: A research synthesis. *Review of Educational Research, 78*(4), 1039–1101. doi:10.3102/0034654308325185.

Powell, D. R., Son, S. H., File, N., & Froiland, J. M. (2012). Changes in parent involvement across the transition from public school prekindergarten to first grade and children's academic outcomes. *Elementary School Journal, 113*(2), 276–300. doi:10.1086/667726.

Selwyn, N., Banaji, S., Hadjithoma-Garstka, C., & Clark, W. (2011). Providing a platform for parents? Exploring the nature of parental engagement with school learning platforms. *Journal of Computer Assisted learning, 27*(4), 314–323. doi:10.1111/j.1365-2729.2011.00428.x.

Sénéchal, M., & LeFevre, J. A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73*(2), 445–460.

Sénéchal, M., & Young, L. (2008). The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3: A meta-analytic review. *Review of Educational Research, 78*(4), 880–907. doi:10.3102/0034654308320319.

Stoep, J., Bakker, J., & Verhoeven, L. (2002). Parental and teacher commitment to emergent literacy development. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 249–264). Amsterdam: John Benjamins.

Stylianides, A. J., & Stylianides, G. J. (2011). A type of parental involvement with an isomorphic effect on urban children's mathematics, reading, science, and social studies achievement at kindergarten entry. *Urban Education, 46*(3), 408–425. doi:10.1177/0042085910377605.

Wilder, S. (2013). Effects of parental involvement on academic achievement: A meta-synthesis. *Educational Review, 66*(3), 377–397. doi:10.1080/00131911.2013.780009.

Xu, M., Benson, S. N. K., Mudrey-Camino, R., & Steiner, R. P. (2010). The relationship between parental involvement, self-regulated learning, and reading achievement of fifth graders: A path analysis using the ECLS-K database. *Social Psychology of Education, 13*(2), 237–269. doi:10.1007/s11218-009-9104-4.

3

Parental Involvement in PIRLS-2011

Our analysis of the PIRLS-2011 data of 41 countries was guided by the analytic framework (Table 3.1), which was based on the general framework (Table 2.1) we used to categorize the studies included in the literature review. The analytic framework matches the available indicators of parental involvement in the PIRLS-2011 data (Table 3.2).

The first dimension, i.e. home-based involvement from the perspective of parents, was split into two components or indicators: early literacy activities and help with homework. The early literacy activities component is especially well measured by the PIRLS home questionnaire. In the international reports of PIRLS-2011, early literacy activities is the only component reported as a scale, with Cronbach's alphas ranging from 0.70 (Czech Republic, Hungary, Italy and Oman) to 0.88 (Romania), indicating high reliability (Martin and Mullis 2012). Although the international report does not report the scale statistics on the items regarding parental help with homework, this component is well addressed, as a total of eight items ask about these practices.

To consider the dimension of school-based involvement and home-school communication from the parent's perspective (component 3 in the analytic framework), three relevant items were selected from the home questionnaire. The number of items for this indicator is low, but the items do seem highly relevant to this context. The student's perception of parental involvement and the school's practices on parental involvement (corresponding to components 4 and 5 in the analytic framework) are measured by five items in the student questionnaire and 15 items in the school questionnaire, respectively. Finally, we established the outcome variable and control variables for the structural multilevel model tested later (Table 3.2): the PIRLS reading literacy achievement scores, gender of the student, and, as approximates for SES, both books at home and highest level of education of the parents.

Table 3.1 Analytic framework for parental involvement for secondary analyses of PIRLS-2011

Perspective	Dimension	Components	Source	Items
Parent	*Dimension 1*	*Component 1*	Home questionnaire	Question 2: A–I
	Home-based involvement	Early literacy activities before beginning primary school		
	Dimension 1	*Component 2*	Home questionnaire	Question 9
	Home-based involvement	Help with homework		
	Dimension 2	*Component 3*	Home questionnaire	Question 10: A, B, E
	School-based involvement and home-school communication	School practices on parental involvement		
Student	*Dimension 3*	*Component 4*	Student questionnaire	Question G7[a] Question R9[a]: C
	Home-based involvement	Student perception of parental involvement		
School	*Dimension 4*	*Component 5*	School questionnaire	Question 11 Question 12: E–F
	School-based involvement and home-school communication	School practices on parental involvement		

Note The PIRLS 2011 questionnaires can be retrieved from http://timssandpirls.bc.edu/pirls2011/international-contextual-q.html (IEA & TIMSS and PIRLS International Study Center 2011)
[a]The PIRLS 2011 student questionnaire consists of multiple sections. Item G7 refers to item 7 in the general part of the questionnaire, R9 to item 9 in the reading section

The benefit of using data from a large-scale assessment study such as PIRLS is not only the richness in data resulting from achievement tests, as well as student, home, teacher and school questionnaires, but also, obviously, the large number of countries for which these data are available. For this study, we initially considered the data from 43 countries participating in PIRLS-2011; countries not meeting the required response rate or for which the average achievement was not reliably measured were excluded. However, two countries, England and the USA, did not administer the home questionnaire and were therefore not included in the scale analyses for components 1–3, using items from the home questionnaire.

Table 3.2 Overview of components and items for secondary analyses of PIRLS-2011

Component	Question	Item number per component	Item in international datasets	Number of response categories
1 Early literacy activities before beginning primary school (*home-based involvement*)	*Before your child began primary school, how often did you or someone else in your home do the following activities with him or her?*			
	Read books	1	ASBH02A	3[a]
	Tell stories	2	ASBH02B	3[a]
	Sing songs	3	ASBH02C	3[a]
	Play with alphabet toys	4	ASBH02D	3[a]
	Talk about things you had done	5	ASBH02E	3[a]
	Talk about what you had read	6	ASBH02F	3[a]
	Play word games	7	ASBH02G	3[a]
	Write letters or words	8	ASBH02H	3[a]
	Read aloud signs and labels	9	ASBH02I	3[a]
	(*Source: Home questionnaire*)			
2 Help with homework (*home-based involvement*)	*How often do you or someone else in your home do the following things with your child?*			
	Discuss my child's schoolwork with him/her	1	ASBH09A	4[b]
	Help my child with his/her schoolwork	2	ASBH09B	4[b]
	Make sure my child sets aside time to do his/her homework	3	ASBH09C	4[b]
	Ask my child what he/she learned in school	4	ASBH09D	4[b]
	Check if my child has done his/her homework	5	ASBH09E	4[b]
	Help my child practice his/her reading	6	ASBH09F	4[b]

(continued)

Table 3.2 (continued)

Component	Question	Item number per component	Item in international datasets	Number of response categories
	Help my child practice his/her math skills	7	ASBH09G	4[b]
	Talk with my child about what he/she is reading	8	ASBH09H	4[b]
	(*Source: Home questionnaire*)			
3 School practices on parental involvement, parent perspective (*home-school communication*)	*What do you think of your child's school?*			
	My child's school includes me in my child's education	1	ASBH10A	4[c]
	My child's school should make a greater effort to include me in my child's education	2	ASBH10B	4[c]
	My child's school should do better at keeping me informed of his/her progress	3	ASBH10E	4[c]
	(*Source: Home questionnaire*)			
4 Parental involvement, student perspective (*home-based involvement*)	*How often do the following things happen at home?*			
	My parents ask me what I am learning in school	1	ASBG07A	4[b]
	I talk about my schoolwork with my parents	2	ASBG07B	4[b]
	My parents make sure that I set aside time for my homework	3	ASBG07C	4[b]
	My parents check if I do my homework	4	ASBG07D	4[b]
	Do you read for any of the following reasons?			
	My parents like it when I read	5	ASBR09C	4[c]
	(*Source: Student questionnaire*)			

(continued)

Table 3.2 (continued)

Component	Question	Item number per component	Item in international datasets	Number of response categories
5 School practices on parental involvement, school perspective (*home-school communication, school-based involvement*)	*How often does your school do the following for parents concerning individual students?*			
	Inform parents about their child's learning progress	1	ACBG11AA	4[d]
	Inform parents about the behavior and well-being of their child at school	2	ACBG11AB	4[d]
	Discuss parents' concerns or wishes about their child's learning	3	ACBG11AC	4[d]
	Support individual parents in helping their child with schoolwork	4	ACBG11AD	4[d]
	How often does your school ask parents to do the following?			
	Volunteer for school projects, programs, and trips	5	ACBG11BA	4[d]
	Serve on school committees	6	ACBG11BB	4[d]
	How often does your school do the following for parents in general?			
	Inform parents about the overall academic achievement of the school	7	ACBG11CA	4[d]
	Inform parents about school accomplishments	8	ACBG11CB	4[d]

(continued)

Table 3.2 (continued)

Component	Question	Item number per component	Item in international datasets	Number of response categories
	Inform parents about the educational goals and pedagogic principles of the school	9	ACBG11CC	4[d]
	Inform parents about the rules of the school	10	ACBG11CD	4[d]
	Discuss parents' concerns or wishes about the school's organization	11	ACBG11CE	4[d]
	Provide parents with additional learning materials	12	ACBG11CF	4[d]
	Organize workshops or seminars for parents on learning or pedagogical issues	13	ACBG11CG	4[d]
	How would you characterize each of the following within your school?			
	Parental support for student achievement	14	ACBG12E	5[e]
	Parental involvement in school activities	15	ACBG12F	5[e]
	(*Source: School questionnaire*)			
Socioeconomic status (*books at home, highest level of education completed by parent*)	About how many books are there in your home?		ASBH14	5[f]
	What is the highest level of education completed by the child's father?		ASBH17A	9[g]
	What is the highest level of education completed by the child's mother?		ASBH17B	9[g]
	(*Source: Home questionnaire*)			

(continued)

Table 3.2 (continued)

Component	Question	Item number per component	Item in international datasets	Number of response categories
Gender	(*Source: Student questionnaire*)		ASBG01	2
Reading literacy	(*Source: PIRLS reading comprehension assessment*)		R11F01M to R31G14M	

Note The datasets are described in detail in Foy and Drucker (2013)
[a]Category labels are: 0 = Often, 1 = Sometimes, 2 = Never or almost never
[b]Category labels are: 0 = Every day or almost every day, 1 = Once or twice a week, 2 = Once or twice a month, 3 = Never or almost never
[c]Category labels are: 0 = Agree a lot, 1 = Agree a little, 2 = Disagree a little, 3 = Disagree a lot
[d]Category labels are, after recoding: 0 = More than three times a year, 1 = Two to three times a year, 2 = Once a year, 3 = Never
[e]Category labels are: 0 = Very high, 1 = High, 2 = Medium, 3 = Low, 4 = Very low
[f]Category labels are: 0 = 1–10, 1 = 11–25, 2 = 26–100, 3 = 101–200, 4 = More than 200
[g]Category labels are: 0 = Did not go to school, 1 = Some ISCED level 1 or 2, 2 = ISCED level 2, 3 = ISCED level 3, 4 = ISCED level 4, 5 = ISCED level 5B, 6 = ISCED level 5A, 7 = Beyond ISCED level 5A, 8 = Not applicable

References

Foy, P., & Drucker, K. T. (Eds.). (2013). *PIRLS 2011 user guide for the international database*. Chestnut Hill, MA, USA: TIMSS & PIRLS International Study Center, Boston College.

IEA, & TIMSS and PIRLS International Study Center (2011). TIMSS and PIRLS press release 2011. PIRLS 2011 contextual questionnaires. Retrieved from http://timssandpirls.bc.edu/pirls2011/international-contextual-q.html.

Martin, M. O., & Mullis, I. V. S. (Eds.). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA, USA: TIMSS & PIRLS International Study Center, Boston College.

Modeling Parental Involvement

In this chapter, we first outline the models used and the estimation and testing procedures employed, and then summarize the results revealed by these models.

4.1 Estimation and Testing Procedures

The procedures we used for parameter estimation and evaluation of model fit are based on marginal maximum likelihood (MML). Most of the procedures we discuss are documented in more detail elsewhere (see Bock and Aitkin 1981; Bock et al. 1988; Gibbons and Hedeker 1992; Glas 1999; Adams and Wu 2006; De Jong et al. 2007; Jennrich and Bentler 2011; Glas and Jehangir 2014). We used the public domain software package MIRT (Glas 2010) in the calculations. Additional estimation and testing procedures were used for the bi-factor model, with unidimensional models as special cases, and random item parameters as a generalization.

4.1.1 MML Estimation

The bi-factor model used in this study was in two parts: a measurement model (i.e., an IRT model) and a structural model. The measurement model pertains to a polytomously-scored response of a student n to an item i. The possible item scores range from 0 to m_i and the score of student n on item i is denoted by the variables x_{nij} ($j = 1, \ldots, m_i$) where $x_{nij} = 1$ if the response is in category 1 and zero otherwise. Note that m_i has an index i, which indicates that the maximum score of items can differ.

We describe the procedure for the bi-factor model, combined with the partial credit model (PCM; Masters 1982) and generalized partial credit model (GPCM; Muraki 1992) as IRT models, since these two models were the ones we selected for the present study. However, the theory also applies to other IRT models, such as the unidimensional PCM and GPCM, the graded response model (Samejima 1969), the

sequential model (Tutz 1990), and other versions of these models with random item parameters instead of fixed item parameters.

In the bi-factor GPCM, the probability of scoring in category j ($j = 0, \ldots, m_i$) is given by

$$p_{ij}(\theta_n) = p(x_{nij} = 1|\theta_n, a, b) = \frac{\exp\left(\sum_{h=1}^{j} a_{i0}\theta_{n0} + a_{ig(n)}\theta_{ng(n)} - b_{ih}\right)}{1 + \sum_{k=1}^{m_i} \exp\left(\sum_{h=1}^{k} a_{i0}\theta_{n0} + a_{ig(n)}\theta_{ng(n)} - b_{ik}\right)} \tag{4.1}$$

where, θ_{n0} is the score of a student n on the latent scale pertaining to all countries, $\theta_{ng(n)}$ is the score on a country specific latent dimension, and the index $g(n)$ indicates the country to which student n belongs. Further, a_{i0} and $a_{ig(n)}$ are the factor loadings of item i on these two dimensions, and b_{ih} ($h = 1, \ldots, m_i$) is the item location parameter. The location parameter b_{ih} is the position on the latent scale, where it is assumed that summations such as $h = 1$ to 0 result in zero. The unidimensional GPCM lacks the country-specific dimensions $\theta_{ng(n)}$ and the associated factor loadings $a_{ig(n)}$. Further, the PCM is obtained by fixing all item parameters a_{i0} to one.

The formula for the response probability and subsequent derivations can be simplified by introducing the re-parametrization $d_{ij} = \Sigma_{h=1}^{j} b_{ih}$ and by defining $a_{ig}^t\theta_n$ as the inner product of the vectors $(a_{i0}, a_{ig(n)})$ and $(\theta_{n0}, \theta_{ng(n)})$, respectively. Thus, Eq. (4.1) becomes

$$p_{ij}(\theta_n) = \frac{\exp\left(ja_{ig}^t\theta_n - d_{ij}\right)}{1 + \sum_{k=1}^{m_i} \exp\left(ka_{ig}^t\theta_n - d_{ik}\right)} \tag{4.2}$$

The θ_0-dimension is the general dimension that pertains to all countries and is the basis for the comparison of the countries. The θ_g-dimensions are the country-specific dimensions, and the factor loadings on these dimensions give an indication of country-by-item interaction. It is assumed that within each country, the dimensions θ_0 and θ_g have a bi-variate normal distribution $N(\theta_{n0}, \theta_{ng}; \mu_g, \Sigma_g)$. For the two-dimensional country mean $\mu_g = (\mu_{g0}, \mu_g)$, it holds that the mean on the second dimension is fixed at zero, that is $\mu_g = 0$. The covariance matrix is given by

$$\Sigma_g = \begin{bmatrix} \sigma_g^2 & 0 \\ 0 & 1 \end{bmatrix}$$

In the unidimensional GPCM and PCM, the latent student parameters θ_0 have a univariate normal distribution with a mean μ_g and a variance σ_g^2. Finally, random item parameters are obtained by introducing independent multivariate normal

distributions on the parameters for each item (for further details, please consult De Jong et al. 2007).

The present application of the bi-factor model is not standard, but an extension of the basic model. Thus, the technical details on the estimation equations, expressions for the covariance matrix of the estimates, and tests of model fit, are also provided (see Appendix A).

4.1.2 *Detection and Modeling of Differential Item Functioning*

Part of the process of establishing the construct validity of a scale may consist of showing that the scale fits an IRT model. In the present study, the focus is on country-specific CDIF. CDIF can be detected using Lagrange multiplier (LM) test statistics (Rao 1947; see also, Aitchison and Silvey 1958) and CDIF can be modeled using country-specific item parameters. Glas and Jehangir (2014) already showed the feasibility of the method using PISA data, although in the slightly simpler framework of one-dimensional IRT models. The method is implemented in the public domain software package MIRT (Glas 2010). LM tests have been previously applied to IRT frameworks (Glas 1999; Glas and Falcón 2003; Glas and Dagohoy 2007). Our primary interest is not in the actual outcome of the LM test, because due to the very large sample sizes in educational surveys even the smallest model violation, that is, the smallest amount of differential item functioning (DIF), will be significant. The reason for adopting the framework of the LM test is that it clarifies the connection between the model violations, and observations and expectations used to detect DIF. Further, because it produces comprehensible and well-founded expressions for model expectations, the value of the LM test statistic can be used as measure of the effect size of DIF, and the procedure can be easily generalized to a broad class of IRT models.

To define the test and the associated residuals, we define a background variable

$$y_{nc} = \begin{cases} 1 & \text{if person } n \text{ belongs to country } c, \\ 0 & \text{if person } n \text{ does not belong to country } c. \end{cases}$$

The LM test targets the null-hypothesis of no DIF, namely the null-hypothesis where $\delta_i = 0$. The LM test statistic is computed using the MML estimates of the null-model, where δ_i is not estimated. The test is based on evaluation of the first-order derivatives of the marginal likelihood with respect to δ_i evaluated at $\delta_i = 0$ (see Glas 1999). If the first-order derivative in this point is large, the MML estimate of δ_i is far removed from zero, and the test is significant. If the first-order derivative in this point is small, the MML estimate of δ_i is probably close to zero and the test is not significant. The actual LM statistic is the squared first-order derivative divided by its estimated variance, and it has an asymptotic chi-squared

distribution with one degree of freedom. However, as already discussed, the primary interest is not so much in the test itself, but in the information it provides regarding the fit between the data and the model.

For a general definition of the approach, which also pertains to polytomously-scored items, the covariates y_{nc} (c = 1, …, C) should be defined. Special cases leading to specific DIF statistics are given later. The covariates may be separately observed person characteristics, but they may also depend on the observed response pattern, but without the response to the item i targeted.

The LM approach can be outlined using the bi-factor GPCM; the special cases for the unidimensional PCM and GPCM are obtained if the restrictions denoted above are invoked. The probability of a response is given by a generalization of the bi-factor GPCM, namely,

$$p_{ij}(\theta_n) = \frac{\exp\left(ja_{ig}^t\theta_n - d_{ij} + j\sum_c y_{nc}\delta_{ic}\right)}{1 + \sum_{k=1}^{m_i}\exp\left(ka_{ig}^t\theta_n - d_{ik} + k\sum_c y_{nc}\delta_{ic}\right)}$$

For one so-called reference country, the covariate y_{nv} is equal to zero. This country serves as a baseline where the bi-factor GPCM with item parameters a and b holds. In the other C-1 countries, the covariates y_{nv} are equal to one. It can be shown (see Glas 1999) that the test statistic is based on the residuals

$$\frac{\sum_{n=1}^{N}\sum_{j=1}^{m_i} y_{nc} jX_{ij}}{\sum_{n=1}^{N} y_{nc}} - \frac{\sum_{n=1}^{N}\sum_{j=1}^{m_i} y_{nc} jE\left(P_{ij}(\theta_n)|x_n;\lambda\right)}{\sum_{n=1}^{N} y_{nc}} \tag{4.3}$$

for c = 1, …, C-1. Dividing this residual by the number of respondents $\Sigma_n y_{nc}$ produces residuals that are the differences between the observed and expected average item-total score in country c = 1, …, C-1. The residual gauges so-called uniform DIF, in other words, the residual indicates whether the item total function (ITF) $\Sigma_j\, jP_{ij}(\theta)$ is shifted for the item, namely whether there is item-by-country interaction.

The LM statistic for the null-hypothesis $\delta_i = 0$ (c = 1, …, C-1) is a quadratic form in the (C-1)-dimensional vector of residuals and the inverse of their covariance matrix (for details, see Glas 1999). It has an asymptotic chi-squared distribution with C-1 degrees of freedom.

A special case of this procedure is obtained if one country serves as the focal country and all other countries serve as reference. Then the model under the alternative hypothesis has only one additional parameter, δ_i, and the associated LM statistic has an asymptotic chi-squared distribution with one degree of freedom.

Items that show the worst misfit, based on their value of the LM statistic and residuals, are given country-specific item parameters. From a practical point of

view, defining country-specific item parameters is equivalent to defining an incomplete design where the DIF item is split into a number of virtual items, and where each virtual item is considered as administered in a specific country. The resulting design can be analyzed using IRT software that supports the analysis of data collected in an incomplete design. We here refer to items with country-specific parameters as split items.

The method is motivated by the assumption that a substantial part of the items function the same in all countries and a limited number of items have CDIF. In the IRT model, it is assumed that all items pertain to the same latent variable θ. Items without CDIF have the same item parameters in every country. However, items with CDIF have item parameters that differ across countries. These items refer to the same latent variable θ as all the other items, but their location on the scale differs across countries. For instance, the number of cars in the family may be a good indicator of wealth, but the actual number of cars at a certain level of wealth may vary across countries, or even within countries. Having a car in the inner city of Amsterdam is clearly a sign of wealth, but, in the rural eastern part of the Netherlands, an equivalent level of wealth would probably result in the ownership of three cars.

The number of items given country-specific item parameters is a matter of choice where two considerations are relevant. First, there should remain a sufficient number of anchor items in the scale. Second, the model including the split items should fit the data. DIF statistics no longer apply to the split items. However, the fit of the item response curve of an individual item, say item *i,* can be evaluated using the test for non-uniform DIF described earlier, but using a model including country-specific items parameters. So, in this application too, test-score ranges are used as proxies for locations on the θ scale, and the test evaluates whether the model with the country-specific item parameters can properly predict the ITF.

4.2 Results of Modeling Country-Specific Differential Item Functioning

We here provide descriptive statistics at country level for each of the five parental involvement components under the PCM and GPCM, including sample size and estimated global reliability (Tables 4.1, 4.2, 4.3, 4.4 and 4.5). Sample sizes for the first four components (early literacy activities, help with homework, school practices on parental involvement from a parental perspective, and parental involvement from a student perspective) were taken from the PIRLS home and student data, providing a significantly larger sample than that available for the last component (school practices on parental involvement, school perspective), where data were derived from the PIRLS school questionnaire. The GPCM rarely improved global reliability. Components 1 (early literacy activities), 2 (help with homework), and 5 (school practices on parental involvement, school perspective) were evaluated using nine, eight, and 15 items, respectively (see also Table 3.2). Their global reliability is

Table 4.1 Country characteristics component 1: early literacy activities before beginning primary school

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Azerbaijan, Republic of	4509	6.56	0.44	1.05	0.74	0.36	0.98	0.74
Australia	3232	4.46	−0.55	1.30	0.77	−0.49	1.19	0.77
Austria	4393	5.90	0.10	1.01	0.73	0.08	0.94	0.74
Belgium (French)	3383	6.46	0.30	1.01	0.74	0.29	0.94	0.74
Bulgaria	5137	6.10	0.12	1.57	0.84	0.12	1.46	0.85
Canada	18848	4.57	−0.49	1.25	0.76	−0.44	1.14	0.76
Chinese Taipei	4242	8.41	0.98	1.11	0.78	0.90	1.03	0.78
Colombia	3798	5.79	0.11	1.19	0.77	0.13	1.10	0.77
Croatia	4539	4.62	−0.38	0.97	0.69	−0.35	0.90	0.69
Czech Republic	4397	5.28	−0.10	0.90	0.68	−0.09	0.84	0.69
Denmark	4322	6.10	0.18	0.96	0.72	0.18	0.90	0.73
Finland	4423	6.23	0.24	0.80	0.65	0.24	0.74	0.65
France	4111	5.94	0.12	1.02	0.74	0.11	0.95	0.74
Georgia	4640	4.46	−0.44	1.11	0.72	−0.44	1.02	0.72
Germany	3197	5.56	−0.01	0.96	0.71	−0.02	0.89	0.71
Hong Kong, SAR	3604	8.45	1.01	0.97	0.73	0.91	0.90	0.74
Hungary	4912	5.27	−0.11	0.92	0.69	−0.12	0.85	0.69
Indonesia	4588	6.90	0.48	1.02	0.74	0.45	0.94	0.75
Iran, Islamic Republic of	5653	7.82	0.82	1.06	0.76	0.75	0.99	0.76
Ireland	4268	4.58	−0.47	1.24	0.76	−0.43	1.14	0.76
Israel	3261	4.81	−0.33	1.11	0.74	−0.30	1.03	0.74
Italy	3873	4.97	−0.23	1.00	0.71	−0.20	0.93	0.71
Lithuania	4406	5.67	0.04	0.96	0.71	0.01	0.90	0.71
Malta	3274	5.24	−0.18	1.14	0.76	−0.17	1.06	0.76
Netherlands	2273	5.53	−0.03	0.96	0.71	−0.02	0.89	0.71
New Zealand	3357	4.37	−0.60	1.33	0.77	−0.54	1.22	0.77
Norway	2909	5.76	0.06	0.97	0.71	0.08	0.90	0.72
Northern Ireland	2107	4.02	−0.74	1.28	0.75	−0.68	1.18	0.75
Poland	4920	5.06	−0.20	0.99	0.71	−0.20	0.92	0.71
Portugal	3887	5.76	0.05	1.09	0.75	0.04	1.01	0.76
Qatar	3650	6.49	0.35	1.08	0.75	0.30	1.00	0.76
Romania	4535	5.59	−0.12	1.57	0.83	−0.12	1.46	0.84
Russian Federation	4412	4.02	−0.70	1.19	0.73	−0.68	1.09	0.73
Saudi Arabia	4369	6.52	0.36	1.04	0.75	0.36	0.97	0.75
Singapore	6194	7.16	0.51	1.24	0.80	0.47	1.15	0.81
Slovak Republic	5481	5.02	−0.24	1.08	0.74	−0.23	1.00	0.74
Slovenia	4313	4.78	−0.33	1.02	0.71	−0.31	0.94	0.71
Spain	7945	5.13	−0.18	1.03	0.72	−0.16	0.95	0.73

(continued)

Table 4.1 (continued)

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Sweden	4013	6.06	0.15	1.03	0.74	0.15	0.96	0.75
Trinidad and Tobago	3497	4.85	−0.33	1.17	0.75	−0.29	1.08	0.76
United Arab Emirates	13305	6.52	0.35	1.03	0.74	0.32	0.96	0.75

Note N is the sample size, and $\bar{X}$ the observed mean score on the component. $\mu(\theta)$ is the estimated mean, $\sigma(\theta)$ is the standard deviation, and ρ is the estimated global reliability under the partial credit model (PCM) or the generalized partial credit model (GPCM)

Table 4.2 Country characteristics component 2: help with homework

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Azerbaijan, Republic of	4541	2.99	−0.95	2.02	0.76	−0.63	1.30	0.76
Australia	3234	5.27	0.53	1.23	0.79	0.33	0.80	0.80
Austria	4430	6.26	0.83	1.22	0.81	0.57	0.81	0.82
Belgium (French)	3356	3.58	−0.44	1.74	0.78	−0.30	1.16	0.79
Bulgaria	5126	4.82	−0.22	2.28	0.83	−0.13	1.50	0.84
Canada	18844	3.99	−0.04	1.41	0.77	−0.02	0.92	0.78
Chinese Taipei	4244	5.73	0.53	1.52	0.83	0.33	1.00	0.84
Colombia	3824	3.03	−0.72	1.74	0.75	−0.46	1.12	0.76
Croatia	4532	5.08	0.44	1.28	0.79	0.32	0.88	0.82
Czech Republic	4418	4.42	0.30	1.10	0.73	0.22	0.74	0.76
Denmark	4303	5.32	0.54	1.23	0.79	0.36	0.82	0.80
Finland	4410	8.31	1.45	0.92	0.78	0.96	0.64	0.80
France	4115	3.63	−0.23	1.48	0.76	−0.15	0.99	0.78
Georgia	4622	3.05	−0.83	1.90	0.76	−0.53	1.23	0.77
Germany	3195	6.05	0.72	1.33	0.82	0.49	0.90	0.84
Hong Kong, SAR	3609	5.94	0.49	1.70	0.85	0.28	1.13	0.85
Hungary	4903	3.91	−0.26	1.71	0.79	−0.15	1.13	0.80
Indonesia	4577	3.99	−0.23	1.70	0.79	−0.21	1.10	0.79
Iran, Islamic Republic of	5650	4.68	0.16	1.53	0.80	0.07	1.01	0.81
Ireland	4268	2.99	−0.69	1.68	0.75	−0.46	1.11	0.76
Israel	3271	5.84	0.63	1.38	0.82	0.43	0.92	0.83
Italy	3867	3.78	−0.22	1.57	0.78	−0.12	1.05	0.80
Lithuania	4395	5.49	0.53	1.35	0.81	0.35	0.92	0.83
Malta	3285	4.23	0.06	1.41	0.78	0.04	0.93	0.79
Netherlands	2280	9.36	1.63	1.10	0.83	1.09	0.76	0.85
New Zealand	3351	5.28	0.43	1.44	0.81	0.26	0.94	0.82
Norway	2105	2.40	−0.82	1.41	0.69	−0.54	0.93	0.70
Northern Ireland	2908	3.56	−0.20	1.39	0.75	−0.15	0.91	0.76

(continued)

Table 4.2 (continued)

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Poland	4923	3.82	−0.10	1.40	0.76	−0.05	0.94	0.78
Portugal	3889	3.82	−0.32	1.72	0.79	−0.23	1.14	0.80
Qatar	3653	3.20	−0.54	1.63	0.76	−0.35	1.04	0.76
Romania	4533	3.71	−0.81	2.32	0.80	−0.50	1.52	0.81
Russian Federation	4417	3.39	−0.40	1.56	0.76	−0.28	1.01	0.77
Saudi Arabia	4256	3.79	−0.33	1.74	0.79	−0.23	1.12	0.79
Singapore	6190	5.83	0.56	1.51	0.83	0.33	0.99	0.84
Slovak Republic	5489	4.99	0.31	1.47	0.81	0.22	0.99	0.83
Slovenia	4340	4.78	0.25	1.42	0.80	0.18	0.96	0.82
Spain	7945	3.15	−0.67	1.76	0.76	−0.43	1.17	0.77
Sweden	3985	4.78	0.31	1.34	0.79	0.19	0.88	0.80
Trinidad and Tobago	3499	2.41	−1.09	1.76	0.72	−0.68	1.12	0.73
United Arab Emirates	13287	3.12	−0.61	1.67	0.76	−0.40	1.07	0.76

Note N is the sample size, and $\bar{X}$ the observed mean score on the component. $\mu(\theta)$ is the estimated mean, $\sigma(\theta)$ is the standard deviation, and ρ is the estimated global reliability under the partial credit model (PCM) or the generalized partial credit model (GPCM)

Table 4.3 Country characteristics component 3: school practices on parental involvement, parent perspective

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Azerbaijan, Republic of	4401	0.79	−2.02	1.40	0.47	−15.97	11.11	0.51
Australia	3185	3.51	0.39	0.13	0.04	3.12	1.04	0.36
Austria	4349	4.11	0.63	0.19	0.10	4.96	1.49	0.49
Belgium (French)	3269	4.14	0.63	0.13	0.04	0.63	1.00	0.35
Bulgaria	5029	2.70	−0.01	0.58	0.42	−0.07	4.62	0.67
Canada	18567	3.66	0.45	0.16	0.07	3.58	1.29	0.43
Chinese Taipei	4189	1.85	−0.35	0.23	0.10	−2.77	1.84	0.54
Colombia	3738	1.31	−1.31	1.31	0.55	−10.37	10.33	0.62
Croatia	4478	3.05	0.18	0.46	0.35	1.43	3.65	0.65
Czech Republic	4316	3.68	0.46	0.15	0.06	3.65	1.16	0.40
Denmark	4243	4.03	0.58	0.13	0.04	4.60	1.01	0.37
Finland	4348	4.53	0.73	0.10	0.02	5.77	0.79	0.28
France	3961	3.86	0.52	0.10	0.02	4.14	0.81	0.27
Georgia	4483	1.63	−0.80	0.96	0.51	−6.31	7.58	0.64
Germany	3097	3.88	0.54	0.13	0.04	4.25	0.99	0.36
Hong Kong, SAR	3593	1.51	−0.60	0.33	0.17	−4.75	2.64	0.59
Hungary	4793	3.38	0.36	0.22	0.13	2.83	1.77	0.52

(continued)

Table 4.3 (continued)

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Indonesia	4549	0.86	−1.65	1.09	0.42	−13.02	8.61	0.57
Iran, Islamic Republic of	5608	1.34	−1.00	0.89	0.45	−7.88	7.07	0.64
Ireland	4187	3.44	0.37	0.36	0.27	2.89	2.87	0.61
Israel	3188	2.47	−0.11	0.55	0.39	−0.87	4.34	0.64
Italy	3755	3.61	0.43	0.11	0.03	3.44	0.86	0.28
Lithuania	4347	3.45	0.38	0.22	0.13	2.99	1.77	0.51
Malta	3188	2.35	−0.25	0.80	0.51	−1.99	6.34	0.66
Netherlands	2265	4.39	0.70	0.13	0.04	5.56	1.01	0.39
New Zealand	3362	3.56	0.42	0.23	0.13	3.29	1.78	0.51
Norway	2091	3.92	0.55	0.19	0.10	4.37	1.51	0.48
Northern Ireland	2884	3.49	0.38	0.11	0.03	3.04	0.89	0.29
Poland	4790	3.25	0.32	0.15	0.05	2.50	1.15	0.38
Portugal	3745	3.60	0.43	0.11	0.03	3.44	0.86	0.28
Qatar	3610	1.87	−0.39	0.40	0.25	−3.06	3.18	0.61
Romania	4477	2.00	−0.52	0.90	0.53	−4.10	7.13	0.66
Russian Federation	4331	3.25	0.31	0.19	0.09	2.44	1.48	0.47
Saudi Arabia	4306	1.39	−1.05	1.03	0.50	−8.29	8.19	0.64
Singapore	6145	2.03	−0.23	0.18	0.07	−1.83	1.42	0.45
Slovak Republic	5344	3.08	0.23	0.23	0.13	1.83	1.79	0.52
Slovenia	4246	4.00	0.55	0.09	0.02	4.32	0.74	0.23
Spain	7699	3.53	0.41	0.20	0.11	3.27	1.62	0.49
Sweden	3974	3.59	0.43	0.13	0.04	3.36	1.02	0.35
Trinidad and Tobago	3328	2.03	−0.51	0.96	0.55	−4.07	7.57	0.67
United Arab Emirates	13061	1.68	−0.59	0.62	0.37	−4.68	4.89	0.64

Note N is the sample size, and $\bar{X}$ the observed mean score on the component. $\mu(\theta)$ is the estimated mean, $\sigma(\theta)$ is the standard deviation, and ρ is the estimated global reliability under the partial credit model (PCM) or the generalized partial credit model (GPCM)

generally >0.70, which is an acceptable level for country inferences. A value of 0.80 is generally considered an acceptable reliability level for individual inferences, and for many combinations of components and countries, this level was attained. Components 3 (school practices on parental involvement, parental perspective) and 4 (parental involvement from a student perspective), were evaluated using three items and five items, respectively; the global reliability of these estimates was thus correspondingly lower.

We also investigated the item characteristics for each component (Tables 4.6, 4.7, 4.8, 4.9 and 4.10). Local reliability, namely the extent to which different θ-values can be distinguished, was assessed using the "slope" parameter. The relatively high

Table 4.4 Country characteristics component 4: student perception of parental involvement

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Azerbaijan, Republic of	4330	1.50	−0.48	0.99	0.51	−0.54	1.18	0.51
Australia	5997	3.31	0.44	0.67	0.57	0.55	0.81	0.58
Austria	4571	1.90	−0.24	0.90	0.54	−0.29	1.09	0.55
Belgium (French)	3680	2.11	−0.12	0.87	0.55	−0.17	1.09	0.57
Bulgaria	5191	2.36	−0.24	1.18	0.64	−0.29	1.44	0.64
Canada	22750	2.46	0.08	0.79	0.56	0.13	0.96	0.57
Chinese Taipei	4276	4.36	0.65	0.84	0.69	0.79	1.04	0.70
Colombia	3793	1.42	−0.74	1.17	0.53	−0.88	1.40	0.53
Croatia	4564	2.08	−0.04	0.74	0.50	−0.06	0.88	0.50
Czech Republic	4483	1.38	−0.52	0.87	0.45	−0.62	1.06	0.47
Denmark	4543	2.58	0.21	0.66	0.51	0.25	0.80	0.52
England	3912	3.30	0.47	0.61	0.53	0.54	0.73	0.54
Finland	4599	3.57	0.55	0.58	0.53	0.67	0.70	0.55
France	4403	2.31	0.01	0.80	0.55	−0.02	1.01	0.57
Georgia	4581	1.56	−0.53	1.05	0.53	−0.62	1.25	0.53
Germany	3600	1.90	−0.16	0.80	0.50	−0.13	0.97	0.53
Hong Kong, SAR	3826	5.34	0.93	0.70	0.67	1.10	0.88	0.68
Hungary	5105	1.95	−0.23	0.91	0.54	−0.31	1.09	0.54
Indonesia	4662	2.32	−0.04	0.90	0.58	−0.01	1.10	0.60
Iran, Islamic Republic of	5727	2.14	−0.07	0.83	0.54	−0.08	0.98	0.54
Ireland	4415	2.27	0.00	0.80	0.54	0.03	0.97	0.56
Israel	4117	2.46	0.06	0.83	0.57	0.08	1.00	0.58
Italy	4100	2.17	−0.01	0.76	0.52	−0.07	0.96	0.54
Lithuania	4591	2.17	−0.02	0.77	0.52	−0.04	0.93	0.53
Malta	3519	2.53	0.09	0.81	0.57	0.14	0.98	0.59
Netherlands	3955	3.56	0.48	0.72	0.61	0.52	0.89	0.61
New Zealand	5549	3.03	0.36	0.67	0.55	0.44	0.80	0.56
Northern Ireland	3523	2.37	0.16	0.61	0.46	0.17	0.75	0.47
Norway	3112	2.50	0.19	0.65	0.49	0.29	0.80	0.53
Poland	4953	2.20	−0.05	0.82	0.54	−0.10	1.01	0.55
Portugal	4037	1.91	−0.19	0.83	0.52	−0.26	1.02	0.53
Qatar	3947	2.82	0.17	0.89	0.62	0.19	1.08	0.62
Romania	4592	1.69	−0.57	1.16	0.57	−0.71	1.39	0.56
Russian Federation	4444	1.82	−0.25	0.86	0.51	−0.31	1.05	0.53
Saudi Arabia	4425	2.55	0.06	0.90	0.60	0.07	1.08	0.61
Singapore	6275	4.25	0.66	0.74	0.65	0.77	0.92	0.66
Slovak Republic	5586	1.76	−0.45	1.06	0.56	−0.57	1.29	0.56
Slovenia	4456	2.13	−0.02	0.75	0.51	−0.02	0.91	0.52
Spain	8501	2.07	−0.15	0.88	0.55	−0.18	1.07	0.56

(continued)

Table 4.4 (continued)

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Sweden	4533	2.45	0.18	0.64	0.49	0.23	0.79	0.51
Trinidad and Tobago	3875	1.52	−0.65	1.12	0.54	−0.75	1.34	0.55
United Arab Emirates	14209	2.23	−0.11	0.94	0.58	−0.14	1.14	0.59
United States	12501	2.72	0.15	0.84	0.60	0.20	1.02	0.61

Note N is the sample size, and $\bar{X}$ the observed mean score on the component. $\mu(\theta)$ is the estimated mean, $\sigma(\theta)$ is the standard deviation, and ρ is the estimated global reliability under the partial credit model (PCM) or the generalized partial credit model (GPCM)

Table 4.5 Country characteristics component 5: school practices on parental involvement, school perspective

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Azerbaijan, Republic of	169	32.89	0.32	0.55	0.70	0.59	0.90	0.73
Australia	269	35.41	0.64	0.64	0.74	1.04	1.06	0.76
Austria	158	31.00	0.05	0.47	0.64	−0.14	0.81	0.71
Belgium (French)	118	23.37	−0.80	0.59	0.73	−1.13	0.92	0.74
Bulgaria	147	30.14	−0.04	0.70	0.79	0.10	1.10	0.81
Canada	1084	33.37	0.38	0.67	0.76	0.56	1.09	0.79
Chinese Taipei	150	34.35	0.54	0.88	0.83	0.77	1.48	0.85
Colombia	149	32.73	0.35	0.77	0.81	0.87	1.29	0.81
Croatia	152	30.50	−0.02	0.46	0.63	0.01	0.74	0.68
Czech Republic	174	28.28	−0.29	0.46	0.64	−0.30	0.81	0.72
Denmark	231	25.93	−0.54	0.42	0.60	−1.03	0.76	0.67
England	120	32.83	0.29	0.57	0.71	0.30	0.89	0.74
Finland	139	25.60	−0.59	0.50	0.68	−1.01	0.85	0.72
France	167	27.52	−0.35	0.57	0.73	−0.60	0.99	0.78
Georgia	171	30.85	0.07	0.69	0.79	0.24	1.17	0.82
Germany	187	30.16	−0.05	0.50	0.68	−0.19	0.82	0.72
Hong Kong, SAR	125	30.18	−0.04	0.63	0.76	−0.31	1.05	0.80
Hungary	143	29.06	−0.18	0.52	0.69	−0.14	0.85	0.73
Indonesia	155	27.53	−0.37	0.75	0.82	−0.60	1.21	0.84
Iran, Islamic Republic of	244	32.60	0.30	0.88	0.85	0.65	1.46	0.85
Ireland	145	27.75	−0.32	0.62	0.76	−0.75	1.07	0.81
Israel	132	32.24	0.25	0.71	0.79	0.31	1.12	0.81
Italy	200	27.80	−0.26	0.65	0.77	−0.38	1.08	0.80
Lithuania	151	30.42	−0.02	0.49	0.67	−0.07	0.83	0.72
Malta	93	30.99	0.09	0.59	0.73	0.02	0.89	0.75
Netherlands	117	26.97	−0.42	0.43	0.61	−0.80	0.77	0.69
New Zealand	175	34.13	0.47	0.57	0.70	0.56	0.88	0.73

(continued)

Table 4.5 (continued)

Country	N	$\bar{X}$	PCM			GPCM		
			$\mu(\theta)$	$\sigma(\theta)$	ρ	$\mu(\theta)$	$\sigma(\theta)$	ρ
Northern Ireland	117	29.23	−0.17	0.55	0.71	−0.53	0.91	0.76
Norway	115	26.03	−0.54	0.39	0.56	−0.93	0.66	0.62
Poland	148	31.57	0.15	0.53	0.69	0.22	0.83	0.71
Portugal	147	29.62	−0.10	0.62	0.76	−0.16	0.98	0.78
Qatar	166	33.86	0.53	0.96	0.85	1.02	1.59	0.85
Romania	147	32.91	0.34	0.71	0.78	0.76	1.21	0.81
Russian Federation	202	34.42	0.46	0.46	0.62	0.69	0.82	0.70
Saudi Arabia	169	26.57	−0.48	0.83	0.85	−0.55	1.43	0.88
Singapore	176	32.40	0.25	0.60	0.73	0.31	1.05	0.79
Slovak Republic	194	28.70	−0.22	0.56	0.72	−0.20	0.97	0.77
Slovenia	191	29.25	−0.14	0.47	0.65	−0.21	0.81	0.71
Spain	302	28.96	−0.18	0.65	0.78	−0.30	1.08	0.81
Sweden	132	27.33	−0.38	0.55	0.71	−0.57	0.89	0.75
Trinidad and Tobago	147	30.84	0.07	0.78	0.82	0.43	1.34	0.85
United Arab Emirates	419	32.42	0.29	0.80	0.82	0.45	1.25	0.83
United States	331	35.36	0.66	0.73	0.78	1.01	1.29	0.81

Note N is the sample size, and $\bar{X}$ the observed mean score on the component. $\mu(\theta)$ is the estimated mean, $\sigma(\theta)$ is the standard deviation, and ρ is the estimated global reliability under the partial credit model (PCM) or the generalized partial credit model (GPCM)

Table 4.6 Response frequencies and item parameter estimates under the generalized partial credit model for items in component 1: early literacy activities

Item	Slope	Intercept	$I(0)$	Relative frequency response categories		
				Cat0	Cat1	Cat2
ASBH02A	1.26	1.84	0.44	0.54	0.41	0.05
ASBH02B	1.24	1.47	0.46	0.48	0.46	0.07
ASBH02C	0.77	0.98	0.23	0.49	0.41	0.11
ASBH02D	1.09	0.85	0.45	0.43	0.44	0.14
ASBH02E	0.95	1.80	0.24	0.62	0.34	0.04
ASBH02F	1.18	0.82	0.45	0.36	0.52	0.12
ASBH02G	1.24	0.57	0.52	0.33	0.51	0.16
ASBH02H	1.06	1.12	0.38	0.47	0.44	0.10
ASBH02I	1.07	0.89	0.43	0.44	0.42	0.13

Note The latent distributions of the countries are normed to an overall mean of zero. Slope and intercept are the parameters ai0 and the mean of the location parameters bi1, bi2, …, bih, etc., respectively. $I(0)$ is the information value of the item at $\theta = 0$. Cat0, Cat1, Cat2 indicate the frequency with which item categories 0, 1 and 2 are endorsed, respectively. The components, items and corresponding category labels are described in Table 3.2

Table 4.7 Response frequencies and item parameter estimates under the generalized partial credit model for items in component 2: help with homework

Item	Slope	Intercept	$I(0)$	Relative frequency response categories			
				Cat0	Cat1	Cat2	Cat3
ASBH09A	1.17	2.44	0.21	0.78	0.18	0.03	0.02
ASBH09B	1.63	2.01	0.78	0.56	0.32	0.07	0.05
ASBH09C	1.15	2.15	0.18	0.81	0.14	0.03	0.03
ASBH09D	1.10	2.27	0.24	0.73	0.22	0.03	0.02
ASBH09E	1.56	2.51	0.31	0.77	0.17	0.03	0.04
ASBH09F	1.69	1.09	1.25	0.43	0.33	0.10	0.14
ASBH09G	2.26	1.87	1.62	0.43	0.37	0.12	0.08
ASBH09H	1.45	1.66	0.77	0.47	0.38	0.11	0.04

Note The latent distributions of the countries are normed to an overall mean of zero. Slope and intercept are the parameters ai0 and the mean of the location parameters bi1, bi2, …, bih, etc., respectively. $I(0)$ is the information value of the item at $\theta = 0$. Cat0, Cat1, Cat2, and Cat3 indicate the frequency with which item categories 0, 1, 2 and 3 are endorsed, respectively. The content of the components, items and corresponding category labels are described in Table 3.2

Table 4.8 Response frequencies and item parameter estimates under the generalized partial credit model for items in component 3: school practices on parental involvement, parent perspective

Item	Slope	Intercept	$I(0)$	Relative frequency response categories			
				Cat0	Cat1	Cat2	Cat3
ASBH10A	0.61	1.41	0.15	0.54	0.37	0.07	0.02
ASBH10B	0.61	0.36	0.34	0.30	0.31	0.23	0.16
ASBH10E	0.58	0.52	0.30	0.38	0.29	0.19	0.14

Note The latent distributions of the countries are normed to an overall mean of zero. Slope and intercept are the parameters ai0 and the mean of the location parameters bi1, bi2, …, bih, etc., respectively. $I(0)$ is the information value of the item at $\theta = 0$. Cat0, Cat1, Cat2, and Cat3 indicate the frequency with which item categories 0, 1, 2 and 3 are endorsed, respectively. The components, items and corresponding category labels are described in Table 3.2

value for PIRLS item ASBH02A ("read books"), indicates that this item of the scale performed best in this respect. Local reliability is further supported if the item location parameters agree closely with the mean of a latent distribution. In this respect, item ASBH02G ("play word games") performed best, because the latent distributions of the countries were normed to an overall mean of zero. Together the intercept and slope parameters determine the information value of an item. Higher values for the information value of an item at $\theta = 0$, namely $I(0)$, indicate the item made a higher contribution to the local reliability of the component.

For component 1 (early literacy activities), the item ASBH02C ("sing songs") has a lower information value than the other items. This should be taken into

Table 4.9 Response frequencies and item parameter estimates under the generalized partial credit model for items in component 4: student perception of parental involvement

Item	Slope	Intercept	$I(0)$	Relative frequency response categories			
				Cat0	Cat1	Cat2	Cat3
ASBG07A	1.01	1.47	0.32	0.67	0.21	0.05	0.07
ASBG07B	0.96	1.15	0.43	0.56	0.27	0.08	0.09
ASBG07C	0.85	1.35	0.22	0.75	0.14	0.04	0.08
ASBG07D	0.77	1.21	0.22	0.73	0.14	0.04	0.09
ASBR09C	0.55	1.55	0.09	0.76	0.18	0.04	0.02

Note The latent distributions of the countries are normed to an overall mean of zero. Slope and intercept are the parameters ai0 and the mean of the location parameters bi1, bi2, …, bih, etc., respectively. $I(0)$ is the information value of the item at $\theta = 0$. Cat0, Cat1, Cat2, and Cat3 indicate the frequency with which item categories 0, 1, 2 and 3 are endorsed, respectively. The components, items and corresponding category labels are described in Table 3.2

Table 4.10 Response frequencies and item parameter estimates under the generalized partial credit model for items in component 5: school practices on parental involvement, school perspective

Item	Slope	Intercept	$I(0)$	Relative frequency response categories				
				Cat0	Cat1	Cat2	Cat3	Cat4
ACBG11AA	0.75	−2.88	0.14	0.00	0.01	0.37	0.62	–
ACBG11AB	0.91	−2.95	0.20	0.00	0.02	0.33	0.65	–
ACBG11AC	0.87	−2.34	0.23	0.00	0.05	0.39	0.56	–
ACBG11AD	0.57	−1.34	0.14	0.03	0.07	0.29	0.62	–
ACBG11BA	0.47	−0.66	0.16	0.07	0.16	0.37	0.40	–
ACBG11BB	0.51	−0.64	0.20	0.06	0.30	0.32	0.32	–
ACBG11CA	0.70	−0.55	0.29	0.05	0.33	0.38	0.23	–
ACBG11CB	0.84	−1.29	0.38	0.03	0.18	0.35	0.44	–
ACBG11CC	1.27	−1.27	0.72	0.01	0.38	0.37	0.24	–
ACBG11CD	1.13	−1.42	0.66	0.01	0.45	0.29	0.25	–
ACBG11CE	1.09	−1.10	0.60	0.03	0.30	0.39	0.29	–
ACBG11CF	0.41	0.02	0.18	0.23	0.25	0.27	0.25	–
ACBG11CG	0.52	0.26	0.23	0.24	0.31	0.30	0.15	–
ACBG12E	0.25	−0.35	0.05	0.02	0.13	0.46	0.31	0.09
ACBG12F	0.20	−0.18	0.03	0.04	0.17	0.46	0.26	0.08

Note The latent distributions of the countries are normed to an overall mean of zero. Slope and intercept are the parameters ai0 and the mean of the location parameters bi1, bi2, …, bih, etc., respectively. $I(0)$ is the information value of the item at $\theta = 0$. Cat0, Cat1, Cat2, Cat3, and Cat4 indicate the frequency with which item categories 0, 1, 2, 3 and 4 are endorsed, respectively. The components, items and corresponding category labels are described in Table 3.2

account when redesigning the instrument for future surveys; in other words, this item may be the first candidate for replacement. Compared to component 1 (early literacy activities), the items in component 2 (helping with homework) were more informative, while items in component 3 (school practices on parental involvement, parent perspective) performed poorly. Components 4 (school practices for parental involvement from a student perspective) and 5 (school practices for parental involvement from a school perspective) provided differing results; in particular, the last two items of component 5 ("parental support for student achievement within school" and "parental involvement in school activities") performed particularly poorly.

Comparing the parameter estimates in the GPCM and the GPCM with random item parameters (henceforth the random GPCM) revealed that the agreement between the slopes and intercepts under the GPCM and the means of the slopes and intercepts under the random GPCM was high (Tables 4.11, 4.12, 4.13, 4.14 and 4.15). A higher variance provides an initial indication that the item functions differently in different countries, a topic we address in more detail later. Here, the effects are global over countries and thus only permit global inferences. For instance, for component 1, the last item, ASBH02I ("read aloud signs and tables") has the lowest CDIF because the variance of the intercepts and slopes across the countries is the lowest among the items (Table 4.11). A low variance indicates that the item parameters do not vary much across countries. Evaluating the relative CDIF of the other eight items is more difficult, because of the trade-off between the standard deviation for the slope and the intercept.

Table 4.11 Item parameter estimates under the generalized partial credit model (GPCM) and GPCM with random item parameters for items in component 1: early literacy activities

Item	GPCM		GPCM random item parameters			
	Slope	Intercept	Slope	SD (Slope)	Intercept	SD (Intercept)
ASBH02A	1.26	1.84	1.37	0.22	2.06	0.66
ASBH02B	1.24	1.47	1.25	0.15	1.50	0.31
ASBH02C	0.77	0.98	0.80	0.12	1.03	0.34
ASBH02D	1.09	0.85	1.21	0.18	0.86	0.44
ASBH02E	0.95	1.80	1.01	0.19	2.01	0.68
ASBH02F	1.18	0.82	1.33	0.23	0.93	0.42
ASBH02G	1.24	0.57	1.35	0.15	0.60	0.27
ASBH02H	1.06	1.12	1.16	0.16	1.17	0.41
ASBH02I	1.07	0.89	1.09	0.11	0.87	0.22

Note The latent distributions of the countries are normed to an overall mean of zero. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

Table 4.12 Item parameter estimates under the generalized partial credit model (GPCM) and GPCM with random item parameters for items in component 2: help with homework

Item	GPCM		GPCM random item parameters			
	Slope	Intercept	Slope	SD (Slope)	Intercept	SD (Intercept)
ASBH09A	1.17	2.44	1.331	0.619	3.686	1.547
ASBH09B	1.63	2.01	1.313	0.534	2.947	1.880
ASBH09C	1.15	2.15	1.396	0.554	2.199	1.203
ASBH09D	1.10	2.27	1.227	0.314	3.736	1.610
ASBH09E	1.56	2.51	1.437	0.634	3.446	1.208
ASBH09F	1.69	1.09	1.477	0.503	0.707	1.251
ASBH09G	2.26	1.87	1.308	0.434	0.796	1.154
ASBH09H	1.45	1.66	1.559	0.224	1.518	1.210

Note The latent distributions of the countries are normed to an overall mean of zero. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

Table 4.13 Item parameter estimates under the generalized partial credit model (GPCM) and GPCM with random item parameters for items in component 3: school practices on parental involvement, parent perspective

Item	GPCM		GPCM random item parameters			
	Slope	Intercept	Slope	SD (Slope)	Intercept	SD (Intercept)
ASBH10A	0.61	1.41	1.218	1.388	4.477	4.172
ASBH10B	0.61	0.36	4.144	1.601	2.751	4.923
ASBH10E	0.58	0.52	3.843	1.791	3.469	5.232

Note The latent distributions of the countries are normed to an overall mean of zero. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

Table 4.14 Item parameter estimates under the generalized partial credit model (GPCM) and GPCM with random item parameters for items in component 4: student perception of parental involvement

Item	GPCM		GPCM random item parameters			
	Slope	Intercept	Slope	SD (Slope)	Intercept	SD (Intercept)
ASBG07A	1.01	1.47	0.924	0.161	1.473	1.102
ASBG07B	0.96	1.15	0.994	0.357	1.155	0.943
ASBG07C	0.85	1.35	0.989	0.316	1.937	2.614
ASBG07D	0.77	1.21	0.990	0.240	1.917	3.017
ASBR09C	0.55	1.55	0.553	0.050	2.100	2.782

Note The latent distributions of the countries are normed to an overall mean of zero. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

Table 4.15 Item parameter estimates under the generalized partial credit model (GPCM) and GPCM with random item parameters for items in component 5: school practices on parental involvement, school perspective

Item	GPCM		GPCM random item parameters			
	Slope	Intercept	Slope	SD (Slope)	Intercept	SD (Intercept)
ACBG11AA	0.75	−2.88	0.689	0.664	−1.667	1.396
ACBG11AB	0.91	−2.95	1.029	0.377	−2.122	0.797
ACBG11AC	0.87	−2.34	0.998	0.506	−2.110	0.778
ACBG11AD	0.57	−1.34	0.466	1.042	−1.480	0.461
ACBG11BA	0.47	−0.66	0.645	0.876	−0.581	1.033
ACBG11BB	0.51	−0.64	0.627	0.807	−0.583	0.462
ACBG11CA	0.70	−0.55	0.887	0.491	−0.576	0.434
ACBG11CB	0.84	−1.29	0.890	0.621	−1.120	0.614
ACBG11CC	1.27	−1.27	1.236	0.620	−0.995	0.682
ACBG11CD	1.13	−1.42	1.194	0.515	−1.122	0.625
ACBG11CE	1.09	−1.10	1.132	0.229	−1.023	0.168
ACBG11CF	0.41	0.02	0.548	0.738	0.029	0.342
ACBG11CG	0.52	0.26	0.737	0.514	0.071	0.781
ACBG12E	0.25	−0.35	0.123	1.453	0.551	1.954
ACBG12F	0.20	−0.18	0.279	1.431	−0.030	1.789

Note The latent distributions of the countries are normed to an overall mean of zero. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

This pattern is repeated for component 2; the items ASBH09F ("helping child practice reading") and ASBH09G ("helping child practice math skills") performed slightly better than the other items (Table 4.12). Conversely, component 3 showed a substantial difference between the item parameters estimated with the GPCM and those estimated using the random GPCM (Table 4.13), indicating this short scale was quite unstable.

The analyses of components 4 and 5 indicated all the items performed comparably with respect to CDIF (Tables 4.14 and 4.15), although questions surrounding specific item-by-country interaction and the influence of the inferences on country means and latent regression remain unanswered.

We compared CDIF as identified by the random GPCM with CDIF as identified using the latent residuals defined by Eq. (4.3) and aggregated over countries (Tables 4.16, 4.17, 4.18, 4.19 and 4.20). Overall the agreement between the methods was high. For instance, item ASBH02I performed strongly in all methods, as did item ASBH02G (Table 4.16). In general, the residuals with the GPCM are smaller than those with the PCM, because the latter model has fewer parameters. Other studies (see e.g., Glas and Jehangir 2014) confirm this expectation. However, we found that differences between the PCM and the GPCM were very small. We

Table 4.16 Absolute differential item functioning (DIF) under the partial credit model (PCM) and the generalized partial credit model (GPCM) and standard deviation random item parameters on items in component 1: early literacy activities

Item	PCM	GPCM	SD (Slope)	SD (Intercept)
ASBH02A	0.12	0.11	0.228	0.667
ASBH02B	0.08	0.08	0.158	0.318
ASBH02C	0.09	0.10	0.126	0.349
ASBH02D	0.12	0.12	0.183	0.443
ASBH02E	0.10	0.10	0.192	0.688
ASBH02F	0.09	0.09	0.239	0.421
ASBH02G	0.07	0.07	0.155	0.279
ASBH02H	0.10	0.10	0.161	0.416
ASBH02I	0.07	0.07	0.112	0.229

Note The columns labeled PCM and GPCM give the mean residuals as estimated under the unidimensional versions of these two models. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

Table 4.17 Absolute differential item functioning (DIF) under the partial credit model (PCM) and the generalized partial credit model (GPCM) and standard deviation random item parameters on items in component 2: help with homework

Item	PCM	GPCM	SD (Slope)	SD (Intercept)
ASBH09A	0.11	0.12	0.619	1.547
ASBH09B	0.07	0.07	0.534	1.880
ASBH09C	0.10	0.10	0.554	1.203
ASBH09D	0.10	0.10	0.314	1.610
ASBH09E	0.08	0.08	0.634	1.208
ASBH09F	0.14	0.12	0.503	1.251
ASBH09G	0.08	0.06	0.434	1.154
ASBH09H	0.07	0.07	0.224	1.210

Note The columns labeled PCM and GPCM give the mean residuals as estimated under the unidimensional versions of these two models. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

Table 4.18 Absolute differential item functioning (DIF) under the partial credit model (PCM) and the generalized partial credit model (GPCM) and standard deviation random item parameters on items in component 3: school practices on parental involvement, parent perspective

Item	PCM	GPCM	SD (Slope)	SD (Intercept)
ASBH10A	0.13	0.47	1.388	4.172
ASBH10B	0.07	0.36	1.601	4.923
ASBH10E	0.09	0.38	1.791	5.232

Note The columns labeled PCM and GPCM give the mean residuals as estimated under the unidimensional versions of these two models. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

Table 4.19 Absolute differential item functioning (DIF) under the partial credit model (PCM) and the generalized partial credit model (GPCM) and standard deviation random item parameters on items in component 4: student perception of parental involvement

Item	PCM	GPCM	SD (Slope)	SD (Intercept)
ASBG07A	0.08	0.07	0.161	1.102
ASBG07B	0.09	0.08	0.357	0.943
ASBG07C	0.07	0.08	0.316	2.614
ASBG07D	0.12	0.12	0.240	3.017
ASBR09C	0.07	0.08	0.050	2.782

Note The columns labeled PCM and GPCM give the mean residuals as estimated under the unidimensional versions of these two models. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

Table 4.20 Absolute differential item functioning (DIF) under the partial credit model (PCM) and the generalized partial credit model (GPCM) and standard deviation random item parameters on items in component 5: school practices on parental involvement, school perspective

Item	PCM	GPCM	SD (Slope)	SD (Intercept)
ACBG11AA	0.23	0.21	0.664	1.396
ACBG11AB	0.19	0.17	0.377	0.797
ACBG11AC	0.17	0.16	0.506	0.778
ACBG11AD	0.16	0.16	1.042	0.461
ACBG11BA	0.32	0.35	0.876	1.033
ACBG11BB	0.24	0.24	0.807	0.462
ACBG11CA	0.20	0.18	0.491	0.434
ACBG11CB	0.22	0.23	0.621	0.614
ACBG11CC	0.15	0.13	0.620	0.682
ACBG11CD	0.21	0.17	0.515	0.625
ACBG11CE	0.11	0.11	0.229	0.168
ACBG11CF	0.29	0.32	0.738	0.342
ACBG11CG	0.32	0.34	0.514	0.781
ACBG12E	0.26	0.27	1.453	1.954
ACBG12F	0.25	0.24	1.431	1.789

Note The columns labeled PCM and GPCM give the mean residuals as estimated under the unidimensional versions of these two models. SD (Slope) indicates the standard deviation of the slope. SD (Intercept) indicates the standard deviation of the intercept. Item descriptions are provided in Table 3.2

tentatively conclude the PCM fits the data quite well. A striking exception, again, was component 3. Here the fit of the GPCM was worse than the fit of the PCM, which leads to the conclusion that the slopes are very hard to estimate. This is in agreement with the reported low global reliability. Obviously, variance in the θ-distribution is too small to support a proper estimate of the slope parameters.

Table 4.21 Residual analysis for country-by-item interactions for component 1: early literacy activities

Country	Item									10 % CDIF	20 % CDIF	Absolute residual
	1	2	3	4	5	6	7	8	9			
Azerbaijan, Republic of			++		++	−−		−−	+	4	5	0.146
Australia	–									0	1	0.072
Austria	–			+						0	2	0.096
Belgium (French)										0	0	0.062
Bulgaria	+									0	1	0.058
Canada										0	0	0.059
Chinese Taipei							−−	++		2	2	0.106
Colombia	++		–						−−	2	3	0.095
Croatia	+									0	1	0.060
Czech Republic								++		1	1	0.104
Denmark	−−			++	−−					3	3	0.150
Finland	−−	+		++	−−		++			4	5	0.160
France						++				1	1	0.080
Georgia					++					1	1	0.080
Germany	–			+	–			+		0	4	0.108
Hong Kong, SAR										0	0	0.068
Hungary		−−					−−	++	++	4	4	0.148
Indonesia		++			+	–	–	−−	+	2	6	0.186
Iran, Islamic Republic of										0	0	0.067
Ireland										0	0	0.073
Israel										0	0	0.066
Italy	+									0	1	0.055
Lithuania			++							1	1	0.072

(continued)

Table 4.21 (continued)

Country	Item									10 % CDIF	20 % CDIF	Absolute residual
	1	2	3	4	5	6	7	8	9			
Malta									+	0	1	0.060
Netherlands	−−		–	+				+		1	4	0.129
New Zealand										0	0	0.068
Northern Ireland	–		−−	++	−−		++			4	5	0.168
Norway								+		0	1	0.086
Poland										0	0	0.040
Portugal						+				0	1	0.070
Qatar	+		++	–						1	3	0.111
Romania	++									1	1	0.074
Russian Federation			++							1	1	0.097
Saudi Arabia	++			–				–		1	3	0.129
Singapore				–	+					0	2	0.075
Slovak Republic										0	0	0.071
Slovenia				+						0	1	0.057
Spain	+					++				1	2	0.091
Sweden	–			++	−−		+			2	4	0.118
Trinidad and Tobago										0	0	0.074
United Arab Emirates	+			–						0	2	0.092

Note + indicates that residual belongs to the 20 % most positive residuals, ++ indicates that residual even belongs to the 10 % most positive residuals. − indicates that residual belongs to the 20 % most negative residuals, −− indicates that residual even belongs to the 10 % most negative residuals. The 10 % cultural differential item functioning (CDIF) and 20 % CDIF columns give the number of outliers in the two respective regions. Absolute residual refers to the means over items of the absolute values of the residuals. The content of items is described in Table 3.2

Table 4.22 Residual analysis for country-by-item interactions for component 2: help with homework

Country	Item								10 % CDIF	20 % CDIF	Absolute residual
	1	2	3	4	5	6	7	8			
Azerbaijan, Republic of	+								0	1	0.084
Australia			+			– –			1	2	0.097
Austria				–					0	1	0.105
Belgium (French)									0	0	0.056
Bulgaria									0	0	0.056
Canada									0	0	0.027
Chinese Taipei	++			+					1	2	0.100
Colombia									0	0	0.032
Croatia	– –		–	– –		++			3	4	0.182
Czech Republic	–				–	++		++	2	4	0.148
Denmark	–				+				0	2	0.057
Finland	–		– –	++		++			3	4	0.131
France		–						+	0	2	0.097
Georgia		+ +							1	1	0.108
Germany									0	0	0.065
Hong Kong, SAR	+			++		–	–		1	4	0.132
Hungary				–	+				0	2	0.083
Indonesia	++	+	++			– –	– –		4	5	0.188
Iran, Islamic Republic of			++					–	1	2	0.125
Ireland									0	0	0.059
Israel					+				0	1	0.076
Italy	–						+		0	2	0.098

(continued)

Table 4.22 (continued)

Country	Item								10 % CDIF	20 % CDIF	Absolute residual
	1	2	3	4	5	6	7	8			
Lithuania	–	–			+	++			1	4	0.134
Malta		+ +							1	1	0.083
Netherlands	++		++	– –	+ +	–	++	– –	6	7	0.249
New Zealand			++						1	1	0.097
Northern Ireland						– –			1	1	0.074
Norway									0	0	0.055
Poland	–					+			0	2	0.080
Portugal	++								1	1	0.069
Qatar									0	0	0.045
Romania									0	0	0.044
Russian Federation				++					1	1	0.068
Saudi Arabia						– –			1	1	0.089
Singapore	++			+			–	+	1	4	0.129
Slovak Republic	– –					++			2	2	0.115
Slovenia			–						0	1	0.080
Spain									0	0	0.064
Sweden									0	0	0.057
Trinidad and Tobago									0	0	0.038
United Arab Emirates									0	0	0.037

Note + indicates that residual belongs to the 20 % most positive residuals, ++ indicates that residual even belongs to the 10 % most positive residuals. − indicates that residual belongs to the 20 % most negative residuals, −− indicates that residual belongs to the 10 % most negative residuals. The 10 % cultural differential item functioning (CDIF) and 20 % CDIF columns give the number of outliers in the two respective regions. Absolute residual refers to the means over items of the absolute values of the residuals. The content of items is described in Table 3.2

Table 4.23 Residual analysis for country-by-item interactions for component 3: school practices on parental involvement, parent perspective

Country	Item			10 % CDIF	20 % CDIF	Absolute residual
	1	2	3			
Azerbaijan, Republic of			+	0	1	0.084
Australia				0	0	0.032
Austria	++			1	1	0.102
Belgium (French)	+			0	1	0.088
Bulgaria	+			0	1	0.110
Canada				0	0	0.058
Chinese Taipei				0	0	0.057
Colombia	−−			1	1	0.112
Croatia	−−			1	1	0.090
Czech Republic	++			1	1	0.085
Denmark	−−			1	1	0.071
Finland	++			1	1	0.096
France	+			0	1	0.081
Georgia	−−			1	1	0.088
Germany	++			1	1	0.164
Hong Kong, SAR	+			0	1	0.054
Hungary				0	0	0.026
Indonesia	–	+		0	2	0.142
Iran, Islamic Republic of				0	0	0.034
Ireland				0	0	0.073
Israel				0	0	0.042
Italy	+			0	1	0.106
Lithuania				0	0	0.029
Malta	−−			1	1	0.082
Netherlands				0	0	0.039
New Zealand				0	0	0.037
Northern Ireland				0	0	0.030
Norway	−−			0	0	0.104
Poland				0	0	0.050
Portugal				0	0	0.037
Qatar	+			0	1	0.075
Romania	−−			1	1	0.127
Russian Federation		+		0	1	0.088
Saudi Arabia	–			0	1	0.048
Singapore	++			1	1	0.083
Slovak Republic				0	0	0.049
Slovenia	++			1	1	0.072
Spain				0	0	0.018

(continued)

Table 4.23 (continued)

Country	Item			10 % CDIF	20 % CDIF	Absolute residual
	1	2	3			
Sweden				0	0	0.016
Trinidad and Tobago	–			0	1	0.109
United Arab Emirates				0	0	0.044

Note + indicates that residual belongs to the 20 % most positive residuals, ++ indicates that residual even belongs to the 10 % most positive residuals. − indicates that residual belongs to the 20 % most negative residuals, −− indicates that residual even belongs to the 10 % most negative residuals. The 10 % cultural differential item functioning (CDIF) and 20 % CDIF columns give the number of outliers in the two respective regions. Absolute residual refers to the means over items of the absolute values of the residuals. Item descriptions are provided in Table 3.2

Table 4.24 Residual analysis for country-by-item interactions for component 4: student perception of parental involvement

Country	Item					10 % CDIF	20 % CDIF	Absolute residual
	1	2	3	4	5			
Azerbaijan, Republic of						0	0	0.040
Australia						0	0	0.060
Austria						0	0	0.037
Belgium (French)				–		0	1	0.076
Bulgaria				+		0	1	0.075
Canada						0	0	0.068
Chinese Taipei	++			−−		2	2	0.117
Colombia						0	0	0.034
Croatia	–			++		1	2	0.094
Czech Republic						0	0	0.051
Denmark						0	0	0.056
England	+		–			0	2	0.088
Finland					++	1	1	0.103
France						0	0	0.068
Georgia			+			0	1	0.075
Germany	+			−−		1	2	0.146
Hong Kong, SAR		–				0	1	0.087
Hungary	–			++		1	2	0.110
Indonesia					−−	1	1	0.080
Iran, Islamic Republic of			+			0	1	0.071
Ireland		+		−−		1	2	0.112

(continued)

Table 4.24 (continued)

Country	Item					10 % CDIF	20 % CDIF	Absolute residual
	1	2	3	4	5			
Israel				++		1	1	0.120
Italy					+	0	1	0.066
Lithuania						0	0	0.061
Malta					–	0	1	0.078
Netherlands		−−	++	++	++	4	4	0.233
New Zealand						0	0	0.027
Northern Ireland	+	++		−−	–	2	4	0.197
Norway			+			0	1	0.087
Poland		+		−−	++	2	3	0.158
Portugal				+		0	1	0.074
Qatar			+			0	1	0.089
Romania				++		1	1	0.082
Russian Federation						0	0	0.073
Saudi Arabia		–	++			1	2	0.133
Singapore			−−		++	2	2	0.104
Slovak Republic						0	0	0.070
Slovenia						0	0	0.059
Spain						0	0	0.056
Sweden	+			−−		1	2	0.090
Trinidad and Tobago						0	0	0.049
United Arab Emirates						0	0	0.060
United States						0	0	0.084

Note + indicates that residual belongs to the 20 % most positive residuals, ++ indicates that residual even belongs to the 10 % most positive residuals. − indicates that residual belongs to the 20 % most negative residuals, −− indicates that residual even belongs to the 10 % most negative residuals. The 10 % cultural differential item functioning (CDIF) and 20 % CDIF columns give the number of outliers in the two respective regions. Absolute residual refers to the means over items of the absolute values of the residuals. Item descriptions are provided in Table 3.2

We then addressed the distribution of country-by-item interaction across countries and items, to determine whether the sizes and directions of the residuals were randomly distributed across all countries and items, or whether they exhibited notable patterns of interaction (Tables 4.21, 4.22, 4.23, 4.24 and 4.25). Residuals were defined by Eq. (4.3), estimated under the GPCM, and calculated for every country, with that country as a focus and all other countries as a reference. To simplify, here we shall not consider the specific values of the residuals, but instead concentrate on the outlying values. For example, if we examine results obtained for the Republic of Azerbaijan and Australia for component 1 (early literacy activities,

Table 4.25 Residual analysis for country-by-item interactions for component 5: school practices on parental involvement, school perspective

Country	Item															10 % CDIF	20 % CDIF	Absolute residual
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
Azerbaijan					−−											1	1	0.164
Australia								+								0	1	0.152
Austria					++		+			–		+	–			1	5	0.294
Belgium (F)	+		+		–								−−			1	4	0.289
Bulgaria	+														–	0	2	0.227
Canada					+											0	1	0.170
Chinese Taipei							–						++	++		2	3	0.186
Colombia	+											−−		−−		2	3	0.268
Croatia	++				−−			–					++			3	4	0.352
Czech Republic	+					–		+					−−		−−	2	5	0.286
Denmark						–		+				++			++	2	4	0.235
England	+						–						–			0	3	0.209
Finland		−−						++				++	++	++	+	5	6	0.294
France					++	++			+			–	−−			3	5	0.271
Georgia	+											−−		–		1	3	0.195
Germany					++	++										2	2	0.173
Hong Kong					+								++			1	2	0.250
Hungary	++				+	–						−−		−−		3	5	0.237
Indonesia					−−	+	+						−−	++	+	3	6	0.302
Iran					−−	+										1	2	0.203
Ireland							–	++				++		++	+	3	5	0.302
Israel																0	0	0.121

(continued)

Table 4.25 (continued)

Country	Item															10 % CDIF	20 % CDIF	Absolute residual
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
Italy	++							–								1	2	0.200
Lithuania												−−				1	1	0.162
Malta													+			0	1	0.124
Netherlands				−−	++							+			+	2	4	0.213
New Zealand					++			++						+	+	2	4	0.247
Northern Ireland												+	–			0	2	0.156
Norway	++				+	–				−−				–		2	5	0.289
Poland					+			++					++			2	3	0.213
Portugal	++		+							–						1	3	0.226
Qatar					–	−−										1	2	0.153
Romania						+								−−	−−	2	3	0.223
Russian Fed.	++					+			–	−−		−−	++	–		4	7	0.367
Saudi Arabia	++				−−			−−		+		–			−−	4	6	0.315
Singapore					+								++			1	2	0.223
Slovak Republic														−−	–	1	2	0.200
Slovenia													++			1	1	0.176
Spain	+							–								0	2	0.176
Sweden									+			++	–			1	3	0.226
Trinidad and Tobago										+		−−	−−	−−		3	4	0.232
United Arab Emirates																0	0	0.128
United States																0	0	0.115

Note + indicates that residual belongs to the 20 % most positive residuals, ++ indicates that residual even belongs to the 10 % most positive residuals. − indicates that residual belongs to the 20 % most negative residuals, −− indicates that residual even belongs to the 10 % most negative residuals. The 10 % cultural differential item functioning (CDIF) and 20 % CDIF columns give the number of outliers in the two respective regions. Absolute residual refers to the means over items of the absolute values of the residuals. Item descriptions are provided in Table 3.2

Table 4.21), it is clear that, aggregated over the items, the mean absolute residual for the Republic of Azerbaijan is much larger than the mean absolute residual for Australia. The responses were coded 0, 1 and 2, so the residuals, which are the differences between a mean observed and expected response are also on a scale from 0 to 2. Closer inspection at the item level for Republic of Azerbaijan reveals that items 3 and 5 have residuals among the 10 % most positive among the countries, while the items 6 and 8 have residuals among the 10 % most negative among the countries. Australia, however, has only one negative residual, and this is among the 20 % most negative residuals among the countries. Checking the absolute residuals further reveals Poland fits the model best with the lowest CDIF, while Indonesia has the most significant CDIF.

In a similar way, component 2 (helping with homework) functions very differently in the Netherlands than in other countries (Table 4.22), probably because giving students homework is not a daily practice in Dutch primary schools. This different item functioning is indicated by both the high mean for the absolute values of the residuals and the large number of outliers among the residuals. Canada fits the model best, having the lowest CDIF for this component. For component 3 (school practices on parental involvement, parents perspective) the highest mean absolute residual was found for Germany. However, the scale for measuring school practices on parental involvement from the school perspective (component 5) showed relatively little evidence of CDIF.

We undertook a marginal count of the outliers for the items aggregated over the countries (Table 4.26). No one item count was prominent, although the first item in component 3 ("my child's school includes me in my child's education") seemed more susceptible to CDIF than other items, since this item had the greatest number of residual outliers among countries: 13 in the 10 % outliers region and 15 in the 20 % outliers region. Items 5 ("volunteering") and 13 ("organize workshops or seminars for parents on learning or pedagogical issues") within component 5 also scored more highly than other items in the component. However, this does not of course mean that these items have CDIF; if 10 and 20 % extreme values are considered, then 10 and 20 % of the residuals must be included, thus such information only serves as a tool to further scrutinize the items.

We also calculated country-specific factor loadings for the bi-factor model, where we first transformed country-specific factor loadings to standard normals, and then identified the 2.5 and 5 % most extreme outlying values (Tables 4.27, 4.28, 4.29, 4.30 and 4.31). This distribution of country-specific factor loadings gives an indication of the extent to which items load on a country-specific factor in addition to the general factor of the item, and can, as in our earlier residual analysis, be used to determine whether the sizes and directions of the factor loadings are randomly distributed across all countries and items, or whether they exhibit notable patterns of interaction.

For component 1, the greatest number of outliers of the country-specific factor loadings and the highest mean absolute factor loading were found for Colombia (Table 4.27), suggesting a high level of CDIF. Interestingly, in the residual analysis for this component, a total of 15 countries showed a higher mean absolute residual

Table 4.26 Distribution of cultural differential item functioning (CDIF) across items on parental involvement

% CDIF	Component	Items														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
10	1	6	2	5	4	6	3	4	5	2	–	–	–	–	–	–
	2	7	2	5	5	1	9	2	2	–	–	–	–	–	–	–
	3	13	0	0	–	–	–	–	–	–	–	–	–	–	–	–
	4	1	2	3	11	5	–	–	–	–	–	–	–	–	–	–
	5	7	1	0	1	10	3	0	5	0	2	0	10	13	9	4
20	1	17	3	7	12	9	5	6	9	5	–	–	–	–	–	–
	2	15	5	8	9	6	12	5	5	–	–	–	–	–	–	–
	3	22	2	1	–	–	–	–	–	–	–	–	–	–	–	–
	4	7	6	8	14	8	–	–	–	–	–	–	–	–	–	–
	5	14	1	2	1	18	11	5	11	3	6	0	15	18	13	11

Note Count of the number of times a residual is in the extreme 10 % and extreme 20 % region of the distribution of residuals. The components, items and corresponding category labels are described in Table 3.2

Table 4.27 Outliers of country-specific factor loadings in the bi-factor model for component 1: early literacy activities

Country	Item									2.5 % Outlier	5 % Outlier	Mean absolute loading
	1	2	3	4	5	6	7	8	9			
Azerbaijan, Republic of					+	−−		−−		3	3	0.062
Australia										0	0	0.034
Austria				−−				−−	−−	3	3	0.071
Belgium (French)	+							++	+	2	3	0.042
Bulgaria										0	0	0.029
Canada								+		1	1	0.050
Chinese Taipei										0	0	0.036
Colombia	−−		−−	–	−−		−−	−−		5	6	0.095
Croatia	−−	−−	–		−−	−−				4	5	0.079
Czech Republic	−−		–	+	−−	−−				4	5	0.083
Denmark	−−				−−	−−				3	3	0.080
Finland	−−	−−								2	2	0.066
France		–			−−					1	2	0.048
Georgia				−−			−−	−−	−−	4	4	0.077
Germany										0	0	0.029
Hong Kong, SAR				−−		−−		−−	−−	4	4	0.082
Hungary		++				–	–	−−		1	4	0.066
Indonesia				++						0	1	0.033
Iran, Islamic Republic of								−−		1	1	0.053
Ireland	−−	−−			–	–				2	4	0.062
Israel	−−	–			–	−−				2	4	0.064
Italy	−−	−−	−−				–			3	4	0.078
Lithuania					–	−−				1	2	0.044

(continued)

Table 4.27 (continued)

Country	Item									2.5 % Outlier	5 % Outlier	Mean absolute loading
	1	2	3	4	5	6	7	8	9			
Malta	−−	–	−−		−−	–				3	5	0.076
Netherlands	++									0	1	0.028
New Zealand	−−		−−		−−	−−				4	4	0.081
Northern Ireland	−−	–		−−	−−				−−	4	5	0.088
Norway	−−		–	–	−−		−−			3	5	0.087
Poland			++	–			−−	−−		2	4	0.056
Portugal										0	0	0.034
Qatar				++						0	1	0.047
Romania				−−			–	−−		2	3	0.060
Russian Federation					++					0	1	0.032
Saudi Arabia	−−	−−				–				2	3	0.065
Singapore				−−			–	−−		2	3	0.057
Slovak Republic	−−				−−	−−				3	3	0.062
Slovenia	−−				−−	−−				3	3	0.065
Spain		–			−−	−−				2	3	0.049
Sweden				−−			–	–		1	3	0.052
Trinidad and Tobago	–	++					–	−−		1	4	0.058
United Arab Emirates	−−	−−			–					2	3	0.065

Note + indicates factor loading belongs to the 5 % most positive loading, ++ indicates factor loading belongs to the 2.5 % most positive loading. − indicates factor loading belongs to the 5 % most negative loading, −− indicates factor loading belongs to the 2.5 % most negative loading. The 2.5 % cultural differential item functioning (CDIF) and 5 % CDIF columns give the number of outliers in the two respective regions. Mean absolute loading refers to the means over items of the absolute values of country-specific factor loadings. Item descriptions are provided in Table 3.2

Table 4.28 Outliers of country-specific factor loadings in the bi-factor model for component 2: help with homework

Country\item	1	2	3	4	5	6	7	8	2.5 % Outlier	5 % Outlier	Mean absolute loading
Azerbaijan, Republic of	+					–			1	2	0.056
Australia						−−	−−	−−	3	3	0.060
Austria								–	0	1	0.037
Belgium (French)	++								0	1	0.043
Bulgaria						−−	−−	−−	3	3	0.051
Canada	+								1	1	0.044
Chinese Taipei						–	−−		1	2	0.037
Colombia	−−			−−		+		–	3	4	0.067
Croatia	−−			–	−−		–	–	2	5	0.068
Czech Republic							–	–	0	2	0.040
Denmark									0	0	0.029
Finland									0	0	0.047
France	+	+				–	–	–	2	5	0.069
Georgia								–	0	1	0.032
Germany	+					–	–		1	3	0.055
Hong Kong, SAR				–				–	0	2	0.044
Hungary	+					−−	−−	–	3	4	0.076
Indonesia						−−	−−	–	2	3	0.057
Iran, Islamic Republic of									0	0	0.048
Ireland									0	0	0.029
Israel	−−			−−				−−	3	3	0.056
Italy									0	0	0.041
Lithuania		++				−−	−−	−−	3	4	0.075

(continued)

Table 4.28 (continued)

Country\item	1	2	3	4	5	6	7	8	2.5 % Outlier	5 % Outlier	Mean absolute loading
Malta	+		+	−−	++		++		3	5	0.072
Netherlands						−−	−−	−−	3	3	0.058
New Zealand						−−	−−	−−	3	3	0.059
Northern Ireland	−−							−−	2	2	0.053
Norway	+					–	−−	−−	3	4	0.066
Poland						−−	−−	–	2	3	0.050
Portugal	++								0	1	0.037
Qatar						−−	−−		2	2	0.041
Romania						−−	−−		2	2	0.051
Russian Federation	+						−−		2	2	0.043
Saudi Arabia				–		+			1	2	0.048
Singapore									0	0	0.036
Slovak Republic								–	0	1	0.043
Slovenia						–	–	–	0	3	0.051
Spain	+					−−			2	2	0.038
Sweden						−−	−−	–	2	3	0.055
Trinidad and Tobago							−−		1	1	0.049
United Arab Emirates								–	0	1	0.036

Note + indicates factor loading belongs to the 5 % most positive loading, ++ indicates factor loading belongs to the 2.5 % most positive loading. − indicates factor loading belongs to the 5 % most negative loading, −− indicates factor loading belongs to the 2.5 % most negative loading. The 2.5 % cultural differential item functioning (CDIF) and 5 % CDIF columns give the number of outliers in the two respective regions. Mean absolute loading refers to the means over items of the absolute values of country-specific factor loadings. Item descriptions are provided in Table 3.2

Table 4.29 Outliers of country-specific factor loadings in the bi-factor model for component 3: school practices on parental involvement, parent perspective

Country	Item			2.5 % Outlier	5 % Outlier	Mean absolute loading
	1	2	3			
Azerbaijan, Republic of	+	++		1	2	1.097
Australia				0	0	0.293
Austria				0	0	0.203
Belgium (French)				0	0	0.223
Bulgaria				0	0	0.262
Canada				0	0	0.423
Chinese Taipei			+	1	1	0.640
Colombia				0	0	0.159
Croatia				0	0	0.194
Czech Republic				0	0	0.393
Denmark				0	0	0.284
Finland				0	0	0.293
France				0	0	0.293
Georgia				0	0	0.240
Germany				0	0	0.409
Hong Kong, SAR				0	0	0.168
Hungary			+	1	1	1.521
Indonesia			++	0	1	0.500
Iran, Islamic Republic of				0	0	0.362
Ireland				0	0	0.279
Israel				0	0	0.216
Italy				0	0	0.131
Lithuania				0	0	0.174
Malta				0	0	0.418
Netherlands				0	0	0.331
New Zealand				0	0	0.260
Northern Ireland				0	0	0.321
Norway				0	0	0.228
Poland				0	0	0.213
Portugal				0	0	0.205
Qatar				0	0	0.297
Romania				0	0	0.430
Russian Federation				0	0	0.153
Saudi Arabia				0	0	0.184
Singapore				0	0	0.150
Slovak Republic				0	0	0.180
Slovenia				0	0	0.228

(continued)

Table 4.29 (continued)

Country	Item			2.5 % Outlier	5 % Outlier	Mean absolute loading
	1	2	3			
Spain				0	0	0.347
Sweden				0	0	0.315
Trinidad and Tobago				0	0	0.517
United Arab Emirates				0	0	0.175

Note + indicates factor loading belongs to the 5 % most positive loading, ++ indicates factor loading belongs to the 2.5 % most positive loading. − indicates factor loading belongs to the 5 % most negative loading, −− indicates factor loading belongs to the 2.5 % most negative loading. The 2.5 % cultural differential item functioning (CDIF) and 5 % CDIF columns give the number of outliers in the two respective regions. Mean absolute loading refers to the means over items of the absolute values of country-specific factor loadings. Item descriptions are provided in Table 3.2

Table 4.30 Outliers of country-specific factor loadings in the bi-factor model for component 4: student perception of parental involvement

Country	Item					2.5 % Outlier	5 % Outlier	Mean absolute loading
	1	2	3	4	5			
Azerbaijan, Republic of						0	0	0.024
Australia						0	0	0.012
Austria						0	0	0.021
Belgium (French)				+		1	1	0.026
Bulgaria						0	0	0.016
Canada	+		++			1	2	0.048
Chinese Taipei						0	0	0.012
Colombia	−−	–				1	2	0.044
Croatia	–	−−	−−	+		3	4	0.084
Czech Republic						0	0	0.018
Denmark						0	0	0.010
England						0	0	0.027
Finland		–				0	1	0.038
France				–	–	0	2	0.035
Georgia	++	++	+	++		1	4	0.057
Germany	−−					1	1	0.034
Hong Kong, SAR						0	0	0.011
Hungary						0	0	0.026
Indonesia				–		0	1	0.029

(continued)

Table 4.30 (continued)

Country	Item					2.5 % Outlier	5 % Outlier	Mean absolute loading
	1	2	3	4	5			
Iran, Islamic Republic of				+		1	1	0.034
Ireland			–			0	1	0.022
Israel						0	0	0.024
Italy						0	0	0.023
Lithuania		−−	++	+		2	3	0.052
Malta						0	0	0.008
Netherlands			–	−−		1	2	0.048
New Zealand						0	0	0.031
Northern Ireland		–				0	1	0.038
Norway						0	0	0.021
Poland				++		0	1	0.025
Portugal				–		0	1	0.030
Qatar			++			0	1	0.034
Romania	++					0	1	0.037
Russian Federation						0	0	0.027
Saudi Arabia			–			0	1	0.028
Singapore						0	0	0.031
Slovak Republic						0	0	0.016
Slovenia						0	0	0.018
Spain						0	0	0.020
Sweden						0	0	0.022
Trinidad and Tobago						0	0	0.020
United Arab Emirates			−−	−−		2	2	0.047
United States						0	0	0.019

Note + indicates factor loading belongs to the 5 % most positive loading, ++ indicates factor loading belongs to the 2.5 % most positive loading. − indicates factor loading belongs to the 5 % most negative loading, −− indicates factor loading belongs to the 2.5 % most negative loading. The 2.5 % cultural differential item functioning (CDIF) and 5 % CDIF columns give the counts of the outliers in the two respective regions. Mean absolute loading refers to the means over items of the absolute values of country-specific factor loadings. Item descriptions are provided in Table 3.2

Table 4.31 Outliers of country-specific factor loadings in the bi-factor model for component 5: school practices on parental involvement, school perspective

Country	Item															2.5 % Outlier	5 % Outlier	Mean absolute loading
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
Azerbaijan																0	0	0.043
Australia								+								1	1	0.078
Austria																0	0	0.048
Belgium (F)	+	+														2	2	0.075
Bulgaria																0	0	0.053
Canada																0	0	0.057
Chinese Taipei	+	+														2	2	0.121
Colombia					−−											1	1	0.066
Croatia																0	0	0.043
Czech Republic																0	0	0.047
Denmark																0	0	0.041
England																0	0	0.041
Finland		+														1	1	0.066
France																0	0	0.028
Georgia													++			0	1	0.047
Germany	+															1	1	0.060
Hong Kong																0	0	0.041
Hungary																0	0	0.070
Indonesia																0	0	0.059
Iran																0	0	0.035
Ireland																0	0	0.050
Israel	+	+														2	2	0.072
Italy																0	0	0.030

(continued)

Table 4.31 (continued)

Country	Item															2.5 % Outlier	5 % Outlier	Mean absolute loading
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
Lithuania																0	0	0.035
Malta					+			+								2	2	0.088
Netherlands																0	0	0.041
New Zealand	+	+														2	2	0.116
Northern Ireland	+	+	+													3	3	0.105
Norway										++						0	1	0.083
Poland	+	+	+													3	3	0.113
Portugal	+	+	+	+												4	4	0.179
Qatar																0	0	0.041
Romania																0	0	0.042
Russian Fed.	+	+														2	2	0.067
Saudi Arabia																0	0	0.040
Singapore	++	++														0	2	0.055
Slovak Republic																0	0	0.043
Slovenia										++						0	1	0.063
Spain																0	0	0.044
Sweden																0	0	0.071
Trinidad and Tobago								+								1	1	0.069
United Arab Emirates																0	0	0.054
United States																0	0	0.053

Note + indicates factor loading belongs to the 5 % most positive loading, ++ indicates factor loading belongs to the 2.5 % most positive loading. − indicates factor loading belongs to the 5 % most negative loading, −− indicates factor loading belongs to the 2.5 % most negative loading. The 2.5 % cultural differential item functioning (CDIF) and 5 % CDIF columns give the number of outliers in the two respective regions. Mean absolute loading refers to the means over items of the absolute values of country-specific factor loadings. Item descriptions are provided in Table 3.2

Table 4.32 Distribution of outliers of country-specific factor loadings in the bi-factor model across items on parental involvement

Region	Component	Items														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2.5 %	1	7	6	4	8	13	11	4	12	5	–	–	–	–	–	–
	2	13	1	1	3	1	14	15	8	–	–	–	–	–	–	–
	3	1	0	2	–	–	–	–	–	–	–	–	–	–	–	–
	4	3	2	3	6	0	–	–	–	–	–	–	–	–	–	–
	5	9	9	3	1	2	0	0	3	0	0	0	0	0	0	0
5 %	1	19	8	13	13	18	15	10	14	5	–	–	–	–	–	–
	2	15	2	1	6	2	20	21	22	–	–	–	–	–	–	–
	3	1	1	3	–	–	–	–	–	–	–	–	–	–	–	–
	4	6	6	9	11	1	–	–	–	–	–	–	–	–	–	–
	5	10	10	3	1	2	0	0	3	0	2	0	0	1	0	0

Note Count of the number of times an outlier of the country-specific factor loading is in the extreme 2.5 % and extreme 5 % region of the distribution of factor loadings. The components, items and corresponding category labels are described in Table 3.2

Table 4.33 Relation between residuals under the generalized partial credit model (GPCM) and country-specific factor loadings in the bi-factor GPCM

Component	Correlation	Kappa classification CDIF		
		Size middle group		
		33 %[a]	40 %[b]	80 %[c]
1	0.228	0.15	0.20	0.24
2	0.603	0.21	0.29	0.27
3	−0.044	0.07	0.17	0.10
4	0.651	0.46	0.41	0.41
5	0.519	0.34	0.31	0.25

Note Correlation between the GPCM residuals and the country-specific factor loadings, over countries and items. The content of the components is described in Table 3.2. Size middle group indicates the classification of the ordered residuals and country-specific factor loadings in three categories according to their size: a category with negative values, a category with positive values and a middle category

Norms for Kappa: poor agreement = 0.00–0.19, fair agreement = 0.20–0.39, and moderate agreement = 0.40–0.59

[a]Three equally-sized categories

[b]The middle category contained 40 % of the values, the two extreme categories each contained 20 %

[c]The middle category contained 80 % of the values, the two extreme categories each contained 10 %

(Table 4.21). Regarding help with homework (component 2), Malta was identified as having the highest number of outliers in country-specific factor loadings (Table 4.28), while The Netherlands, which we earlier identified as exhibiting CDIF for component 2 (Table 4.22), also had a high number of outliers. For component 3, counting the number of outliers provided little information, as only three outliers were counted in the 2.5 % region (Table 4.29). Hungary did show a high mean absolute country-specific factor loading on this component, though the questionable reliability of the scale must be kept in mind. Student perception of parental involvement (component 4) was measured with the least CDIF in Denmark, whereas the school practices on parental involvement from the school perspective showed the least CDIF for Italy (Tables 4.30 and 4.31).

Aggregating the items over the countries provides a tool for further investigation of items (Table 4.32), with the same caveats as before; if the 2.5 and 5 % most extreme values are considered, then similarly 2.5 and 5 % of the residuals must fall in this region, but this does not imply that 2.5 and 5 % of the items have CDIF. No item count is prominent. Item 5 (“talk about things you had done”) in component 1 did seem more susceptible to CDIF than other items, since this item revealed the greatest number of outliers in country-specific factor loadings over countries.

We then addressed whether the residual analyses using the GPCM and the bi-factor GPCM analyses led to the same conclusions (see Table 4.33). A priori, this would be unexpected. The residual analyses target so-called uniform CDIF, namely a shift in the item location (item intercept) parameters over countries. The

Table 4.34 Correlation and rank correlation between country means estimated with no, 10 and 20 % cultural differential item functioning (CDIF) parameters, and random item parameters

Component	Parameter	Correlation			Rank correlation		
		No CDIF	10 % CDIF	20 % CDIF	No CDIF	10 % CDIF	20 % CDIF
1	10 % CDIF	0.99	–	–	0.98	–	–
	20 % CDIF	0.99	0.99	–	0.98	0.98	–
	Random	0.98	0.97	0.97	0.97	0.96	0.97
2	10 % CDIF	0.99	–	–	0.98	–	–
	20 % CDIF	0.98	0.99	–	0.98	0.99	–
	Random	0.66	0.64	0.58	0.95	0.93	0.95
3	10 % CDIF	0.83	–	–	0.94	–	–
	20 % CDIF	0.80	0.82	–	0.93	1.00	–
	Random	0.53	0.38	0.33	0.62	0.64	0.63
4	10 % CDIF	0.98	–	–	0.97	–	–
	20 % CDIF	0.97	0.98	–	0.95	0.95	–
	Random	0.50	0.44	0.37	0.94	0.92	0.89
5	10 % CDIF	0.97	–	–	0.97	–	–
	20 % CDIF	0.97	1.00	–	0.97	1.00	–
	Random	0.97	0.98	0.98	0.97	0.99	0.99

Note The components are described in Table 3.2

bi-factor analyses target non-uniform CDIF, namely differences in the slopes and the dimensionality across items. The correlations for components 2, 4 and 5 were moderate, while for component 1, the correlation was much lower, and for component 3, the correlation completely vanished. The result for component 3 is probably because both the residuals and the country-specific factor loadings are poorly estimated for a test containing only three items.

Though the correlation between the residuals and the country-specific factor loadings is a reasonable estimate between the two measures, it does not properly indicate to what extent the two measures have the same outliers. To investigate this, we ordered and classified the residuals and country-specific factor loadings in three categories according to their size (a category with negative values, a category with positive values and a middle category). Further, we varied the definition of what constituted an outlying value by varying the size of the middle group (assigning it variously as 33, 40, or 80 % of values). The calculation of Kappa establishes the agreement in categorization between the residual analyses using the GPCM and the bi-factor GPCM. This revealed that agreement was poor throughout for component 3, while, for component 1, the agreement was poor in the 33 % category; for other categories in component 1 the agreement was only fair to moderate. In general, the results indicate that it is not a good policy to rely on one approach for the investigation of CDIF.

We investigated the influence of CDIF by calculating the correlation and rank correlation between country means estimated with no, 10, and 20 % CDIF

parameters, and with random item parameters (Table 4.34). Estimates of the means using the unidimensional GPCM without country-specific item parameters and using the bi-factor GPCM could not be distinguished, so we exclude them from further discussion. In general, correlations were high, indicating that, in the estimation of the country means and the rank order of the country means, CDIF had little impact. Component 3 remained the exception; both correlations and rank correlations were low. Further, for components 2 and 4, the correlations between the means estimated using the GPCM with random item parameters and the other three models were also low; however this was not the case for the rank correlations. This is because the relationship between means is not linear. We discuss the possible influence of CDIF further in the next chapter.

References

Adams, R., & Wu, M. (2006). The mixed-coefficients multinomial logit model: A generalized form of the Rasch model. In M. von Davier & C. H. Carstensen (Eds.), *Multivariate and mixture distribution Rasch models: Extensions and applications* (pp. 57–75). New York: Springer.

Aitchison, J., & Silvey, S. D. (1958). Maximum likelihood estimation of parameters subject to restraints. *Annals of Mathematical Statistics, 29*, 813–828.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM-algorithm. *Psychometrika, 46*, 443–459.

Bock, R. D., Gibbons, R. D., & Muraki, E. (1988). Full-information factor analysis. *Applied Psychological Measurement, 12*, 261–280.

De Jong, M. G., Steenkamp, J. B. E. M., & Fox, J. P. (2007). Relaxing measurement invariance in cross-national consumer research using a hierarchical IRT model. *Journal of Consumer Research, 34*(2), 260–278. doi:10.1086/518532.

Gibbons, R. D., & Hedeker, D. R. (1992). Full-information item bi-factor analysis. *Psychometrika, 57*(3), 423–436. doi:10.1007/BF02295430.

Glas, C. A. W. (1999). Modification indices for the 2-PL and the nominal response model. *Psychometrika, 64*(3), 273–294. doi:10.1007/bf02294296.

Glas, C. A. W. (2010). Multidimensional item response theory (MIRT), manual and computer program. Retrieved from http://www.utwente.nl/gw/omd/Medewerkers/temp_test/mirt_package.zip, http://www.utwente.nl/gw/omd/Medewerkers/temp_test/mirt-manual.pdf.

Glas, C. A. W., & Dagohoy, A. V. T. (2007). A person fit test for IRT models for polytomous items. *Psychometrika, 72*(2), 159–180. doi:10.1007/s11336-003-1081-5.

Glas, C. A. W., & Falcón, J. C. S. (2003). A comparison of item-fit statistics for the three-parameter logistic model. *Applied Psychological Measurement, 27*(2), 87–106. doi:10.1177/0146621602250530.

Glas, C. A. W., & Jehangir, K. (2014). Modeling country-specific differential item functioning. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 97–115). New York: Springer.

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika, 76*(4), 537–549. doi:10.1007/s11336-011-9218-4.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149–174.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16*, 159–176.

Rao, C. R. (1947). Large sample tests of statistical hypothesis concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society, 44*, 50–57.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph, 34*(4, Pt. 2), 1–100.

Tutz, G. (1990). Sequential item response models with an ordered response. *British Journal of Mathematical and Statistical Psychology, 43*, 39–55.

5 Relation Between Parental Involvement and Student Achievement in PIRLS-2011

5.1 Method and Rationale for Latent Regression Models

After modeling parental involvement, we investigated its relationship with reading literacy using a three-level regression model in which students (level 1) were clustered within schools (level 2) and schools were clustered within countries (level 3). Although it is important to recognize that the countries participating in PIRLS-2011 cannot necessarily be regarded as being representative of the whole world, incorporating a country level in our analyses provides some indication whether the parental involvement influences reading achievement in countries worldwide. The majority of previous studies on parental involvement applied only to the USA.

In our multilevel model, the dependent variable was the reading literacy variable from the PIRLS dataset. To account for the unreliability of this outcome variable, five plausible values are available in the PIRLS dataset. All analyses were repeated for all five plausible values and then aggregated to overall estimates of fixed and random factors, thus incorporating the differences in standard errors for the different effect sizes (Von Davier et al. 2009). The sampling procedure was accounted for by including a student-class weight at the student level and a school weight at the school level. It is important to concurrently use school and student weights in the analyses, because schools were sampled first and then students were sampled within schools.

To assess the relationships between the components of interest and the reading literacy outcome, important determinants of student achievement such as gender and socioeconomic status (SES) were also included. The analytic framework described in Chap. 3 included questions that could be regarded as proxies for SES, such as “number of books in the home” and “highest educational level attained by one of the parents”. Together with gender, these variables were added as control variables.

We evaluated the different multilevel models for each set of estimates resulting from the measurement models described in Chap. 4. Having already compared the outcomes for the means on the latent scales for the different corrections for CDIF, comparing the results of the structural multilevel model reveals the extent to which CDIF may influence the relationship between parental involvement and students' reading literacy, and level of control provided by the measurement models.

The five variables for the five parental involvement components were generated using the five IRT models considered in the previous chapter: the GPCM, the GPCM with 10 and 20 % country-by-item interaction, the GPCM with random item parameters and the bi-factor GPCM. We included the last four models to assess whether taking potential CDIF into account would direct toward different conclusions. For all five models, we obtained a posteriori estimates for the student parameters and entered these estimates as independent variables into the multilevel model.

For the GPCM, first an empty model (model 0) was estimated to see how the variance in the outcome variable is distributed over the three levels. Subsequently, control variables (student background characteristics) were added as fixed effects (model 1). The resulting model can be seen as a baseline to which the models, including the parental involvement variables of interest, can be compared. The separate parental involvement components were added as fixed effects on either the student level (i.e., components 1–4) or school level (component 5), resulting in models 2A–2E. We also created a model that included all five components simultaneously (model 3).

By entering the five components as fixed factors, the factor was assumed to have the same effect across all countries. However, in the context of this study, we also wanted to determine the extent of differences in the effects of parental involvement across countries. Therefore, we also considered a model with random slopes at the country level for the parental involvement components (model 4). A random slopes model includes a variance component for the slope of one or more predictor variables, while the other models may be considered special versions obtained by fixing parameters. The full model, a random intercepts-and-slopes model, with one component for parental involvement is given by:

$$Y_{ijk} = \beta_{0jk} + \beta_1 Gender + \beta_2 Books + \beta_3 Educ + \beta_{4jk} Construct + \epsilon_{ijk},$$

where ϵ_{ijk} is a normally distributed error term with variance $VAR(\epsilon_{ijk}) = \sigma^2$ that is independent over students i, schools j, and countries k. The first term on the right-hand side is a random intercept, decomposed as

$$\beta_{0jk} = \gamma_{000} + u_{0jk} + v_k,$$

where γ_{000} is the grand mean, and the other two terms are independent normally distributed error components with mean zero and variance $VAR(u_{0jk}) = \xi_0^2$ and $VAR(v_k) = \tau_0^2$. The regression coefficients β_1, β_2, and β_3 pertain to gender, the

number of books in the home, and the highest educational level attained by one of the parents, respectively. The regression coefficient for the parental involvement component is decomposed as

$$\beta_{4jk} = \gamma_{400} + u_{4jk} + u_{4k},$$

where γ_{400} is the average slope for the component over all countries and schools, and the two error terms have variances $VAR(u_{4jk}) = \xi_4^2$ and $VAR(u) = \tau_4^2$, respectively. Finally, random intercepts and random slopes are allowed to covary; at the country level this leads to a parameter $COV(v_k, u_{4k}) = \tau_{04}^2$.

The fixed effects models are obtained by setting $VAR(u) = \tau_4^2 = 0$, and the baseline models, model 0 and model 1, are obtained by removing the appropriate predictors.

To keep the model interpretable and relevant, only the components showing a meaningful effect in the fixed model were entered as covariates in the random-intercepts-and-slopes model.

We conducted all analyses using the software package Mplus version 7.11 (Muthén and Muthén 1998–2012).

5.2 Results of Latent Regression Models

The effects on student reading literacy were first modelled using the GPCM without correction for CDIF (Tables 5.1, 5.2, 5.3 and 5.4). Model 0 indicates that most of the variance in student achievement in reading literacy was situated at the student level (44 %; Table 5.1). Differences between countries were also considerable; 39 % of the variance could be accounted for by between-country differences. This was to be expected based on the large range of average country scores reported in the international report of PIRLS-2011 (Mullis et al. 2012).

As expected, model 1 indicated that gender and the two SES-indicators are important predictors of reading literacy. On average, girls outperformed boys by almost 13 points on the PIRLS-test. The number of books at home and the educational level of the parents are both positively related to reading achievement. The three background variables explain a considerable amount of variance; 40 % at school level and 62 % at country level. This suggests that a substantial part of the differences in achievement scores between PIRLS countries can be attributed to individual differences in student background characteristics.

In models 2A–2E, we explored the fixed effects of the different components of parental involvement, taken into account the effect of the three background variables (Table 5.2). The effect sizes of the background variables did not change noteworthy when the different components of parental involvement were included in the model. Parental report of literacy activities before their child starts in first grade, and helping with homework were both related to a student's reading literacy,

Table 5.1 Effects of student background characteristics and components of parental involvement on reading literacy achievement of grade 4 students in 41 PIRLS countries, using a random intercept model, without correction for cultural differences

Effects	Model 0		Model 1	
	Empty model		Student background characteristics	
	Effect	SE	Effect	SE
Fixed effects				
Intercept γ_{000}	525.49	7.12	541.75	6.30
Male difference β_1			−12.83	0.85
Books at home (low-high) β_2			8.55	0.49
Parental education (high-low) β_3			−13.50	0.77
Random effects				
Variance between students σ^2	4305.31 (44 %)	168.27	3878.90 (61 %)	165.14
Variance between schools ξ_0^2	1658.19 (17 %)	269.24	1001.59 (16 %)	177.31
Variance between countries τ_0^2	3887.53 (39 %)	0.30	1497.43 (23 %)	327.48
Explained by predictors				
At student level			10 %	
At school level			40 %	
At country level			62 %	

but each in a different way. Students of parents reporting spending more time on early literacy activities with their child showed higher achievement levels than those whose parents spent less time on these activities (the scale runs from high involvement to low involvement, therefore the effect in Table 5.2 appears as a negative score). This is in agreement with the results presented in the international report (Mullis et al. 2012). With regard to helping with homework, there is a negative relationship (this scale also runs from high involvement to low involvement, therefore the effect in Table 5.2 appears as a positive). Because of the large number of respondents recorded in the data (over 200,000 students), each relationship with achievement, even when very weak, is significant. Therefore, the relevance of these relationships was assessed in terms of changes in the achievement score if the predictor increased by one standard deviation. The standard deviation of early literacy activities was almost 1 (0.97, and thus excluded from this report). If parents' perceptions of the time they spent on early literacy activities increased one point (i.e., from average to one standard deviation above the mean), the score of the student on the PIRLS test increased by nine points. On a scale with a mean of 500 and a standard deviation of 100, this could be considered a small effect. This also applies to the negative association between helping with homework and reading achievement. The standard deviation of this component was also 1, and the reduction in student achievement was 9.3 if parents reported that they spent one standard deviation more in time in helping their child.

Table 5.2 Effects of student background characteristics and components of parental involvement on reading literacy achievement of grade 4 students in 41 PIRLS countries, random intercept model, without correction for cultural differences

Effects	Model 2A		Model 2B		Model 2C		Model 2D		Model 2E	
	Early literacy activities		Help with homework		Parent's view of school practices		Student's perception		School's perception	
	Effect	SE	Effect	SE	Effect	SE	Effect	SE	Effect	SE
Fixed effects										
Intercept γ_{000}	542.69	6.20	541.53	5.88	541.89	5.81	541.62	6.36	541.59	6.34
Male difference β_1	−11.73	0.83	−12.73	0.84	−12.73	0.86	−12.43	0.82	−12.78	0.85
Books at home (low-high) β_2	7.17	0.51	8.72	0.48	8.32	0.50	8.53	0.49	8.52	0.50
Parental education (high-low) β_3	−12.78	0.72	−13.51	0.79	−13.26	0.75	−13.53	0.75	−13.53	0.78
Early literacy activities before primary school (high-low) γ_{400}	−9.03	0.57								
Help with homework (high-low) γ_{500}			9.32	1.20						
Parent's view of school practices on parental involvement (high-low) γ_{600}					5.15	1.75				
Student's perception of parental involvement (high-low) γ_{700}							−2.64	0.68		
School's perception of parental involvement (high-low) γ_{800}									3.10	0.74

(continued)

Table 5.2 (continued)

Effects	Model 2A		Model 2B		Model 2C		Model 2D		Model 2E	
	Early literacy activities		Help with homework		Parent's view of school practices		Student's perception		School's perception	
	Effect	SE	Effect	SE	Effect	SE	Effect	SE	Effect	SE
Random effects										
Variance between Students σ^2	3825.10	161.13	3813.42	171.89	3849.27	164.10	3861.90	161.85	3886.61	162.96
Variance between schools ξ_0^2	977.68	173.60	993.75	176.31	977.54	172.83	977.60	172.91	978.71	168.07
Variance between countries τ_0^2	1469.37	319.06	1395.61	330.34	1301.80	315.67	1507.00	326.07	1545.41	346.23
Explained variance by predictors										
At student level	11 %		11 %		11 %		10 %		10 %	
At school level	41 %		41 %		41 %		41 %		41 %	
At country level	62 %		64 %		67 %		61 %		60 %	

Table 5.3 Effects of student background characteristics and parental involvement on reading literacy achievement of grade 4 students in 41 PIRLS countries, random intercept model, without correction for cultural differences

Effects	Model 0 Empty model		Model 1 Student background characteristics		Model 3 Parental involvement	
	Effect	SE	Effect	SE	Effect	SE
Fixed effects						
Intercept γ_{000}	525.49	7.12	541.75	6.30	541.82	5.46
Male difference β_1			−12.83	0.85	−10.45	0.81
Books at home (low-high) β_2			8.55	0.49	6.60	0.48
Parental education (high-low) β_3			−13.50	0.77	−12.15	0.70
Early literacy activities before primary school (high-low) γ_{400}					−12.91	0.56
Help with homework (high-low) γ_{500}					13.08	1.14
Parent's view of school practices on parental involvement (high-low) γ_{600}					4.12	1.50
Student's perception of parental involvement (high-low) γ_{700}					−4.09	0.49
School's perception of parental involvement (high-low) γ_{800}					2.7	0.75
Random effects						
Variance between students σ^2	4305.31 (44 %)	168.27	3878.90	165.14	3675.41	163.91
Variance between schools ξ_0^2	1658.19 (17 %)	269.24	1001.59	177.31	888.24	153.69
Variance between countries τ_0^2	3887.53 (39 %)	0.30	1497.43	327.48	1190.78	346.23
Explained variance by predictors						
At student level			11 %		15 %	
At school level			41 %		46 %	
At country level			64 %		69 %	

Table 5.4 Effects of student background characteristics and parental involvement on reading literacy achievement of grade 4 students in 41 PIRLS countries, random slopes model, without correction for cultural differences

Effects	Component 1		Component 2		Component 1 + 2	
	Early literacy activities		Help with homework			
	Estimate	SE	Estimate	SE	Estimate	SE
Fixed effects						
Intercept γ_{000}	542.01	6.42	538.54	5.99	538.67	5.78
Male difference β_1	−11.72	0.84	−12.52	0.84	−10.93	0.83
Books at home (low-high) β_2	7.19	0.51	8.45	0.45	6.58	0.45
Parental education (high-low) β_3	−12.76	0.72	−12.96	0.78	−11.97	0.72
Early literacy activities γ_{400}	−8.67	0.57			−12.66	0.68
Help with homework γ_{500}			11.75	1.38	15.15	1.40
Random effects						
Students σ^2	3789.18	160.03	3741.12	171.88	3618.13	16.048
Variance intercepts schools ξ_0^2	960.26	170.61	979.91	180.05	930.05	171.372
Variance slopes schools ξ_4^2	33.62	6.04			30.86	5.57
Variance slopes schools ξ_5^2			33.47	6.79	22.74	4.86
Variance intercepts countries τ_0^2	1386.39	292.43	1402.06	321.18	1447.48	371.59
Variance slopes countries τ_4^2	14.57	3.80			20.09	5.59
Covariance intercepts and slopes τ_{04}^2	27.25	28.10			41.40	30.56
Variance slopes countries τ_5^2			74.42	14.32	73.96	14.25
Covariance intercepts and slopes τ_{05}^2			109.23	47.79	134.36	48.19

At first glance, it seems that students whose parents had less positive views about school practices outperformed the classmates whose parents held more positive views. However, as the standard deviation of this scale was 1.6 and the effect size 5.1, the increase in scores was only three points. The same was true for students' perception of parental involvement (decrease of almost three points) and school perception of parental involvement (increase of three points). These are very small effects.

Models 2A–2E revealed that for each of the five components, the percentages suggested hardly any alteration in the variance, as compared with model 1. Thus, in model 3, we entered the fixed effects for all components of parental involvement simultaneously (Table 5.3). The influence of early literacy activities and helping with homework increased slightly when the effects of the other components were held constant; variance increased to 15 % at student level, 46 % at school level, and 69 % at country level.

The next step was to estimate two models with random slopes at country level for early literacy activities and helping with homework, to determine whether the effects of components of parental involvement differed across countries. While recognizing this is still open for discussion, we considered these two components as showing a small, but meaningful relation with reading achievement. We included the three background variables as fixed effects in the random model (Table 5.4).

For early literacy activities we see a very small increase in the average overall effect, from –9.0 in model 2A to –8.7 in model 4. The variance over countries is 14.57; relative to the total variance in the outcome variable this is very small, but relative to the effect of early literacy activities, the effect is clearly larger. Ninety-five percent of the range of the slope over countries lay roughly between –37.5 and 20.5. A covariance of 27.25 indicated there was a relationship between the intercept of a country and the steepness of the slope within a country. A positive covariance means that the relationship between parental involvement and reading achievement is stronger in countries that performed strongly in the PIRLS test; a negative covariance means that the association between the predictor and dependent variable becomes stronger as the country average of reading achievement decreases. The standard error of the covariance for early literacy activities was larger than the covariance, indicating that there was no relation between the intercept and slope. From the variance components and the covariance, we obtained a correlation of 0.07, which must be considered small.

For helping with homework, the average effect size increased from 9.3 to 11.8. The variance of the slope over countries was 74.42; 95 % of the range of the slope over countries lay between –134.0 and 159.5, which can be considered substantial. Further, a positive covariance of 109.23 led to a correlation of 0.34, which is also substantial. As this scale runs from high involvement to low involvement, although the effect reported was positive, in truth it is a negative effect (more help = lower achievement). The positive covariance suggests that this negative association of helping with homework with achievement was stronger in high-performing countries.

To assess the impact of CDIF, we replicated the last analysis (Table 5.4) with the a posteriori estimates of the latent student parameters from all five IRT models (Table 5.5). Estimates from all models were very close and never more than one standard deviation away from the estimates under the GPCM. We conclude that CDIF did not bias the inferences.

Table 5.5 Random-intercepts-and-slopes model for effects of student background characteristics and parental involvement on reading literacy achievement, without and with correction for cultural differences

Effects	GPCM		GPCM 10 % Split	GPCM 20 % Split	Random GPCM	Bi-factor GPCM
	Estimate	SE	Estimate	Estimate	Estimate	Estimate
Fixed effects						
Intercept γ_{000}	538.67	5.78	539.01	539.31	539.39	538.27
Male difference β_1	−10.93	0.83	−10.91	−10.90	−10.92	−11.11
Books at home (low-high) β_2	6.58	0.45	6.59	6.58	6.58	6.74
Parental education (high-low) β_3	−11.97	0.72	−11.99	−11.98	−11.98	−12.10
Early literacy activities γ_{400}	−12.66	0.68	−12.58	−12.56	−12.52	−12.106
Help with homework γ_{500}	15.15	1.40	15.05	15.51	15.59	13.084
Random effects						
Students σ^2	3618.13	16.048	3619.85	3615.16	3621.08	3628.78
Variance intercepts schools ξ_0^2	930.05	171.372	927.67	928.90	928.49	931.63
Variance slopes schools ξ_4^2	30.86	5.57	30.36	30.91	29.84	29.07
Variance slopes schools ξ_5^2	22.74	4.86	25.34	26.09	24.91	23.49
Variance intercepts countries τ_0^2	1447.48	371.59	1469.70	1418.14	1463.75	1347.119
Variance slopes countries τ_4^2	20.09	5.59	20.47	20.52	21.49	25.418
Covariance intercepts and slopes τ_{04}^2	41.40	30.56	40.32	32.40	40.43	21.73
Variance slopes countries τ_5^2	73.96	14.25	72.36	82.66	85.51	47.521
Covariance intercepts and slopes τ_{05}^2	134.36	48.19	138.30	132.56	145.50	101.94

References

Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 international results in reading*. Chestnut Hill, MA, USA: TIMSS&PIRLS International Study Center, Boston College.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide.* (7th ed.) Retrieved from http://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r3_web.pdf.

Von Davier, M, Gonzalez, E, & Mislevy, R. J. (2009). What are plausible values and why are they useful? In M. Von Davier & D. Hastedt (Eds.), *IERI. monograph series. issues and methodologies in large-scale assessments* (pp. 9–36). Hamburg: IER Institute.

6

Discussion and Conclusions

The main purpose of this research was to develop a psychometric framework to assess the relationship between parental involvement and reading literacy. The framework incorporates country specific differences, both at the item level and the scale level, to gain insight into cultural differences in the parental involvement component and its relation to student achievement in reading literacy. We conducted secondary analyses on the PIRLS-2011 data of 41 countries. A review of the research literature distinguished four dimensions of parental involvement: (1) home-based involvement from a parent perspective; (2) school-based involvement and home-school communication from a parent perspective; (3) home-based involvement from a student perspective; and (4) school-based involvement and home-school communication from a school perspective. Based on items available in the PIRLS data, the first dimension was split in two components: early literacy activities and helping with homework. IRT analyses provided item-by-country interactions indicating CDIF. The five components were first modeled using the unidimensional GPCM. Using these analyses, potential items with CDIF were identified and subsequently modeled using country-specific parameters for the 10 and 20 % most extreme interactions. These methods for identifying and modeling CDIF were compared with two other models. The first was the GPCM with random item parameters, where the variance of the parameters across countries provided an indication of possible CDIF. The second was a bi-factor GPCM where a country-specific covariance matrix gave an indication of the extent to which the scale loaded on the intended latent variable and the extent of loading on a country-specific dimension. Finally, multilevel analyses were conducted to explore the association between parental involvement and student achievement for all countries that participated in PIRLS-2011. A three-level (student, school and country) random intercept model was explored, as well as a random three-level model.

This study addressed three central research questions.

(1) *Which dimensions of parental involvement can be discerned and to what extent is there empirical evidence that these dimensions are related to student attainment?*

From the literature review, we constructed a framework of four dimensions that combined the different perspectives from which parental involvement can be perceived (i.e., parent, student and school perspectives) with frequently-mentioned dimensions of the construct, such as home-based involvement, school-based involvement, and home-school communication. The literature review and analysis of the meta-studies in particular, further indicated an overall positive effect of parental involvement on student achievement, but considering the individual dimensions separately leads to more variable interpretations. For example, the relation between helping with homework and student achievement is positive in some studies, but non-existent or negative in others (Hoover-Dempsey et al. 2001; Patall et al. 2008). These contrasting results were explained by the complexity of measurement and lack of agreement between scholars in measuring (the dimensions of) parental involvement and the use of single-source data. None of the meta-studies recognized other possible reasons for the variable results, such as cultural differences in how parents perceive parental involvement or cultural differences in their attitude towards their child's education. In this study, we wanted to examine whether these cultural differences could be identified and, if so, whether controlling for these differences revealed new information regarding the association between components of parental involvement and student's reading literacy.

(2) *To what extent are there any cultural differences (differences between countries) in the components that measure dimensions of parental involvement?*

We developed tools for the identification and modeling of CDIF that were based on five models: the GPCM, the GPCM with 10 and 20 % country-specific parameters, GPCM with random item parameters and the bi-factor GPCM. Firstly, we found that all models clearly and consistently supported the identification of CDIF. However, we also found the results obtained by the models varied. There was reasonable agreement for components 2 (helping with homework), 4 (student's perception of parental involvement) and 5 (school practices for parental involvement from a school perspective). The methods clearly disagreed for component 1 (early literacy activities) and for component 3 (school practices on parental involvement, parent perspective); the latter was likely because of the poor reliability of this component, probably due to the shortness of the instrument. Disagreement in the other four tests is because different aspects of model fit are assessed by the models. In fact, the method using residuals specifically targets uniform CDIF, while the bi-factor GPCM specifically targets non-uniform CDIF. In conclusion, practitioners should not rely on one model and one approach to investigate CDIF, but diversify in their methods.

Finally, and most importantly, analyzing the influence of CDIF on the estimates of country means and on the outcomes of latent regression analyses led to the conclusion that CDIF did not influence the results. Considering all the differing components of parental involvement, CDIF had no influence on these items within the PIRLS survey.

(3) *To what extent are the different dimensions of parental involvement related to student achievement in reading literacy, taking into account student background characteristics and differences between countries?*

The results of the three-level models with a random intercept showed that, controlled for student's gender and SES, and taking into account between-schools and between-countries variance, there is a rather weak but positive relationship between early literacy activities and student achievement in reading literacy at grade 4. This positive association supports the Dutch study of Kloosterman et al. (2011), who also found that early literacy activities were positively related to student reading achievement at primary school. We may here only confirm a positive association and cannot make any claims about causality, as PIRLS is cross-sectional. The results only indicate that other types of studies (experimental studies) measuring the real effects of early literacy activities on reading achievement are relevant, assuming that there is agreement among scholars in how these activities should be measured. Our analyses have shown that if written questionnaires are applied, the current PIRLS scale seems to work identically in a large number of countries and cultures. Further, the reliability of the first two components, early literacy activities and help with homework, meets the minimum standard for a survey of 0.70 within all countries. The scale for parental involvement from the school perspective often met this standard. The scale for parental involvement from the student perspective did not meet this standard, though it consisted of 15 items. Component 3 (school practices on parental involvement, parent perspective) turned out to be an unreliable scale, probably because it contained only three items. There is clearly margin for improvement in these last three scales.

The results of two meta-studies on homework involvement and its relation with student attainment were inconclusive (Hoover-Dempsey et al. 2001; Patall et al. 2008). The exploration of the PIRLS data revealed helping with homework had a small negative effect. This may be explained by the so-called reactive hypothesis, suggesting that parents tend to react with a higher level of involvement if their child is falling behind at school (McNeal Jr. 2012). McNeal Jr. (2012) and Cooper et al. (2000) suggested another explanation for this negative association; helping with homework might also interfere with learning if parents are not sufficiently equipped to help, if they are too eager (which affects the self-confidence of their child), or if their instruction is very different from the instruction of the teacher. Again, based on the PIRLS-2011 data, it is difficult to analyze how helping with homework affects student reading literacy, but further exploration of the indirect effect of helping with home via some measure of student self-confidence in reading would be useful. Another suggestion for future PIRLS studies would be to ask parents how confident they feel about helping their child with homework and whether they feel sufficiently informed about how reading is taught at school.

Both early literacy activities and helping with homework are home-based activities, confirming that what parents do at home with their child is important for student achievement. In this study we found school-based involvement from the perspective of the school (component 5) had negligible effect. As the constructed

scale for school-based involvement from the perspective of the parent (component 3) turned out to be unreliable, we are unable to draw valid conclusions for this component regarding its relation with student achievement.

Overall, the impact of parental involvement on reading literacy is not large. When all five components were entered into the model, it explained approximately 15 % at the student, 46 % at the school level, and 69 % at the country level. However, the impact differences across countries proved to be quite large, especially for helping with homework, where regression coefficients, with a mean value of 11.8, range over countries from −134.0 to 159.5. Finally, the country-level intercept and slopes for helping with homework have a substantial positive correlation of 0.34. In low-achieving PIRLS countries, the effect of helping with homework is smaller than in high-performing countries. This means that, in exploring the achievement effect of helping with homework, the educational context should be taken into account. The sometimes contradictory results of earlier studies on this subject (Hoover-Dempsey et al. 2001) may also be explained by such differing effects between countries.

Another explanation for the positive correlation between intercepts and slopes on the country level would be that, in low-achieving countries, parents' reading competency will also be low, so parents are themselves less able to read and hence provide effective support. However, it is beyond the possibilities of the present research to draw conclusions in this respect.

In PIRLS, the literacy test and background questionnaires are translated and adapted for each country. Considerable effort is devoted to guaranteeing the international validity of these instruments. For example, the translations and the layout of the instruments are thoroughly reviewed by independent verifiers, and all necessary adaptations are documented in detail. However, it is not unlikely that there are cultural differences in the way respondents interpreted some of the questionnaire items. The main purpose of this study was to establish whether there were any cultural differences in the measurement of parental involvement in PIRLS and, if so, whether correction for these differences led to different results with regard to its relation with reading literacy. Although some of the PIRLS scales for parental involvement may require improvements to increase their reliability, the overall conclusion is that these scales are internationally valid.

References

Cooper, H., Lindsay, J. J., & Nye, B. (2000). Homework in the home: how student, family, and parenting-style differences relate to the homework process. *Contemporary Educational Psychology, 25*(4), 464–487.

Hoover-Dempsey, K. V., Battiato, A. C., Walker, J. M. T., Reed, R. P., DeJong, J. M., & Jones, K. P. (2001). Parental involvement in homework. *Educational Psychologist, 36*(3), 195–209. doi:10.1207/S15326985EP3603_5.

Kloosterman, R., Notten, N., Tolsma, J., & Kraaykamp, G. (2011). The effects of parental reading socialization and early school involvement on children's academic performance: A panel study of primary school pupils in the Netherlands. *European Sociological Review, 27*(3), 291–306. doi:10.1093/Esr/Jcq007.

McNeal, R. B, Jr. (2012). Checking in or checking out? Investigating the parent involvement reactive hypothesis. *Journal of Educational Research, 105*(2), 79–89. doi:10.1080/00220671.2010.519410.

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). Parent involvement in homework: A research synthesis. *Review of Educational Research, 78*(4), 1039–1101. doi:10.3102/0034654308325185.

Appendix A
Technical Details on the Implementation of the Bi-factor Model

The estimation equations and expressions for the covariance matrix of the estimates are easily derived using Fisher's identity (Efron 1977; Louis 1982; Glas 1999). The identity plays an important role in the framework of the EM algorithm, which is an algorithm for finding the maximum of a likelihood marginalized over unobserved data. The principle can be summarized as follows. Let $L_0(\lambda)$ be the log-likelihood function of parameters λ given observed data x_0, and let $L_c(\lambda)$ be the log-likelihood function given both observed data x_0 and unobserved missing data x_m. The latter is called the complete data log-likelihood. The interest is in finding expressions for the first-order derivatives of $L_0(\lambda)$, say, the expressions for $L'_0(\lambda)$. Define the first-order derivatives with respect to the complete data log-likelihood as $L'_m(\lambda)$. Then Fisher's identity entails that $L'_0(\lambda)$ is equal to the expectation of $L'_m(\lambda)$ with respect to the posterior distribution of the missing data given the observed data, $p(x_m|x_0; \lambda)$, that is,

$$L'_0(\lambda) = E(L'_m(\lambda)|x_0, \lambda) = \int L'_m(\lambda)p(x_m|x_0, \lambda)dx_m.$$

To apply this framework to IRT, a very general definition of an IRT model is adopted. Assume an IRT model is defined by the probability of a response pattern x_n, which is a function of a, possibly vector-valued, student parameters θ_n, and item parameters a and b, which are item discrimination and item location parameters of an IRT model. So the IRT model is given by $p(x_n \mid \theta_n, a, b)$. Assume further that the student parameter θ_n has a normal density $N(\theta_n; \mu_{g(n)}, \Sigma_{g(n)})$ where, again, $g(n)$ is the country to which student n belongs. The key idea is to view the student parameters θ_n as missing data and the item and population parameters a, b, $\mu_{g(n)}$, and $\Sigma_{g(n)}$ as structural parameters λ to be estimated. Then the complete data log-likelihood for a student n is

$$L_{c(n)}(\lambda) = log\, p\big(\, \mathrm{x}_n |\, \theta_{n0}, \theta_{ng},\, a, b\,\big) + log\, N(\theta_{n0}, \theta_{ng};\, \mu_{g(n)}, \Sigma_{g(n)})$$

and so the estimation equations are given by

$$\frac{\partial L_0(\lambda)}{\partial a_{i0}} = \sum_n E\left(\theta_{n0}\left(\sum_{j=1}^{m_i} x_{nij} - p_{ij}(\theta_n)\right)\middle| x_n, \lambda\right) = 0$$

$$\frac{\partial L_0(\lambda)}{\partial a_{ig}} = \sum_{n|g(n)=g} E\left(\theta_{ng}\left(\sum_{j=1}^{m_i} x_{nij} - p_{ij}(\theta_n)\right)\middle| x_n, \lambda\right) = 0$$

$$\frac{\partial L_0(\lambda)}{\partial d_{ij}} = \sum_n \left[E\big(p_{ij}(\theta_n)\big|x_n, \lambda\big) - x_{nij}\right] = 0$$

$$\frac{\partial L_0(\lambda)}{\partial \mu_g} = \sum_{n|g(n)=g} \mu_g - E(\theta_{n0}|x_n, \lambda) = 0$$

$$\frac{\partial L_0(\lambda)}{\partial \sigma_g^2} = \sum_{n|g(n)=g} \sigma_g^2 - E\Big(\theta_{n0}^2 - \mu_g^2|x_n, \lambda\Big) = 0$$

where all the expectations are relative to the posterior distribution

$$p(\theta_{n0}, \theta_{ng}|x_n, \lambda) \propto p\big(\, \mathrm{x}_n |\, \theta_{n0}, \theta_{ng},\, a, b\,\big) N(\theta_{n0}, \theta_{ng}; \mu_g, \Sigma_g).$$

We undertook all calculations using the public domain software package MIRT (Glas 2010). The program uses the EM-algorithm to solve the estimation equations and Gaussian quadrature to evaluate the integrals.

References

Efron, B. (1977). Discussion on maximum likelihood from incomplete data via the EM algorithm (by A. Dempster, N. Liard, and D. Rubin). *Journal of the Royal Statistical Society, Series B, 39*, 1–38.

Glas, C. A. W. (1999). Modification indices for the 2-PL and the nominal response model. *Psychometrika, 64*(3), 273–294. doi:10.1007/bf02294296.

Glas, C. A. W. (2010). Multidimensional item response theory (MIRT), manual and computer program. Retrieved from http://www.utwente.nl/gw/omd/Medewerkers/temp_test/mirt_package.zip, http://www.utwente.nl/gw/omd/Medewerkers/temp_test/mirt-manual.pdf.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society, Series B, 44*, 226–233.