



ELSEVIER

Contents lists available at [ScienceDirect](#)

Journal of Adolescence

journal homepage: www.elsevier.com/locate/jado

Within- and between-person and group variance in behavior and beliefs in cross-cultural longitudinal data[☆]

Kirby Deater-Deckard^{a,*}, Jennifer Godwin^b, Jennifer E. Lansford^b, Dario Bacchini^c, Anna Silvia Bombi^d, Marc H. Bornstein^e, Lei Chang^f, Laura Di Giunta^d, Kenneth A. Dodge^b, Patrick S. Malone^b, Paul Oburu^g, Concetta Pastorelli^d, Ann T. Skinner^b, Emma Sorbring^h, Laurence Steinberg^{i,j}, Sombat Tapanya^k, Liane Peña Alampay^l, Liliana Maria Uribe Tirado^m, Arnaldo Zelliⁿ, Suha M. Al-Hassan^{o,p}

^a University of Massachusetts Amherst, Amherst, MA, USA

^b Duke University, Durham, NC, USA

^c Second University of Naples, Caserta, Italy

^d Università di Roma "La Sapienza", Rome, Italy

^e Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD, USA

^f University of Macau, China

^g Maseno University, Maseno, Kenya

^h University West, Trollhättan, Sweden

ⁱ Temple University, Philadelphia, PA, USA

^j King Abdulaziz University, Jeddah, Saudi Arabia

^k Chiang Mai University, Chiang Mai, Thailand

^l Ateneo de Manila University, Quezon City, Philippines

^m Universidad San Buenaventura, Medellín, Colombia

ⁿ University of Rome "Foro Italico", Rome, Italy

^o Hashemite University, Zarqa, Jordan

^p Emirates College for Advanced Education, Abu Dhabi, United Arab Emirates

ARTICLE INFO

Article history:

Received 14 November 2016

Received in revised form 9 June 2017

Accepted 11 June 2017

Available online xxx

Keywords:

Acculturation

Adolescence

Families

Multilevel modeling

ABSTRACT

This study grapples with what it means to be part of a cultural group, from a statistical modeling perspective. The method we present compares within- and between-cultural group variability, in behaviors in families. We demonstrate the method using a cross-cultural study of adolescent development and parenting, involving three biennial waves of longitudinal data from 1296 eight-year-olds and their parents (multiple cultures in nine countries). Family members completed surveys about parental negativity and positivity, child academic and social-emotional adjustment, and attitudes about parenting and adolescent behavior. Variance estimates were computed at the cultural group, person, and within-person level using multilevel models. Of the longitudinally consistent variance, most was within and not between cultural groups—although there was a wide range of between-group differences. This approach to quantifying cultural group variability may prove valuable when applied to quantitative studies of acculturation.

[☆] This article is part of a Special Issue entitled Explaining Positive Adaptation of Immigrant Youth across Cultures.

* Corresponding author. Department of Psychological and Brain Sciences, Tobin Hall, 135 Hicks Way, University of Massachusetts, Amherst, MA, 01003, USA

E-mail address: kdeaterdeck@umass.edu (K. Deater-Deckard).

<http://dx.doi.org/10.1016/j.adolescence.2017.06.002>

0140-1971/© 2017 The Foundation for Professionals in Services for Adolescents. Published by Elsevier Ltd. All rights reserved.

Please cite this article in press as: Deater-Deckard, K., et al., Within- and between-person and group variance in behavior and beliefs in cross-cultural longitudinal data, *Journal of Adolescence* (2017), <http://dx.doi.org/10.1016/j.adolescence.2017.06.002>

1. Introduction

Acculturation is a complex intra- and inter-personal process by which an individual who comes into contact with one or more additional cultures modifies her or his own behaviors, beliefs and self-construals (including identity) in response to the experience in ways that may be adaptive or maladaptive (Bornstein, 2017; Sam & Berry, 2010). Ward and Geeraert (2016) recently offered a highly articulated framework to conceptualize, measure, and test competing hypotheses about acculturation processes that involves the individual's heritage/home culture, the settlement/host culture, and global culture. Of relevance for the current paper, Ward and Geeraert (2016) placed a strong emphasis on family context and relationships, and psychosocial adjustment and well-being, as essential elements of understanding the acculturation process.

The emphasis in the current paper is not on defining, measuring, or statistically testing acculturation. Instead, our goal is to provide an example of an efficient and thorough quantitative approach that we hope will be useful in future studies of acculturation processes. Our example utilizes data analysis of a cross-cultural longitudinal study of child and adolescent development in 12 “cultural groups” in nine countries. (Note that we use the term “cultural group” in the present study to describe the samples that we assessed at each site around the globe; as we describe in the Method section, the sites varied in how culturally homogeneous and representative they were of the broader cultural and geographic context). The approach we present is useful because it permits researchers to address the relative proportions of variation in constructs of interest (in our case, parent and youth behavioral constructs as well as some constructs capturing beliefs)—between-group, within-group, and within-person (over time). Estimating and comparing these “pockets” of variability across a broad range of constructs can enhance researchers' understanding of what it means to be a member of a group, to be an individual within that group, and to change over time following experiences. Our data were not collected to develop new measurement approaches for acculturation research and do not capture change following inter-cultural interaction. However, we believe our longitudinal, cross-cultural study design and data analysis approach provide a useful opportunity to extrapolate how measuring within and between cultural variation may yield insights in future acculturation studies.

1.1. Quantifying behaviors and beliefs: between and within the “group”

The quantitative measurement and analysis of a group-level mean score on a dimensional variable is well reasoned and useful, *only to the extent that it informs the inference we make about the culturally based norm for that behavior or belief*. That is, the key assumption is that the mean score is the best quantitative representation of that group's norm on that behavior or belief. For example, the widely studied dimensions of interdependence and independence (Hofstede, 2001) are often quantified, and a mean score for each cultural group is computed that can be interpreted as the “norm” for that group (but see McSweeney, 2002, for a critique of the approach). These behavior and belief norms, as computed mean scores, can then be compared between groups, and the likelihood by chance of finding a difference between the groups' norms can be estimated (i.e., p , the probability of falsely rejecting the null hypothesis of no mean difference between groups). Looking at it another way, for scientists who do *not* presume that the mean score on a dimension of interest represents the norm for that group, it is unclear why quantitative tests of group mean differences would be useful at all.

The challenge for quantifying norms in behavior and beliefs is developing a common approach for estimating and interpreting between-group and within-group variation. Although there is increasing emphasis throughout the social and behavioral science disciplines on reporting and interpreting all aspects of the distributions of variables being studied, there is little consideration of having a common approach to *interpreting* the within-group variance that is found. For cross-cultural studies, this remains a major gap in much of the intracultural and intercultural research being conducted (Taras, Rowney, & Steel, 2009). As Taras et al. (2009) report, in a meta-analysis of studies that had quantified cultural variables as dimensional scores (e.g., traditionalism, group loyalty, family integration), only one in 50 of the over 500 studies that were examined described and interpreted within-group (i.e., culture) variation. Over half of the studies did not even report descriptive statistics on score dispersion (i.e., variance, standard deviation) in each group.

With respect to studies of acculturation, we searched the literature for relevant studies that reported and compared within- and between-group variation, but we were not able to identify any that did so using the approach we present. It is not that acculturation researchers are unaware of within-group variation. On the contrary, many studies of acculturation acknowledge and sometimes examine within-group differences, but it is uncommon for studies to present relative estimates of within- and between-group variation. For example, studies often make note of (and even quantify differences between) identifiable sub-groups within heterogeneous groups—for example, country of origin distinctions among Asian Americans or Latino Americans (e.g., Xia, Do, & Xie, 2013). However, that kind of an approach typically does not parse within-versus-between-group variance. A second way in which studies of acculturation have addressed within-group differences is by

examining individual differences in acculturation measures and processes within a group (e.g., integration, assimilation, separation, and marginalization; Sam & Berry, 2010). However, this approach also differs from the method we present, in that it does not quantify relative within- and between-group estimates of variation across a broad set of constructs of potential relevance and interest to acculturation processes. A third way in which studies of acculturation have addressed within-group variation is by using multilevel modeling as an analytic approach, to account for the clustering of individuals within families or different groups. A whole volume on this topic (edited by Van de Vijver, van Hemert, and Poortinga, 2014), provides an outstanding collection of chapters about multilevel models used in studying individuals and cultures, including a chapter on acculturation. However, the studies of acculturation using multilevel modeling that we have found, and our own prior cross-cultural multi-level modeling papers, did not parse within- and between-group variance in the way we are advocating. Thus, to our knowledge, the suggested modeling method and approach in the present study—to describe within- and between-level variability—has not been published previously.

The gap in consideration of relative comparisons of within-group and between-group variation arises in part from a closely related issue regarding the interpretation of effect sizes for between-group comparisons (see Matsumoto, Grissom, & Dinnel, 2001). For variables with distributions that vary between individuals within (as well as between) any given group, most of the research that is conducted examines the statistical significance of the mean difference between groups of interest, but usually does not report the effect size—and even less often actually interprets the effect size and the implications of that interpretation (Fritz, Morris, & Richler, 2012). The smaller the mean-difference effect size representing the average difference between two groups, the greater the overlap in the distributions of the two groups (expressed as an intra-class correlation or overlapping coefficient; e.g., Inman & Bradley, 1989; Snijders & Bosker, 2012).

Our view is that the field of cross-cultural and acculturation research can benefit from the application of commonly used methods that efficiently and thoroughly estimate and describe within and between group variances, as well as effect sizes. Although no method is “one size fits all”, having a small set of commonly applied approaches can promote standardization of how researchers present and describe within and between group variation and effect sizes (Barry et al., 2016; Vacha-Haase, Nilsson, Reetz, Lance, & Thompson, 2000). For cross-cultural quantitative studies of any dimensionally distributed construct, one essential question is: what is the relative magnitude of the between-group difference effect, in comparison to the within-group between-person variation, and the within-group within-person variation? The answer to this question has implications for acculturation research, with competing predictions arising. From a statistical standpoint, on the one hand one could argue that it would be relatively easier to detect changes in behavior and beliefs due to interactions in the new host culture, among constructs that consistently show lower within-group between person variance relative to between-group variance. On the other hand, greater within-culture variation in behaviors and beliefs could indicate that people are more receptive to changing their beliefs and behaviors during acculturation—that is, when within-culture variation is relatively high on a particular construct, it may be easier to acculturate on that construct.

Consideration of the number of groups being compared also has potential implications for acculturation researchers. Although group-by-group binary comparisons are standard in acculturation research for good reason (i.e., it is necessary to see how individual change occurs at the intersection of two or more cultures; Ward & Geeraert, 2016), these binary comparisons are more likely to produce biased results that are harder to replicate, compared to estimation of the same effects while incorporating multiple groups that intersect. In the current paper, we also recommend and demonstrate the use of *multi-group* multi-level statistical models whenever possible, because doing so provides a more complete and reliable estimation of the relative within- and between-group and person variances when multiple distinct groups are incorporated (Maas & Hox, 2005)—in particular, the estimation of the intra-class correlation (Castro, 2002), the statistical parameter that we use for making relative comparisons of within- and between-group variability.

A final key point is that it can be very helpful to expand the constellation of variables under consideration, to situate the key variables of interest for testing theories of acculturation within the broader context of relative within- and between-group variation, study wide. Our view is that it benefits rigor in theory testing to also include analysis of a broad variety of other dimensionally distributed variables—ranging from those that are conceptually related to acculturation, to those that have no expected relevance. This also includes consideration of possible differences in effects between informants (e.g., self, parent, friend, observer), and across time points to estimate effects with respect to change over time. A “study wide” approach provides a broader view of the patterns of relative differences in within- and between-group variances, and can reveal useful information for hypothesis generation as well as refinement of inferences about group norms and variation on the sub-set of constructs and variables of primary interest to the researcher.

2. Method

2.1. Participants

The sample for this cross-cultural study design included 1296 children (51% female) across three biennial waves (i.e., two years between waves) (age, $M = 8.28$ years, $SD = 0.64$ years in wave one) living in: China, Shanghai: $n = 120$; Colombia, Medellín: 108; Italy, Naples: 100; Italy, Rome: 102; Jordan, Zarqa: 114; Kenya, Kisumu: 100; Philippines, Manila: 120; Sweden, Trollhättan and Vänersborg: 103; Thailand, Chiang Mai: 119; and the United States, Durham: European American: 102, African American: 110, and Hispanic or Latino: 98. At the follow-up interviews four years after the initial interviews, 83% of the original sample provided data. Participants who provided follow-up data did not differ from the original sample with respect

to child gender, parents' marital status, or mothers' education. Informed consent was obtained, and human subjects ethics approval was granted by the university IRB for the co-authors at each site. Eight years of age was selected as the starting point for the study, primarily because the consensus of the research team was that it is at this age that most children are regarded as reliable informants using self-report instruments, stemming in part from cognitive development that accelerates between ages five to seven years (e.g., Sameroff & Haith, 1996). Sampling focused on including families from the majority ethnic group in each country; exceptions were Kenya where we sampled the Luo ethnic group (third largest, 13% of population) and the United States, where we sampled European American, African American, and Hispanic or Latino American families. To ensure economic diversity, we included students from private and public schools and from high- to low-income families, sampled in proportions representative of each recruitment area. Child age and gender did not vary across countries. Participants were recruited through letters sent from schools. Initial enrollment response rates varied across countries (from 24% to nearly 100%) primarily because of differences in the schools' roles in recruiting. For more details on the sample and recruitment procedures see Lansford et al. (2015).

These locations were selected to represent a wide range of geopolitical, sociocultural contexts. Countries in which we sampled families ranged from eighth to 145th ranking in the 2015 Human Development Index, with 0% to 43% of the population living on less than US\$1.25 per day from 2002–2012 (United Nations Development Program, 2015); Hofstede (2016) Individualism Score values from 13 to 91; 17% to 97% reporting that religion is important in daily life (Gallup, 2015); 4% to 94% of population enrolling in tertiary education and an average age at first marriage for women from 22 to 31 years (World Bank, 2016). There was marked ethnic and cultural group variation between sites and sometimes within sites (i.e., United States).

2.2. Procedure

The data summarized in the current analyses were based on youth and one or both parents' surveys that were collected in face-to-face or telephone interviews or written questionnaires following a rigorous procedure of forward- and back-translation to ensure the linguistic and conceptual equivalence of the measures across languages (Maxwell, 1996). In addition to translating the measures, translators noted items that did not translate well, were inappropriate for the participants, were culturally insensitive, or elicited multiple meanings and suggested improvements (Peña, 2007). Country coordinators and the translators reviewed the discrepant items and made appropriate modifications. Given the complexities of testing for and establishing measurement invariance on so many variables and in 12 cultural groups (see Huang et al., 2012, for an example involving just one variable), we decided not to analyze and present measurement invariance tests.

2.3. Measures

Parents and children completed the Parental Acceptance-Rejection/Control Questionnaire-Short Form (Rohner, 2005) in which respondents rated the frequency with which various parental behaviors were used (1 = *never or almost never*, 2 = *once a month*, 3 = *once a week*, or 4 = *every day*). Children completed the measure twice, once for each parent. The items are divided into parental behavior categories and averaged within category to create 5 scales: *Parental Warmth, Hostility, Rejection, Control, and Neglect*.

Parents and children also completed the Discipline Interview (Lansford et al., 2005) which captured the frequency of parental use (1 = *never*, 5 = *almost every day*) of various discipline behaviors (e.g., spanking, shaming, taking away privileges). Respondents also rated their approval of, perceived effectiveness of, and the community normativeness of each discipline technique. Harsh Physical and Harsh Verbal Discipline scales were created by averaging across dichotomous indicators of any harsh physical discipline (spanked, slapped, or hit the child; grabbed or shook the child; and threw something at the child) and any harsh verbal discipline (told the child he/she wouldn't love him/her; threatened to leave the child; and tried to scare the child into behaving), respectively. Mean scales were also created for Approval of, Effectiveness of, Normativeness of Harsh Physical and Harsh Verbal Punishment by averaging items in the corresponding categories. Parents also completed the Multiple Indicator Cluster Survey (UNICEF, 2005) which includes 6 dichotomous items indicating whether the parent administered different types of mild (spanked/hit child with bare hand, hit child on the bottom, slapped/hit child on the hand, shook the child) or harsh (slapped/hit child on the face, or beat the child) discipline. Averaging the three items within each (mild, harsh) yielded two corresponding scales.

All respondents completed the Child Behavior Checklist (CBCL; Achenbach, 1991) measuring the extent to which the child exhibited particular behaviors or emotions in the previous 6 months (0 = *not true*, 1 = *somewhat or sometimes true*, and 2 = *very often or often true*). Items were divided broadly into externalizing and internalizing problem behaviors as well as into more specific groups capturing aggressive, delinquent, anxious/depressed, somatic behavior. Items are summed within each of the 6 categories to create corresponding scales.

Children also completed the Behavior Frequency Scale (BFS), which consists of items compiled from Farrell, Danish, and Howard (1992), Crick and Bigbee (1998), and Orpinas and Frankowski (2001). Children indicated how often in the last 30 days they engaged in various behaviors (0 = *never* to 5 = *20 or more times*). Items were divided into categories and average to create 4 scales: Delinquent Behavior (e.g., *been suspended*); Non-Physical Aggression (e.g., *teased someone to make them angry*); Physical Aggression (e.g., *shoved and pushed another kid*); and Relational Aggression (e.g., *spread a false rumor about someone*).

3. Results

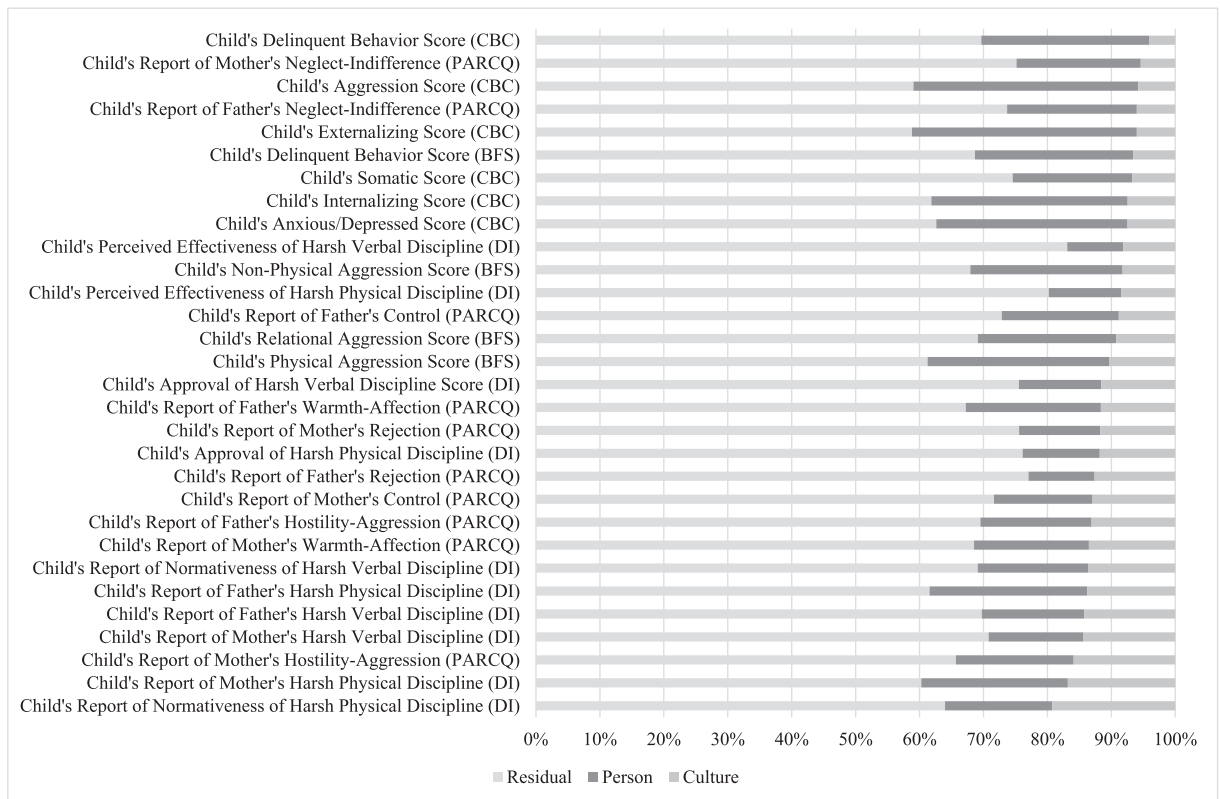
To estimate within- and between-cultural group variability in each of the instrument scales, we estimated separate models for youth report (see Table 1 and Fig. 1), mother report (Table 2 and Fig. 2), and father report (Table 3 and Fig. 3). The data structure was time point (three biennial waves) at level one, person at level two, and cultural group at level three; for each variable, we estimated a full information maximum likelihood multilevel model with a random intercept for person and culture using SAS PROC MIXED. Using the estimated variances (residual/within-person, between person within culture, and between culture), we calculated (1) the ICC_{1,2} (level 1 within level 2) which measures the percent of variance between persons and (2) ICC_{2,3} (level 2 within level 3) which measures the percent of the between person variance accounted for by culture. The statistical significance of ICC_{1,2} was assessed using a chi square test comparing the log likelihood of the “level one only” model to that of the two-level model (time points nested within person ignoring culture). The statistical significance of ICC_{2,3} was assessed using a chi square test comparing the log likelihood of the two-level model (time points nested within person ignoring culture) to that of the complete three-level model (time points nested within person nested within culture). To visually examine the division of variance across levels, we also generated bar graphs depicting the percentage of the estimated variance at each level (i.e., residual, person, and culture) (Hoffman, 2015).

For each child-reported variable, most of the variance was within person (Table 1). The ICC_{1,2}, the proportion of variance between persons, ranged from 0.17 to 0.41 and was always statistically significant at the 0.01 level. The proportion of variance between persons was less than one-third of the total, for 21 of the 30 child-reported scales. The ICC_{2,3}, the proportion of the between-person variance accounted for by culture, ranged from 0.13 to 0.55 and was also always statistically significant at the 0.01 level. For 12 of the 30 scales, culture accounted for less than one-third of the between-person variance. For 16 of the 30 scales, culture accounted for one-third to one-half of the between-person variance. For only two scales, culture accounted for over half of the between-person variance. Fig. 1 displays the variance estimates at each level of the model, with variables ranked from top to bottom of each figure in order of least to most between-cultural group variability. Variance at the between-culture level ranged from 4% to 19%, variance at the between-person within-culture level ranged from 9% to 35%, and the rest was residual within-person variance—by far the largest source of variability in the data. For 16 of the 30 scales, the percentage of variability found within-culture exceeded the between-culture variance by 5 percentage points or more. For 12

Table 1
Estimated Variances and Intra-class Correlations for Child-Reported Variables.

	Variances (SE)			ICC(1,2)	ICC(2,3)
	Residual	Person	Culture	(Chi Sq, 1 dof)	(Chi Sq, 1 dof)
Child's Delinquent Behavior Score (BFS)	0.04 (0.00)	0.01 (0.00)	0.00 (0.00)	0.31 (278.89)	0.07 (126.65)
Child's Non-Physical Aggression Score (BFS)	0.18 (0.01)	0.06 (0.01)	0.02 (0.01)	0.32 (305.72)	0.08 (163.80)
Child's Physical Aggression Score (BFS)	0.08 (0.00)	0.04 (0.00)	0.01 (0.01)	0.39 (455.01)	0.10 (203.67)
Child's Relational Aggression Score (BFS)	0.14 (0.00)	0.04 (0.00)	0.02 (0.01)	0.31 (279.68)	0.09 (185.88)
Child's Aggression Score (CBC)	16.6 (0.49)	9.88 (0.66)	1.63 (0.72)	0.41 (511.05)	0.06 (99.22)
Child's Anxious/Depressed Score (CBC)	14.04 (0.42)	6.71 (0.5)	1.67 (0.73)	0.37 (418.41)	0.07 (140.06)
Child's Delinquent Behavior Score (CBC)	2.98 (0.09)	1.12 (0.09)	0.17 (0.08)	0.3 (270.07)	0.04 (73.00)
Child's Externalizing Score (CBC)	27.1 (0.81)	16.2 (1.08)	2.77 (1.23)	0.41 (513.90)	0.06 (104.49)
Child's Internalizing Score (CBC)	43.86 (1.31)	21.76 (1.59)	5.28 (2.30)	0.38 (428.82)	0.07 (140.31)
Child's Somatic Score (CBC)	7.97 (0.24)	2 (0.22)	0.72 (0.31)	0.25 (188.11)	0.07 (147.01)
Child's Approval of Harsh Physical Discipline (DI)	0.30 (0.01)	0.05 (0.01)	0.05 (0.02)	0.24 (171.53)	0.12 (289.95)
Child's Approval of Harsh Verbal Discipline (DI)	0.28 (0.01)	0.05 (0.01)	0.04 (0.02)	0.24 (182.24)	0.12 (286.75)
Child's Perceived Effectiveness of Harsh Physical Discipline (DI)	0.51 (0.02)	0.07 (0.01)	0.05 (0.02)	0.20 (114.98)	0.08 (209.14)
Child's Perceived Effectiveness of Harsh Verbal Discipline (DI)	0.60 (0.02)	0.06 (0.01)	0.06 (0.03)	0.17 (87.66)	0.08 (206.15)
Child's Report of Father's Harsh Physical Discipline (DI)	0.06 (0.00)	0.02 (0.00)	0.01 (0.01)	0.38 (412.11)	0.14 (274.36)
Child's Report of Father's Harsh Verbal Discipline (DI)	0.04 (0.00)	0.01 (0.00)	0.01 (0.00)	0.30 (263.57)	0.14 (322.05)
Child's Report of Mother's Harsh Physical Discipline (DI)	0.06 (0.00)	0.02 (0.00)	0.02 (0.01)	0.40 (481.25)	0.17 (363.83)
Child's Report of Mother's Harsh Verbal Discipline (DI)	0.05 (0.00)	0.01 (0.00)	0.01 (0.00)	0.29 (265.52)	0.14 (347.27)
Child's Report of Normativeness of Harsh Physical Discipline (DI)	0.61 (0.02)	0.16 (0.02)	0.18 (0.08)	0.36 (387.61)	0.19 (455.96)
Child's Report of Normativeness of Harsh Verbal Discipline (DI)	0.47 (0.01)	0.12 (0.01)	0.09 (0.04)	0.31 (284.36)	0.14 (318.63)
Child's Report of Father's Control (PARCQ)	0.27 (0.01)	0.07 (0.01)	0.03 (0.01)	0.27 (199.03)	0.09 (187.08)
Child's Report of Father's Hostility-Aggression (PARCQ)	0.15 (0.00)	0.04 (0.00)	0.03 (0.01)	0.30 (254.30)	0.13 (288.37)
Child's Report of Father's Neglect-Indifference (PARCQ)	0.21 (0.01)	0.06 (0.01)	0.02 (0.01)	0.26 (187.27)	0.06 (113.87)
Child's Report of Father's Rejection (PARCQ)	0.14 (0.00)	0.02 (0.00)	0.02 (0.01)	0.23 (160.50)	0.13 (318.69)
Child's Report of Father's Warmth-Affection (PARCQ)	0.22 (0.01)	0.07 (0.01)	0.04 (0.02)	0.33 (283.17)	0.12 (235.11)
Child's Report of Mother's Control (PARCQ)	0.22 (0.01)	0.05 (0.01)	0.04 (0.02)	0.28 (230.68)	0.13 (309.17)
Child's Report of Mother's Hostility-Aggression (PARCQ)	0.15 (0.00)	0.04 (0.00)	0.04 (0.01)	0.34 (353.95)	0.16 (366.23)
Child's Report of Mother's Neglect-Indifference (PARCQ)	0.19 (0.01)	0.05 (0.01)	0.01 (0.01)	0.25 (187.94)	0.05 (110.61)
Child's Report of Mother's Rejection (PARCQ)	0.16 (0.00)	0.03 (0.00)	0.02 (0.01)	0.24 (195.78)	0.12 (297.59)
Child's Report of Mother's Warmth-Affection (PARCQ)	0.18 (0.01)	0.05 (0.00)	0.03 (0.01)	0.31 (295.24)	0.14 (309.37)

Note: SE = standard error, ICC(1,2) = interclass correlation for level 1 within level 2, ICC(2,3) = interclass correlation for level 2 within level 3, Chi Sq = chi-square test statistics, dof = degrees of freedom, CBC=Child Behavior Checklist, DI = Discipline Interview, BFS=Behavior Frequency Scale, PARCQ=Parental Acceptance-Rejection/Control Questionnaire.



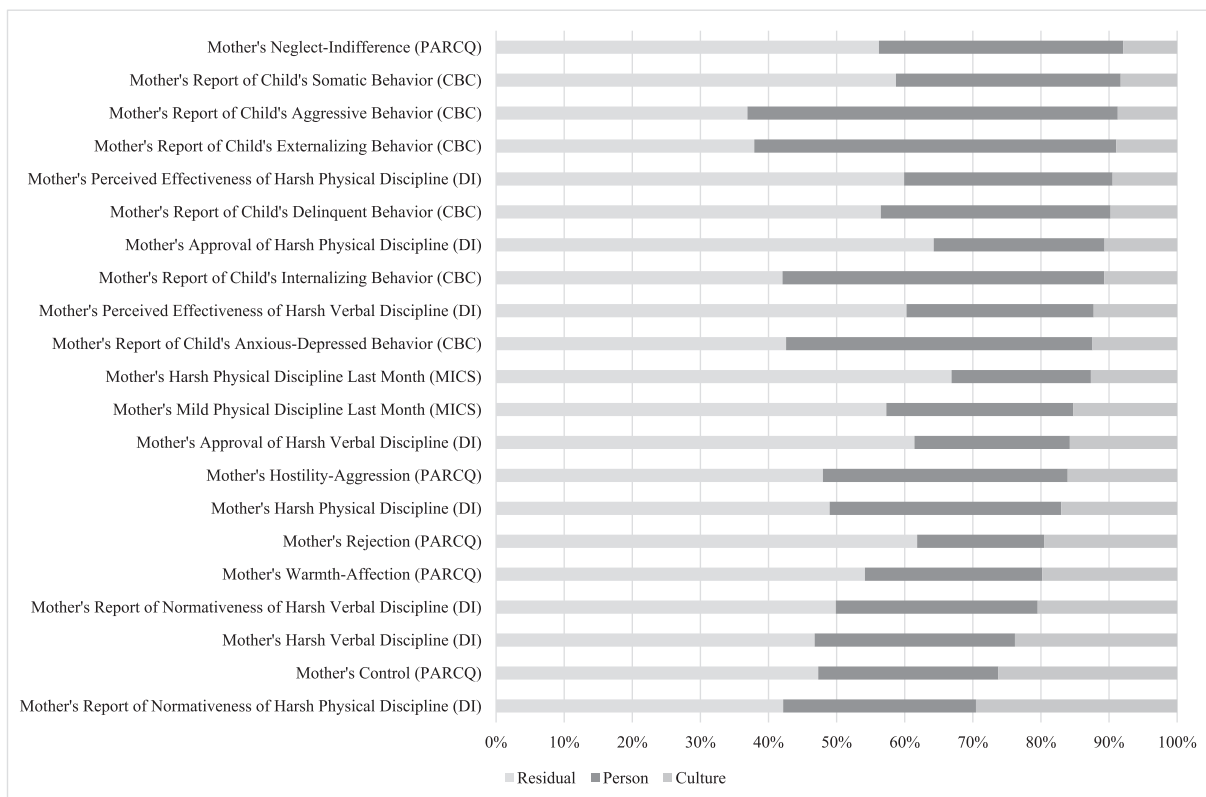
Note: CBC=Child Behavior Checklist, DI=Discipline Interview, BFS=Behavior Frequency Scale, PARCQ=Parental Acceptance-Rejection/Control Questionnaire.

Fig. 1. Percentage of Variance at Each Level for Child-Reported Variables. Note: CBC=Child Behavior Checklist, DI = Discipline Interview, BFS=Behavior Frequency Scale, PARCQ=Parental Acceptance-Rejection/Control Questionnaire.

Table 2
Estimated Variances and Intra-class Correlations for Mother-Reported Variables.

	Variances (SE)			ICC(1,2)	ICC(2,3)
	Residual	Person	Culture	(Chi Sq, 1 dof)	(Chi Sq, 1 dof)
Mother's Report of Child's Aggressive Behavior (CBC)	13.5 (0.41)	19.89 (1.01)	3.19 (1.39)	0.63 (1287.9)	0.09 (123.2)
Mother's Report of Child's Anxious-Depressed Behavior (CBC)	7.39 (0.22)	7.80 (0.43)	2.16 (0.92)	0.57 (1014.7)	0.12 (197.0)
Mother's Report of Child's Delinquent Behavior (CBC)	2.44 (0.07)	1.46 (0.10)	0.42 (0.18)	0.43 (578.5)	0.10 (182.6)
Mother's Report of Child's Externalizing Behavior (CBC)	21.66 (0.65)	30.37 (1.55)	5.10 (2.22)	0.62 (1245.9)	0.09 (129.7)
Mother's Report of Child's Internalizing Behavior (CBC)	21.94 (0.66)	24.65 (1.34)	5.57 (2.40)	0.58 (1050.3)	0.11 (162.0)
Mother's Report of Child's Somatic Behavior (CBC)	3.02 (0.09)	1.70 (0.12)	0.43 (0.18)	0.41 (513.5)	0.08 (149.7)
Mother's Approval of Harsh Physical Discipline (DI)	0.27 (0.01)	0.10 (0.01)	0.04 (0.02)	0.36 (383.0)	0.11 (209.4)
Mother's Approval of Harsh Verbal Discipline (DI)	0.27 (0.01)	0.10 (0.01)	0.07 (0.03)	0.39 (439.9)	0.16 (333.7)
Mother's Harsh Physical Discipline (DI)	0.05 (0.00)	0.04 (0.00)	0.02 (0.01)	0.51 (802.2)	0.17 (317.6)
Mother's Harsh Verbal Discipline (DI)	0.05 (0.00)	0.03 (0.00)	0.02 (0.01)	0.53 (857.1)	0.24 (480.8)
Mother's Perceived Effectiveness of Harsh Physical Discipline (DI)	0.28 (0.01)	0.14 (0.01)	0.04 (0.02)	0.40 (469.6)	0.09 (170.2)
Mother's Perceived Effectiveness of Harsh Verbal Discipline (DI)	0.32 (0.01)	0.15 (0.01)	0.07 (0.03)	0.40 (455.8)	0.12 (238.1)
Mother's Report of Normativeness of Harsh Physical Discipline (DI)	0.44 (0.01)	0.29 (0.02)	0.30 (0.13)	0.58 (1019.3)	0.30 (607.3)
Mother's Report of Normativeness of Harsh Verbal Discipline (DI)	0.39 (0.01)	0.23 (0.02)	0.16 (0.07)	0.50 (735.1)	0.20 (402.0)
Mother's Harsh Physical Discipline Last Month (MICS)	0.02 (0.00)	0.01 (0.00)	0.00 (0.00)	0.33 (326.7)	0.13 (264.9)
Mother's Mild Physical Discipline Last Month (MICS)	0.04 (0.00)	0.02 (0.00)	0.01 (0.00)	0.43 (531.9)	0.15 (298.4)
Mother's Control (PARCQ)	0.15 (0.00)	0.09 (0.01)	0.09 (0.04)	0.53 (822.5)	0.26 (548.3)
Mother's Hostility-Aggression (PARCQ)	0.07 (0.00)	0.05 (0.00)	0.02 (0.01)	0.52 (835.8)	0.16 (291.9)
Mother's Neglect-Indifference (PARCQ)	0.1 (0.00)	0.06 (0.00)	0.01 (0.01)	0.44 (565.0)	0.08 (137.2)
Mother's Rejection (PARCQ)	0.07 (0.00)	0.02 (0.00)	0.02 (0.01)	0.38 (439.6)	0.20 (453.9)
Mother's Warmth-Affection (PARCQ)	0.09 (0.00)	0.04 (0.00)	0.03 (0.01)	0.46 (629.9)	0.20 (417.7)

Note: SE = standard error, ICC(1,2) = interclass correlation for level 1 within level 2, ICC(2,3) = interclass correlation for level 2 within level 3, Chi Sq = chi-square test statistics, dof = degrees of freedom, CBC=Child Behavior Checklist, DI = Discipline Interview, MICS = Multiple Indicator Cluster Survey, PARCQ=Parental Acceptance-Rejection/Control Questionnaire.



Note: CBC=Child Behavior Checklist, DI=Discipline Interview, MICS=Multiple Indicator Cluster Survey, PARCQ=Parental Acceptance-Rejection/Control Questionnaire.

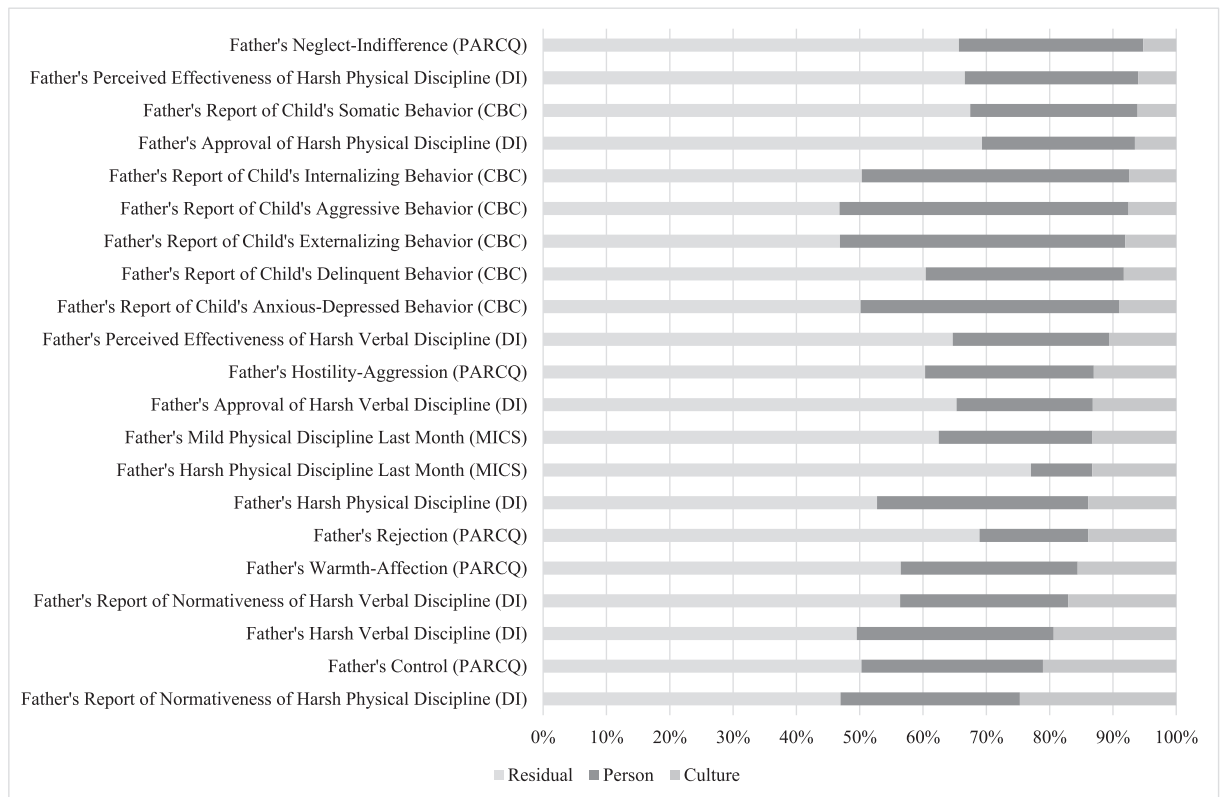
Fig. 2. Percentage of Variance at Each Level for Mother-Reported Variables. Note: CBC=Child Behavior Checklist, DI = Discipline Interview, MICS = Multiple Indicator Cluster Survey, PARCQ=Parental Acceptance-Rejection/Control Questionnaire.

Table 3
Estimated Variances and Intra-class Correlations for Father-Reported Variables.

	Variances (SE)			ICC(1,2)	ICC(2,3)
	Residual	Person	Culture	(Chi Sq, 1 dof)	(Chi Sq, 1 dof)
Father's Report of Child's Aggressive Behavior (CBC)	14.44 (0.49)	14.08 (0.88)	2.32 (1.04)	0.53 (662.4)	0.14 (95.1)
Father's Report of Child's Anxious-Depressed Behavior (CBC)	7.01 (0.24)	5.71 (0.38)	1.27 (0.56)	0.50 (562.8)	0.18 (116.8)
Father's Report of Child's Delinquent Behavior (CBC)	2.38 (0.08)	1.23 (0.10)	0.33 (0.14)	0.40 (385.2)	0.21 (137.6)
Father's Report of Child's Externalizing Behavior (CBC)	22.59 (0.77)	21.72 (1.36)	3.88 (1.71)	0.53 (671.3)	0.15 (105.8)
Father's Report of Child's Internalizing Behavior (CBC)	21.78 (0.74)	18.27 (1.19)	3.22 (1.44)	0.50 (565.5)	0.15 (92.3)
Father's Report of Child's Somatic Behavior (CBC)	2.86 (0.10)	1.12 (0.10)	0.26 (0.12)	0.33 (241.5)	0.19 (94.1)
Father's Approval of Harsh Physical Discipline (DI)	0.28 (0.01)	0.10 (0.01)	0.03 (0.01)	0.31 (208.1)	0.21 (102.1)
Father's Approval of Harsh Verbal Discipline (DI)	0.30 (0.01)	0.10 (0.01)	0.06 (0.03)	0.35 (246.6)	0.38 (222.2)
Father's Harsh Physical Discipline (DI)	0.05 (0.00)	0.03 (0.00)	0.01 (0.01)	0.47 (526.3)	0.29 (221.1)
Father's Harsh Verbal Discipline (DI)	0.04 (0.00)	0.03 (0.00)	0.02 (0.01)	0.50 (580.4)	0.38 (311.1)
Father's Perceived Effectiveness of Harsh Physical Discipline (DI)	0.30 (0.01)	0.12 (0.01)	0.03 (0.01)	0.33 (228.9)	0.18 (84.8)
Father's Perceived Effectiveness of Harsh Verbal Discipline (DI)	0.34 (0.01)	0.13 (0.01)	0.06 (0.02)	0.35 (250.8)	0.30 (163.3)
Father's Report of Normativeness of Harsh Physical Discipline (DI)	0.43 (0.01)	0.26 (0.02)	0.23 (0.09)	0.53 (674.2)	0.47 (437.9)
Father's Report of Normativeness of Harsh Verbal Discipline (DI)	0.43 (0.01)	0.20 (0.02)	0.13 (0.06)	0.44 (418.7)	0.39 (285.4)
Father's Harsh Physical Discipline Last Month (MICS)	0.02 (0.00)	0.00 (0.00)	0.00 (0.00)	0.23 (109.0)	0.58 (309.4)
Father's Mild Physical Discipline Last Month (MICS)	0.04 (0.00)	0.02 (0.00)	0.01 (0.00)	0.37 (304.9)	0.35 (232.5)
Father's Control (PARCQ)	0.17 (0.01)	0.10 (0.01)	0.07 (0.03)	0.50 (576.5)	0.42 (358.2)
Father's Hostility-Aggression (PARCQ)	0.09 (0.00)	0.04 (0.00)	0.02 (0.01)	0.40 (372.7)	0.33 (223.8)
Father's Neglect-Indifference (PARCQ)	0.13 (0.00)	0.06 (0.00)	0.01 (0.00)	0.34 (265.2)	0.15 (73.1)
Father's Rejection (PARCQ)	0.09 (0.00)	0.02 (0.00)	0.02 (0.01)	0.31 (231.1)	0.45 (279.8)
Father's Warmth-Affection (PARCQ)	0.13 (0.00)	0.06 (0.01)	0.04 (0.01)	0.43 (422.9)	0.36 (279.6)

Note: SE = standard error, ICC(1,2) = interclass correlation for level 1 within level 2, ICC(2,3) = interclass correlation for level 2 within level 3, Chi Sq = chi-square test statistics, dof = degrees of freedom, CBC=Child Behavior Checklist, DI = Discipline Interview, MICS = Multiple Indicator Cluster Survey, PARCQ=Parental Acceptance-Rejection/Control Questionnaire.

Please cite this article in press as: Deater-Deckard, K., et al., Within- and between-person and group variance in behavior and beliefs in cross-cultural longitudinal data, *Journal of Adolescence* (2017), <http://dx.doi.org/10.1016/j.adolescence.2017.06.002>



Note: CBC=Child Behavior Checklist, DI=Discipline Interview, MICS=Multiple Indicator Cluster Survey, PARCQ=Parental Acceptance-Rejection/Control Questionnaire.

Fig. 3. Percentage of Variance at Each Level for Father-Reported Variables. Note: CBC=Child Behavior Checklist, DI = Discipline Interview, MICS = Multiple Indicator Cluster Survey, PARCQ=Parental Acceptance-Rejection/Control Questionnaire.

of the 30 scales, the percentage of variability within-culture exceeded the between-culture variance by less than five percentage points. For only two scales, the percentage of variability within-culture was less than the between-culture variance.

For each mother-reported variable, most of the variance was within person (Table 2). The $ICC_{1,2}$, the proportion of variance between persons, ranges from 0.33 to 0.63 and was always statistically significant at the 0.01 level. The proportion of variance between-persons was less than 0.5, for 11 of the 21 mother-reported scales. The $ICC_{2,3}$, the proportion of the between person variance accounted for by culture, ranged from 0.14 to 0.51 and was always statistically significant at the 0.01 level. For 11 of the 21 scales, culture accounted for less than one-third of the between-person variance. For eight of the 21 scales, culture accounted for one-third to one-half of the between-person variance. For only two scales, culture accounted for over half of the between-person variance. As seen in Fig. 2, the between-culture variance ranged from 8% to 30%, variance at the between-person/within-culture level ranged from 19% to 54%, and the rest was residual within-person variance. For 18 of the 21 scales, the percentage of variability within-culture exceeded the between-culture variance by five percentage points or more. For one of the 21 scales, the percentage of variability within-culture exceeded that found between-culture, by less than five percentage points. For only two scales, the percentage of variability within-culture was less than the between-culture variance.

Finally, for each father-reported variable, most of the variance was within person (Table 3). The $ICC_{1,2}$, the proportion of variance between persons, ranged from 0.23 to 0.53 and was always statistically significant at the 0.01 level. The proportion of variance between-persons was less than 0.5, for 17 of the 21 father-reported scales, and less than 0.33 for 4 of those scales. The $ICC_{2,3}$, the proportion of the between-person variance accounted for by culture, ranged from 0.14 to 0.58 and was always statistically significant at the 0.01 level. For 12 of the 21 scales, culture accounted for less than one-third of the between-person variance. For eight of the 21 scales, culture accounted for one-third to one-half of the between-person variance. For only one scale, culture accounted for over half of the between-person variance. As shown in Fig. 3, the between-culture variance ranged from 5% to 24%, variance at the between-person/within-culture level ranged from 9% to 44%, and the rest was residual within-person variance. For 18 of the 21 scales, the percentage of variability within-culture exceeded that between-culture by five percentage points or more. For two of the 21 scales, the percentage of variability within-culture exceeded that between-culture by less than five percentage points. For only one scale, the percentage of variability within-culture was less than that found between cultures.

4. Discussion

In the current analysis of a large longitudinal cross-cultural study of youth and their parents in nine countries, we examined variation in a wide range of adolescent adjustment and parenting environment indicators. Adolescence is a time of remarkable growth in independence and risk-taking that can promote healthy development or endanger youth, depending on the antecedents and consequences of their decision making and actions—a pattern that is seen in many cultures around the globe (Steinberg et al., 2017). This developmental growth in autonomy, agency, and risk-taking also has differential effects on adolescent adjustment and health depending on features of parenting and the parent-adolescent relationship. Much of the research in this area has focused on parental monitoring, control, and autonomy support, along with levels of acceptance versus rejection; this literature spans quantitative and qualitative studies of many cultural groups, some in cross-cultural designs (for a recent review see Lansford, 2017).

One of the broad findings emerging from this literature is that while there are cultural-universal and culture-specific effects linking parenting and adolescent adjustment *on average*, there is noteworthy within-culture variations in relevant parenting environments and indicators of adolescent functioning that are often overlooked. In the present paper, we used a multilevel modeling approach to examine variability. More specifically, we computed the relative proportions of total variance that were found between cultural groups, between people within each cultural group, and residual variance within each person, arising from within-person change in scores over time (which also included measurement error). It is important to emphasize that this method derives variance estimates that reflect *relative* variances within and between groups—not actual *absolute* effect size estimates based on some population-wide standard. The purpose of our demonstrated approach is to estimate and describe the pockets of variability within and between cultural groups, relative to each other. Though not an acculturation study, the present longitudinal cross-cultural study illustrates how analyzing and describing within- and between-culture variation in behaviors and beliefs may offer new insights in future quantitative acculturation studies.

In this cross-cultural study of parenting environments and youth behavioral and emotional adjustment, we estimated these effects for every variable on which we had data, across three biennial waves of data when the children were eight, 10, and 12 years of age on average. Using this approach, we found that overall, most of the variation in the sample across the study variables was within-person variability, followed by between-people within-group variability, and lastly between-group or culture variability. Specifically, more than half of the observed between-person variation was accounted for by cultural group for only two of thirty variables (youth report), two of twenty-one variables (mother report), and one of twenty-one variables (father report). Still, there was noteworthy between-culture variation spanning the variables we studied. The proportion of variance at the level of cultural group ranged across variables from 4% to 19% for youth reports, 8% to 30% for mothers' reports, and 5% to 24% for fathers' reports. These findings illustrate the importance of utilizing efficient multilevel modeling methods generally (Van de Vijver et al., 2014), and conducting thorough examination of within- and between-group variation across multiple variables—including within-person variability, in the case of longitudinal designs.

Aside from the broad patterns of differences in between/within group variability just described and presented in detail in Results, there also were clear patterns of consistency across all three informants that are worth noting. For the descriptions and interpretations that follow, it is most useful to examine the three figures. Bear in mind, this summary is offered just to exemplify the analysis method, not because the content of these variables has particular implications for acculturation research.

First, for all three informants, beliefs about the Normativeness of Harsh Discipline had the largest between-cultural group variation, yet for all three informants the perceived effectiveness of harsh discipline had some of the smallest between-cultural group variance estimates. Beyond this, there were no other apparent similarities across informants in the relative ordering of variables from smallest to largest variance estimates at the cultural group level. Second, there also was a clear pattern of similarity across informants with respect to magnitude of between-person variability. For youth, mother, and father reports alike, some of the largest between-person variance estimates were found for reports of externalizing and internalizing behavior problems and their pertinent subcomponents (e.g., aggression, delinquency, anxiety/depression, somatic problems), which suggests that these scales are capturing the most temporally and situationally stable behaviors that we assessed. Another consistency across the three informants was that the smallest or nearly smallest between-person variance estimates were found for parental rejection of the child. This suggests that much of the variation in that construct was either between groups, or within children and families reflecting changes over time. Third, with respect to residual within-person variability, again there was a consistent pattern across the three informants. The largest or nearly largest within-person residual variance estimates were found for parents' reports of the child's actual exposure to physical punishment and discipline. The smallest or nearly smallest within-person variance estimates were found for reports of the child's externalizing symptoms and aggressive behavior.

4.1. Application to acculturation research

Our broad goal in conducting and presenting the findings from this cross-cultural study is to exemplify one way to elaborate multilevel model methods, to fully describe between- and within-group variability in constructs of interest. But how would such an approach be applied, in studies of acculturation? To begin answering this question, we offer two ideas. First, there are a variety of acculturation constructs or variables that are measured (e.g., integration, assimilation,

separation, marginalization; Sam & Berry, 2010) that show variation within as well as between cultural groups, and that can be analyzed using our method. Doing so within two or more groups, and analyzing within-person change over time as well, would be a particularly fruitful application of this method. The groups might be geographically separated members of the same cultural group (e.g., acculturation for a cultural group in two distinct, separated locations), or multiple groups who are interacting in the same place (e.g., acculturation of two or three distinct cultural groups in one geographic location). The method we present would allow researchers to more fully describe and understand how much of the variability they are observing in acculturation variables of interest, is within the groups relative to between the groups. Such comparisons could lead to new insights about how acculturation happens among members of those groups in those geographic areas.

Second, the approach we have presented can provide essential information about within- and between-group variability in cultural groups that are not interacting with each other, but that informs acculturation research involving the same groups that are interacting in other geographic locations. For instance, we examined within and between group variability in mainland Chinese youth's behavioral and emotional problems, in the context of a large multi-nation study including a sample of youth in the United States. Our results could inform future acculturation research examining behavioral and emotional adjustment of Chinese youth in the United States (and, for that matter, any of the other countries we have studied) following arrival to those new cultural contexts. Our basic viewpoint is that it is helpful, whenever possible, to sample broadly across multiple country and culture groups, and across informants, even when striving to understand the comparison or intersection of just two groups who are involved in an acculturation process (Ward & Geeraert, 2016). It is worth noting that there is considerable knowledge already about likely informant differences (e.g., parents versus youth) in acculturation processes and rates of change—another level of variability (i.e., within-family) in perceptions that could be incorporated into extensions of the present modeling approach. More broadly, we wish to emphasize the value of multilevel statistical models that take full advantage of the measured differences at multiple levels of variability that exist between and within culture groups, and within and between families, given the demonstrated precision and accuracy of estimation using such methods (Castro, 2002; Maas & Hox, 2005).

4.2. Limitations and conclusions

There are several limitations to consider. First, we caution that the samples from each site were not nationally representative (although they were locally representative of the cities from which they were recruited). Second, the analyses were limited to the measures for which we had multi-informant reports for three points in time. Thus, the specific patterns of between-culture and between-person within-culture distributions of variance would depend on the variables under investigation. Third, and perhaps most importantly, ours is not a study of acculturation processes or acculturation variables, so the study does not address directly the ways in which acculturation functions.

With these caveats in mind, the general conclusion based on the approach we have demonstrated in the current paper is this: Depending on the acculturation constructs or variables in question, there may be wide differentiation between constructs in terms of which ones show particularly large between-group (relative to within-group) variation. This leads to two competing predictions. First, within a defined universe of conceptually related and pertinent variables under consideration (given one's theory or specific hypothesis regarding an acculturation process), it will be more difficult to quantify statistically significant acculturation changes for constructs that are known already to have substantial within-person and within-group variability compared to between-group variability. Put another way, if most of the variability between people is found within the two or three groups being studied, it will be that much more challenging to demonstrate empirically that within-person change can be attributed to a between-group transactional process as that individual acculturates. Accordingly, it would be easiest to detect *systematic* (i.e., attributable to the interaction between the person and the new host culture) within-person and within-group change over time, for those constructs or variables that show the largest between-group variability relative to within-group variability. Second, in contrast, greater within-culture variation in behaviors and beliefs could indicate greater flexibility for changing beliefs and behavior arising from interactions with people from a different culture. That is, individuals may more readily acculturate on constructs that show the most heterogeneity within groups.

In conclusion, we have demonstrated the simultaneous examination of nine samples in different countries—but one need not stop there. For instance, “simultaneous triangulation” has greatly informed rigorous mixed and multi-method designs, to find information that converges or diverges across the methods (e.g., Jick, 1979; Morse, 2003). In much the same way, simultaneous triangulation can be applied *within a quantitative method*. In the case of the current study, the between- and within-group variation next could be estimated and compared across different pairings of country or cultural groups—and, any given pairing could be interpreted in light of the overall patterns of within and between country or culture group variation. Such an approach, if used in similarly structured datasets examining acculturation, would have the potential to inform whether and how a pattern of acculturative change that is observed for a specific pair of cultures is alike or distinct from a pattern of change seen in other pairings (see also implications of testing this and other aspects of the “specificity principle” articulated by Bornstein, 2017). The same logic would apply when comparing between and within group variance estimates between informants, given that effects that differ by informant can be instrumental to inferences and hypothesis generation (De Los Reyes, Thomas, Goodman, & Kundey, 2013).

Acknowledgements

We are grateful to the participating families and research staff. This research was supported by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, HD054805. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or NICHD.

References

- Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL 14-18, YSR, and TRF profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Barry, A. E., Szucs, L. E., Reyes, J. V., Ji, Q., Wilson, K. L., & Thompson, B. (2016). Failure to Report effect sizes: the handling of quantitative results in published health education and behavior research. *Health Education & Behavior*, 43(5), 518–527.
- Bornstein, M. H. (2017). The specificity principle in acculturation science. *Perspectives on Psychological Science*, 12(1), 3–45.
- Castro, S. L. (2002). Data analytic methods for the analysis of multilevel questions. *The Leadership Quarterly*, 13(1), 69–93.
- Crick, N. R., & Bigbee, M. A. (1998). Relational and overt forms of peer victimization: A multi-informant approach. *Journal of Consulting and Clinical Psychology*, 66(2), 337–347.
- De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kunder, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, 9, 123–149.
- Farrell, A. D., Danish, S. J., & Howard, C. W. (1992). Risk factors for drug use in urban adolescents: Identification and cross-validation. *American Journal of Community Psychology*, 20(3), 263–286.
- Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1), 2–18.
- Gallup (2015). <http://www.gallup.com/poll/1690/religion.aspx>. Accessed 8 November 2016.
- Hoffman, L. (2015). *Longitudinal Analysis: Modeling within-person fluctuation and change*. New York, NY: Routledge.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, CA: SAGE Publications.
- Hofstede, G. (2016). <https://geert-hofstede.com/countries.html>. Accessed 8 November 2016.
- Huang, L., Malone, P. S., Lansford, J. E., Deater-Deckard, K., Di Giunta, L., Bombi, A. S., et al. (2012). Measurement invariance of mother reports of discipline in different cultural contexts. *Family Science*, 2, 212–219.
- Inman, H. F., & Bradley, E. L. (1989). The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. *Communications in Statistics - Theory and Methods*, 18(10), 3851–3874.
- Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24(4), 602–611.
- Lansford, J. E. (2017). An international perspective on parenting and children's adjustment. In N. J. Cabrera, & B. Leyendecker (Eds.), *Handbook of positive development of minority children and youth*. New York, NY: Springer.
- Lansford, J. E., Chang, L., Dodge, K. A., Malone, P. S., Oburu, P., ... Quinn, N. (2005). Physical discipline and children's adjustment: Cultural normativeness as a moderator. *Child Development*, 76, 1234–1246.
- Lansford, J. E., Godwin, J., Uribe Tirado, L. M., Zelli, A., Al-Hassan, S. M., ... Alampay, L. P. (2015). Individual, family, and culture level contributions to child physical abuse and neglect: A longitudinal study in nine countries. *Development and Psychopathology*, 27(4), 1417–1428.
- Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, 1(3), 86–92.
- Matsumoto, D., Grissom, R. J., & Dinnel, D. L. (2001). Do between-culture differences really mean that people are Different?: a look at some measures of cultural effect size. *Journal of Cross-Cultural Psychology*, 32(4), 478–490.
- Maxwell, B. (1996). Translation and cultural adaptation of the survey instruments. In M. O. Martin, & D. L. Kelly (Eds.), *Third international mathematics and science study (TIMSS) technical report, volume 1: Design and development*. Chestnut Hill, MA: Boston College.
- McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith—a failure of analysis. *Human Relations*, 55(1), 89–118.
- Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.
- Orpinas, P., & Frankowski, R. (2001). The aggression scale: A self-report measure of aggressive behavior for young adolescents. *The Journal of Early Adolescence*, 21(1), 50–67.
- Peña, E. D. (2007). Lost in Translation: Methodological considerations in cross-cultural research. *Child Development*, 78(4), 1255–1264.
- Rohner, R. P. (2005). *Handbook for the study of parental acceptance and rejection* (4th ed.). Storrs, CT: Rohner Research Publications.
- Sam, D. L., & Berry, J. W. (2010). Acculturation: When individuals and groups of different cultural backgrounds meet. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 5(4), 472–481.
- Sameroff, A. J., & Haith, M. M. (1996). *The five to seven year shift: The age of reason and responsibility*. Chicago, IL: The University of Chicago Press.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage.
- Steinberg, L., Icenogle, G., Shulman, E. P., Breiner, K., Chein, J., ... Takash, H. M. (2017). Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation. *Developmental Science*. <http://dx.doi.org/10.1111/desc.12532>. Advance online publication.
- Taras, V., Rowney, J., & Steel, P. (2009). Half a century of measuring culture: Review of approaches, challenges, and limitations based on the analysis of 121 instruments for quantifying culture. *Journal of International Management*, 15(4), 357–373.
- United Nations Children's Fund/UNICEF. (2005). *Multiple indicator cluster survey manual 2005: Monitoring the situation of children and women*. New York, NY: UNICEF.
- United Nations Development Program/UNDP. (2015). *Human development report 2015: Work for human development*. New York, NY: UNDP.
- Vacha-Haase, T., Nilsson, J. E., Reetz, D. R., Lance, T. S., & Thompson, B. (2000). Reporting practices and APA editorial policies regarding statistical significance and effect size. *Theory & Psychology*, 10(3), 413–425.
- Van de Vijver, F. J., Van Hemert, D. A., & Poortinga, Y. H. (2014). *Multilevel analysis of individuals and cultures*. New York: Psychology Press.
- Ward, C., & Geeraert, N. (2016). Advancing acculturation theory and research: The acculturation process in its ecological context. *Current Opinion in Psychology*, 8, 98–104.
- World Bank (2015). <http://www.worldbank.org/en/publication/wdr2015>. Accessed 8 November 2016.
- Xia, Y., Do, K. A., & Xie, X. (2013). *The adjustment of Asian American families to the U.S. context: The ecology of strengths and stress*. Faculty Publications, Department of Child, Youth, and Family Studies. Paper 80.