

Empirical Article

Culture, Temperament, and the “Difficult Child”: A Study in Seven Western Cultures*

Charles M. Super¹, Giovanna Axia², Sara Harkness¹, Barbara Welles-Nyström³,
Piotr Olaf Zylicz⁴, Parminder Parmar⁵, Sabrina Bonichini², Moisés Ríos Bermúdez⁶,
Ughetta Moscardino², Violet Kolar⁷, Jesús Palacios⁶, Andrzej Elias⁴, and Harry
McGurk⁷

¹University of Connecticut, USA

²University of Padua, Italy

³Karolinska Institute, Sweden

⁴Warsaw School of Social Psychology, Poland

⁵The Pennsylvania State University, USA

⁶University of Seville, Spain

⁷Australian Institute of Family Studies, Australia

This study explores parental ethnotheories of children’s temperament through mothers’ responses to McDevitt and Carey’s Behavioral Style Questionnaire (1978) for 299 children aged 3 to 8 years and interviews with their parents, in Australia, Italy, the Netherlands, Poland, Spain, Sweden, and the United States. We first established a standardized, “derived etic” version of the questionnaire with adequate reliability for 8 of the original 9 scales. Cross-cultural comparisons of the scales’ means showed generally similar perceptions of children’s behavior. However, intercorrelations of the mean ratings with each other and with global “difficulty,” as presented through multidimensional scaling, showed both general tendencies and culture-specific patterns, which are further illustrated by parental discourse about “difficult” children in each sample. The findings underline the importance of parental ethnotheories for shaping the expression of temperament in development.

Keywords: culture, temperament, child development, parents’ ideas

* The work reported here was supported by grants from the Spencer Foundation and the National Science Foundation (award number BNS 83-11084), and a Fulbright Senior Fellowship to the first author; all statements made and views expressed are the sole responsibility of the authors. The authors especially appreciate the participation of the many families in this research, which was part of the International Study of Parents, Children, and Schools. The following research teams conducted that study, contributing essential effort and ideas: in Australia, Harry McGurk, Violet Kolar, and Grace Soriano; in Italy, Giovanna Axia, Sabrina Bonichini, and Ughetta Moscardino; in the Netherlands, Sara Harkness, Charles Super, Hesje Andersson, Marjolijn Blom, Hanneke Vrijenhoek Diekhuis, Jarissa Dijkstra, Karina Grijzen, Mariël Jacobs, Lieke Meijer, Edwin Mons, Reina Rijdsdam, Nathalie van Tijen, and Ellen van der Vlugt, with much appreciated advice from Gedolph and Rita Kohnstamm; in Poland, Andrzej Elias, Agnieszka Carrasco-Zylicz, and Piotr Olaf Zylicz; in Spain, Jesús Palacios, Victoria Hidalgo, María Carmen Moreno, Alfredo Oliva, and Moisés Ríos Bermúdez; in Sweden, Barbara Welles-Nyström and Caroline Tovatt; in the United States of America, Sara Harkness, Charles Super, Xin Feng, Marcia Hughes, Archana Khattar, Amy Miller, Beth Muller, and Parminder Parmar.

The concept of temperament has gained prominence in recent years as a persuasive and powerful framework for understanding individual differences among children. Described as the “how” of behavior, temperament refers to constitutional differences in the ways that individuals respond to their environments. These differences, which are conceptualized as stable across different situations and developmental periods, are believed to be based in aspects of physiology and neuroanatomy that underlie broad patterns of activity and response. Such differences, in interaction with the growing person’s experience, are thought ultimately to play an important role in the development of personality, mental health, and adaptation to the social environment (Bates, 1998; Chess & Thomas, 1984; Rothbart, Chew, & Gartstein, 2001).

Although there is general agreement about the importance of temperament for behavior and development, there has been less consensus among researchers about how to conceptualize its structure. Thomas and Chess’s formulation, which sparked developments in this field in the last half-century, postulated nine dimensions of temperament: Activity, Regularity, Adaptability, Approach/Withdrawal, Intensity, Mood, Persistence, Distractibility, and Threshold (Thomas & Chess, 1977). Although these dimensions were conceptualized as independent, psychometric analyses of questionnaires used to assess temperament dimensions have generally failed to support a nine-factor solution (see review by Martin, Wisenbaker & Huttunen, 1994). Thus, other approaches have proposed a smaller number of dimensions that correspond to the solution provided by factor analysis of questionnaire responses. Additionally, some researchers have focused on a more narrow set of temperamental features that can be indexed through multiple behavioral and physiological measures (Kagan, Snidman, Arcus & Reznick, 1994; Rothbart, Ellis, & Posner, 2004).

Thomas and Chess’s formulation of temperament is distinctive in its origins in clinical/ethnographic research with parents and children, and this fact was a determining force in their insistence on the validity of the nine dimensions despite a lack of psychometric support. According to this view, dimensions such as Approach and Adaptability, which may meld into one factor when questionnaires are subjected to factor analysis, are nonetheless separable because they are meaningful as distinct concepts to parents and are therefore clinically useful. An early product of this approach, based on Thomas and Chess’s experience with middle-class, mostly Jewish families in the New York Longitudinal Study, was the definition of the “difficult infant” in terms of a particular constellation of temperament characteristics which, when found together, would be challenging to parents and other caregivers: negative mood, irregularity in biological rhythms (e.g. in sleeping and eating), intensity of response, withdrawal from new situations, and slowness in adaptation to changes in the environment. A subsequent study of mostly working-class Puerto Rican families in New York, however, showed a somewhat different constellation of temperament characteristics that parents found difficult (Korn & Gannon, 1983). Irregularity did not present problems for these families, for example, because infants’ bedtimes were apt to be contingent on whatever was happening with other members of the family. On the other hand, high activity was problematic for them, living as they did in

smaller apartments. Thus, the concept of the “difficult infant” soon gave way to the idea of “goodness of fit” between the individual and the environment. This idea has remained a central tenet of temperament theory (Super & Harkness, 1993), and other approaches to dispositional differences among children have demonstrated the same principle. Chen and his colleagues, for example, have documented in a series of studies that shyness or behavioral inhibition in children is related to maternal approval, peer acceptance, and academic adjustment in China, whereas the opposite holds true in Canada (Chen, Hastings, Rubin, Chen, Cen, & Stewart, 1998; Chen, Rubin, & Li, 1995; Chen, Rubin, & Sun, 1992). Contemporary research has assumed a coordination of temperament, environment, and developmental trajectory (van den Boom & Hoeksma, 1994), and the clinical literature acknowledges that perceived difficult temperament is a risk factor in children’s mental health and psychological well-being (Carey & McDevitt, 1989; Maziade, Caron, Côté, Mérette, Bernier, Laplante, Boutin, & Thivierge, 1990; Rutter, 1989).

Despite the inherent focus on parents’ ratings in much temperament research, there has been little attention to parents’ own ideas about children’s temperament as they may relate to the questionnaire-based assessments. This poses special issues for the cross-cultural study of children’s temperament, because parents might respond cogently to a questionnaire even though their local concepts differ from those presumably embedded in the items (Nakagawa & Sukigara, 2005; Shwalb, Shwalb, & Shoji, 1996). Comparative research shows ample cultural variability in parental responses to the nine-dimension questionnaires; each of the nine dimensions has at least one significant group difference reported, but Approach is the most frequently cited. The Australian Temperament Project (Prior, Garino, Sanson & Oberklaid, 1987), for example, studied longitudinally native-born Anglo-Australian and immigrant children and found, in general, that the Anglo-Australian children were rated as more approaching, more adaptable, and less distractible than immigrant children. Related studies in the countries of origin demonstrate some degree of cultural continuity with the Australian immigrant groups (Axia, Prior, & Carelli, 1992; Kyrios, Prior, Oberklaid, & Demetriou, 1989).

Variability in average ratings of children’s temperaments in different cultures can be interpreted in at least three ways (Hsu, 1981; Prior, Kyrios & Oberklaid, 1986; Super & Harkness, 1994). First, it is possible that children’s temperaments really do vary across cultural groups, for a variety of reasons including genetic and early (including prenatal) environmental factors. Second, parents in different cultures may have different expectations and values regarding children’s temperament, which are reflected in their responses to the questionnaires; thus, particular kinds of child behavior may have dramatically different meaning in different cultures. Third, parental ethnotheories of child temperament may shape parenting practices in such a way as to influence the development and expression of child temperament over time; this may be particularly true as the ethnotheories are shared with other important social groups and institutions. We believe that all three of these explanations are plausible. Nevertheless, a central assumption underlying the work presented here is that children’s temperament is best understood within the cultural context in which their inner dispositions become manifest.

Like the concept of temperament, approaches to the study of “culture” in relation to children’s development vary considerably. In cross-cultural research, including that cited above, there is a straightforward assumption that samples of respondents in different parts of the world, or who originate from different parts of the world, belong to different “cultures;” in other words, culture in this paradigm is a kind of social address. Within this tradition, sample differences in response are attributed to culture although there may be an effort to “control” for other predictor variables such as social class, education, age, employment status, or rural versus urban residence. A contrasting approach, reflecting anthropological traditions, is to consider “culture” as a package of customs, beliefs and practices characteristic of a particular population that is usually but not always geographically defined. Within this paradigm, requiring that different cultural samples—say, from rural Africa and the metropolitan U.S.—be matched on variables such as socioeconomic status is pointless since such features as education and wealth are integral aspects of the way of life of a particular community (Rogoff & Angelillo, 2002). A third approach, currently popular in psychology although not anthropology, is to seek simplification and clarity in cross-cultural comparisons through the creation of higher-order categories such as “individualistic” or “collectivistic,” assign various cultural groups to these categories, and explain sample differences with reference to the supposed characteristics of these general orientations.

Our own approach in the present study draws from both anthropological and psychological traditions. Like anthropologists, we define “culture” as a constellation of particular beliefs and practices that function as a system, and that are shared by a group of people who recognize each other as members of a definable (usually geographic) entity. Our concept of culture reflects, in particular, cognitive anthropological approaches that draw attention to the central role of cultural models (belief systems) as organizers of individual experience and social life (e.g., D’Andrade & Strauss, 1992). At the same time, we recognize that most “cultures” (in the common sense of populations who share such characteristics) have fuzzy boundaries and that in fact many cultural beliefs and customs have arrived at their current location through a process of migration or dispersion from other cultures. On that basis, we would therefore expect some cultural groups to be more similar to each other than are others; for example, we would expect more similarity among cultural communities in various parts of continental Europe than between any of them and communities in sub-Saharan Africa or east Asia; but we make no firm assumptions about the specific patterning of similarity or difference across cultures.

The question of how widely one could draw the boundaries around each culturally based, geographically defined sample without encountering more internal variability than homogeneity (that is, how “representative” a given cultural sample is of the larger socio-political entity to which it belongs) is of less theoretical interest to us because our central concern is with how local cultures, as imperfectly but powerfully integrated systems, interact with the developing child. This context can be defined in terms of the child’s “developmental niche,” a construct encompassing the child’s

physical and social settings of daily life, the customs and practices of child care, and the psychology of the caretakers (Harkness & Super, 2005; Super & Harkness, 1986). Parents' cultural beliefs, in particular, may have an important influence on parents' evaluation of their children's temperaments. At the cross-cultural level, these parental ethnotheories—and the patterns of child care and daily life that instantiate them—may create different possibilities for “goodness of fit” between children of varying temperaments and their culturally structured developmental niches. They may also lead to socialization practices that ultimately shape children's patterns of response along particular dimensions of temperament.

The present study focuses on the interplay among children's temperamental dispositions, parents' understanding of specific behaviors and concepts of individual variation, and parental evaluations of their children's behavioral styles and “difficulty.” Like many others, we see the issue of psychosocial risk to center on how well the child's disposition fits with the demands and opportunities presented by the environment. The cross-cultural approach used here allows us to explore the idea that children's temperamental “difficultness” is at least in part a function of their niche.

Methods

Samples

As a cross-cultural study, the present report is unusual in that all the samples are drawn from within the Western world. Our rationale in constructing such a sample base was that a great deal of variability among Western cultures has gone largely undocumented in the developmental literature. Contrary to the claim of a “Western mind,” we sought to discover both similarities and differences—themes and variations—among cultural communities within Europe and the European diaspora (Harkness & Super, 2005; Harkness, Super, & Pai, 2000). Because the countries from which our particular communities were chosen share a broadly similar socioeconomic structure, however, it made sense to control for a few sources of possible variation such as poverty or major health issues, as well age, sex, and birth order (first versus later) of the target children. However, we let other sources of variability such as maternal employment or use of non-parental childcare vary freely as these tend to be integral aspects of different cultural communities.

The data are drawn from the International Study of Parents, Children, and Schools, a collaborative study of parental ethnotheories and practices related to children's development and learning in seven countries: Australia, Italy, The Netherlands, Poland, Spain, Sweden, and the United States. These countries were selected to sample the broad East-West and North-South variation within the European continent, as well as the British diaspora. The research team in each country was headed by (or in the Dutch case, advised by) an established, local social scientist; this allowed deci-

sions regarding methodology and interpretation to be made with equitable consultation between the “insiders” and “outsiders” of any one culture, a process that seems critical for comparative research (Harkness, Moscardino, Ríos Bermúdez, Zylicz, Welles-Nyström, Blom, Parmar, Axia, & Super, 2006). Study sites in each country were chosen to be broadly representative of a local middle-class population in a city or region. Because the central purpose of the project was to identify shared cultural models and their role in parenting, the samples were restricted to families in which both parents were native-born and native speakers of the local language.

With the exception of the US, samples in each country were recruited in one geographic area, using a variety of methods including snowball sampling, recruitment through parent-teacher associations, and advertisements in child care centers, health clinics, and schools. Parents who expressed interest in participating were called by a member of the research team to check for appropriateness in terms of both the general parameters mentioned above and the age of the “target child.” The nature and purpose of the study was then explained by a member of the research team, and normal consent procedures in each site were followed.

The Australian sample consists of Anglo-Celtic families residing in the metropolitan area of Melbourne, who were recruited through public announcements. The Italian families all resided in Padua, and were recruited through their membership in a parents’ civic organization. “Bloemenheim,” the Dutch research site, lies in the densely populated area between Amsterdam and the Hague. Families were recruited through social networks based initially in a neighborhood school, and thereafter through the snowball method as well as through announcements in schools and medical practices. Families in the Polish sample resided in a suburb of Warsaw, and were recruited through informal networks as well as the local primary school. The Spanish families were all residents of Seville, and were recruited through schools, national health centers, and child care centers in a particular section of the city. Families in the Swedish sample lived in a suburb of Stockholm. They were recruited through informal and school-based networks. Finally, the US sample combines data from three subsamples: families in the Boston area who were recruited through a health maintenance organization, families living in central Pennsylvania, and families in central and eastern Connecticut. The second and third subsamples were recruited through schools, community centers, and health care practices.

We have hypothesized that parental ethnotheories are constructed and elaborated in the context of actual practices with specific children, and that the age of the target child will influence parental behavior and discourse (Harkness & Super, 1996). We therefore specified subsamples within each community to tap several key ages during the preschool and school-transitional years, specifically choosing target ages that do not fall at known periods of rapid change (and thus, presumably, minimizing within-subsample variance due to developmental shifts).

The complete study sample at all sites consists of families with a child in one of the five target age groups: 6 months, 18 months, 3 years, 4.5 years, and 7 to 8 years. At each age, there are at least twelve children, approximately evenly divided by sex and

birth order (first-born versus later-born). This report presents data from mothers of children in the oldest three age groups, for whom the same temperament questionnaire was used (the 3-year-olds in Australia and the 7-year-olds in Italy are omitted because they were given a different questionnaire). Statistical analysis of demographic measures¹ on the families with sufficient temperament data to be included in the present analyses indicates significant differences in maternal employment, maternal and paternal age and education, number of children in the family, and religious orientation ($p < .01$). Some of these differences reflect obvious national variation (e.g. high Roman Catholic affiliation in Italy, Spain, and Poland), but others are more particular to our samples (e.g. mothers in our Italian sample average 2 more years education than the other groups, except the U.S.). As demonstrated below, however, these demographic differences have no bearing on our conclusions.

Procedures

Mothers were asked to complete the Behavioral Style Questionnaire (BSQ) developed by McDevitt and Carey (1978) for assessing temperament in children aged 3 to 7 years. Designed to be filled out by parents, the questionnaire contains 100 items that describe a wide variety of specific behaviors, which are to be rated for their perceived frequency of occurrence on a six-point scale, from “almost never” (1) to “almost always” (6). For example, the Activity dimension includes items such as “The child enjoys games that involve running or jumping,” and “The child sits calmly while watching TV or listening to music.” Relevant items are then summed to construct scales for the nine temperament dimensions. In addition, the last page of the questionnaire asks the respondent for global impressions of the child along each of the dimensions, as well as how “difficult” the child is. In order to explore how parental perceptions of their children’s temperament as derived from questionnaire results might compare with what parents actually say about their children naturalistically, we also examined discourse transcribed from semi-structured interviews, focusing particularly on interviews with parents who had rated their child relatively high (within their sample) on “difficulty.”

Results

Deriving a Culturally Optimized Version of the Behavioral Style Questionnaire

The specificity of behavioral items in the BSQ is an important asset, as it allows us to formulate operational definitions of temperament dimensions as perceived by parents in the seven cultural groups. Our approach to analysis of these data is driven both by the necessity of dealing with small samples, and by a readiness to accept at face value sample differences in the inter-item correlations. As a first step, we calculated the internal reliability (Cronbach’s Alpha) for each of the nine tempera-

ment dimensions separately for each site, starting with the original scale definitions provided by McDevitt and Carey (1978). We then successively dropped poorly fitting items to maximize the internal cohesion of each scale, separately for each site. This procedure enabled us to convert an "etic" instrument (that is, a questionnaire developed in a cultural context different from where it is used) to an "emic" set of items specific to each sample, which presumably should correspond more closely to local understandings of children's temperament (Berry, 1989).

Using this procedure, the US sample, not surprisingly, seems to have the best fit with the original BSQ (92 items retained); however, the Polish and Spanish samples are close behind (87 and 86 items retained), and the Dutch and Swedish samples follow them with 81 and 80 items retained. The Italian set of items is much lower (59), but 27 of the 41 omitted items were previously removed from this version of the BSQ, following a standardization of the instrument that has been used in other research (Attili, 1989). The Australian sample, with only 73 items of the original 100 items retained, seems to contain greater individual variability.

The patterning of dropped and retained items across the seven samples indicates both sample-specific and general patterns. Some items were dropped from only one sample (not counting items previously omitted in the Italian version), suggesting that they may not be ecologically valid in particular cultural groups. More typically, however, the same items from the BSQ were dropped from more than one sample, suggesting that they may not work so well in general even though very few of these were dropped from the US sample. For all but one of the dimensions (Threshold), there is moderate agreement about which items were perceived as belonging together within each specified dimension. Overall, 64 (of 100) items were retained in at least six of the seven sample-specific, emic versions. This subset of 64 items constitutes a "derived etic" set of scales that can be considered culturally valid for our samples (Berry, 1989). Their median Cronbach Alphas², across all seven sites, are: for Activity, .74; Regularity, .57; Adaptability, .67; Approach, .77; Intensity, .63; Mood, .63; Persistence, .66; and Distractability, .77. As the final Threshold scale was left with only one item, no reliability can be calculated. These measures of internal coherence are in most cases very similar to Alphas for the locally optimized (emic) versions; the median difference between the Cronbach Alpha coefficients for the two versions is .03. In summary, then, the behavioral dimensions summarized by the retained items seem to be generally recognizable by parents of young children across the seven cultural samples.

Variation in Mean Scores on the Temperament Dimensions

In order to avoid the ethnocentricity of basing comparisons on US norms, we standardized scores on the full, seven-site corpus, with appropriate weights to represent equally each age group in each sample. The resulting z-scores, representing deviations within a hypothetical seven-community population, are used for all subsequent analyses. Table 1 presents the mean rating for each dimension, in each of the samples, ex-

pressed in standardized scores (pooled mean = 0, s.d. = 1, weighted for equal *n* with age and sample), as well as the raw (non-standardized) means and standard deviations. Also shown are statistics for the Sample effect in a three-way analysis of variance (other main effects—Child Sex and Child Birth Order—and all interactions failed to reach the .05 level of significance), and the related significant pair-wise comparisons.

As is evident in Table 1, parents in all cultural samples rated their children as being rather similar in terms of Activity, Adaptability, Intensity, Mood, Persistence, and possibly Approach, whereas there are cultural differences on the dimensions of Regularity, Distractibility, and Threshold, as well as the global Difficulty rating. In terms of absolute ratings (see bottom of Table 1), parents generally perceived their children as being active, regular, fairly adaptable, somewhat approaching, quite intense, mostly pleasant in mood, somewhat distractible, and somewhat persistent. Tukey-Kramer (Kramer, 1956) post-hoc comparisons between all pairs indicate which samples are responsible for the significant effects of culture (see the lower part of Table 1). One of the strongest differences is found in the Regularity dimension, where the Dutch mothers rated their children as more regular in their daily rhythms of sleep and activity than did mothers in all other samples except Australia. Two other significant findings highlight differences between Sweden and the other samples, regarding Threshold and global Difficulty. Swedish children are evidently perceived by their mothers as being less sensitive and also less difficult than children in four of the six other samples. The Italian rating of their children as low in Intensity is intriguing, but seems to be due to the more limited set of items used in the Italian questionnaire, as the cultural difference disappears when other samples are compared using the same reduced set. The last difference shows cross-cultural variability in Distractibility, with the Australian children rated as most and the Swedish children as least distractible.

To examine the degree to which family-level demographic variables might relate to temperament ratings both within and among the cultural groups, we correlated, separately for each sample, the temperament scores (excluding Threshold) and difficulty rating with seven background variables: maternal employment status, father's hours at work per week, mother's and father's educational level, mother's and father's age, and the number of children in the family. Of the resulting 560 correlation coefficients, only 38, or 6.7 percent, reach the .05 level of significance (that is, not many more than would be expected by chance). The correlations are small, averaging .16 in absolute value. There is little patterning to the statistically significant ones, and little generality across samples.

Most importantly for the present analysis, inclusion of the seven background variables as covariates in the central Analysis of Variance does not functionally alter the between-group conclusions. Maternal age does contribute significantly as a negative covariate to ratings on Activity and Threshold; maternal education contributes negatively to Distractibility; and number of children contributes negatively to Distractibility. However, although these covariates slightly alter the *F* values for the Sample effects, the pattern of statistical significance among groups and the relative distribution of group means are not altered from those reported in Table 1. In short, the group differences cannot be accounted for by these family-level variables.

Table 1. Mean (and Standard Deviation) of Global Difficulty Rating and Standardized Temperament Scores, with ANOVA Results and Post-hoc Comparisons.

site	n	Difficulty	Activity (high)	Regularity (low)	Approach (neg)	Adaptability (low)	Intensity (high)	Mood (neg)	Persistence (low)	Distraction (high)	Threshold (low)
Australia	22	2.6 (1.3)	-.05 (.32)	-.06 (.31)	.09 (.34)	-.04 (.40)	.11 (.38)	.09 (.38)	.04 (.39)	.19 (.33)	.10 (.35)
Italy	24	2.8 (0.8)	.09 (.48)	.11 (.36)	.09 (.33)	.16 (.38)	-.24 (.41)	.07 (.34)	.09 (.43)	-.02 (.36)	---
Netherlands	64	2.7 (0.9)	-.07 (.33)	-.21 (.33)	-.05 (.33)	.04 (.42)	.04 (.34)	.01 (.38)	.04 (.35)	.04 (.34)	.09 (.31)
Poland	45	3.0 (1.0)	.04 (.37)	.09 (.35)	.11 (.41)	-.04 (.40)	.13 (.37)	.01 (.42)	.10 (.40)	.11 (.39)	.09 (.42)
Spain	41	2.4 (1.2)	.09 (.43)	-.09 (.48)	-.05 (.47)	.06 (.42)	.13 (.48)	.03 (.46)	-.07 (.56)	.10 (.48)	.18 (.38)
Sweden	35	1.9 (0.9)	.07 (.41)	.03 (.43)	-.17 (.40)	-.10 (.35)	-.04 (.25)	-.13 (.36)	-.05 (.32)	-.14 (.38)	-.22 (.45)
USA	68	2.8 (1.3)	.02 (.37)	.02 (.37)	.01 (.38)	.01 (.36)	-.02 (.37)	.06 (.40)	.09 (.33)	-.05 (.42)	-.08 (.35)
Overall Mean (not standardized)		2.62 (2.85)	3.95 (2.06)	2.67 (2.43)	3.02 (2.09)	3.02 (1.90)	4.26 (1.73)	3.55 (1.74)	2.83 (2.13)	4.03 (2.32)	4.49 (3.72)
Sample effect											
<i>F</i> (6,220)		3.88	1.10	4.32	1.98	1.67	4.11	1.35	0.96	2.59	5.56
<i>p</i>		.001	ns	.0004	.07	ns	.0006	ns	ns	.02	.004
significant pair-wise comparisons (Tukey-Kramer)		SW < IT; NL, PO, US	.	NL < IT; PO, SW, US	.	.	IT < AU, NL, PO, SP	.	.	AU > SW	SW < AU, NL, PO, SP and US < SP

Note: Original scoring used, thus high scores imply "difficult" behavior.

Relationships Between Temperament Dimensions and Mothers' Perceptions of "Difficulty"

In order to explore both general and culture-specific patterns of "goodness of fit," we first correlated children's scores on the eight derived-etic dimensions with mothers' global ratings of how difficult their child was. The results, shown in Table 2, indicate some general trends. First, Adaptability and Mood are significantly related to Difficulty ratings in all samples but one. The Swedish correlation between Adaptability and Difficulty is almost the same as in the Dutch sample, but due to smaller sample size does not reach significance; both the Swedish and Dutch samples are characterized by a lower correlation between Adaptability and Difficulty than found in the other samples. The Italian sample is strikingly different from all other samples in that the correlation between Mood and Difficulty is virtually nil. Other temperament dimensions are related to Difficulty in some but not all samples: Activity in four, and Persistence in three. The remaining dimensions are related to Difficulty in only one or two samples: Intensity in The Netherlands and US samples, Approach only in the Italian sample, and Distractibility only in the Dutch sample (In all these contrasts, the identified differences between the correlations are at least marginally significant, $p < .10$).

Table 2. Correlations of Temperament Dimensions with "Difficulty" Rating.

Site	N	Acti- vity (high)	Regu- larity (low)	Ap- proach (neg)	Adap- tability (low)	Inten- sity (high)	Mood (neg)	Persis- tence (low)	Dis- tracti- bility (high)	Thres- hold (low)
Australia	22	.44*	.18	.24	.52*	.06	.63***	.47*	.06	-.02
Italy	24	.09	.15	.42*	.43*	.19	-.09	.12	-.05	-
Netherlands	63	.38**	.04	.06	.29*	.33**	.30**	.49***	.27*	.10
Poland	44	.32*	.17	.14	.42**	.14	.49***	.24	.13	.13
Spain	41	.28	-.10	.11	.50***	.12	.32*	.27	-.05	.26
Sweden	33	.06	.02	.14	.27	-.02	.59***	.44*	.15	.30
USA	66	.35**	.19	.01	.51***	.38**	.60***	.18	.06	.00

Notes: Original scoring used, thus high scores imply "difficult" behavior.

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

Examination of the inter-correlations among the eight temperament dimensions plus Difficulty rating within each site (not shown) suggests that not only does the temperament-difficulty relationship vary from group to group, but also that the interrelations among the eight temperament dimensions differ. In order to represent these sets of relationships, we applied multidimensional scaling to the 9 x 9 correlation matrix (eight temperament dimensions—omitting the single-item Threshold—plus Difficulty) for each culture. Figures 1 and 2 illustrate the results for two contrasting patterns, in Italy and the Netherlands. Temperament dimensions that

correlate .30 or greater with Difficulty for each site are enclosed in a circle for greater ease of interpretation. (The statistical power in each sample varies with the number of subjects, but this criterion of $r \geq .30$ excludes all non-significant correlations and includes all significant ones, with single exception of Distractability in the Dutch sample, which is significant but not included in the circle.)

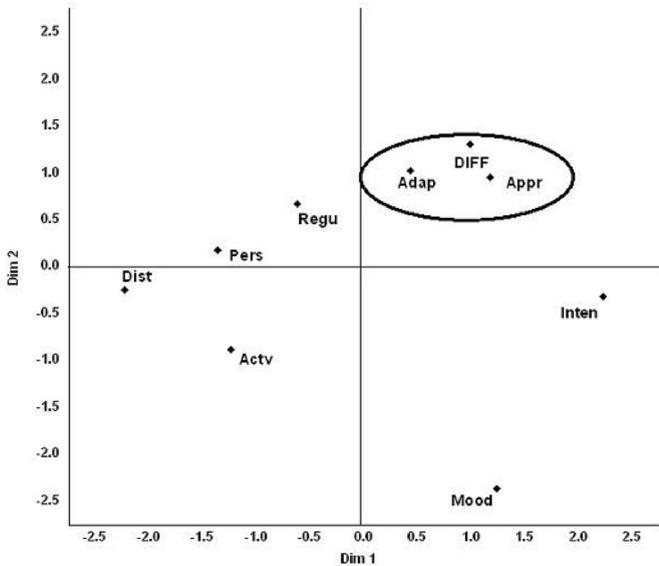


Figure 1. Multidimensional scaling of correlations among eight temperament dimensions and difficulty rating: Italy.

The Italian scaling (Figure 1) is distinctive, reflecting the fact that only Adaptability and Approach are significantly related to Difficulty. Most striking in this scaling is the absence of Mood in the Difficult grouping, as well as the distance between Mood and all the other dimensions except for Intensity (with which it is correlated at .35). The Spanish scaling³ (not shown), like the Italian one, includes only two temperament dimensions related to Difficulty; but Mood appears here instead of Approach. The Swedish scaling (not shown) is also distinctive in that Persistence and Mood are the only dimensions in the Difficulty group (as noted above, Adaptability is moderately but not significantly related to Difficulty). Four samples (Australia, The Netherlands, Poland, and the US) are all characterized by a trio of the dimensions Mood, Activity, and Adaptability in the Difficult grouping (see Dutch example in Figure 2). There are also some important differences within this group, however, especially with regard to the inclusion of Intensity, which is significantly correlated with Difficulty in the US and The Netherlands, but not

in Poland or Australia. Finally, Persistence is found in the Difficult grouping in three samples (Sweden, The Netherlands, and Australia) but not in the others.

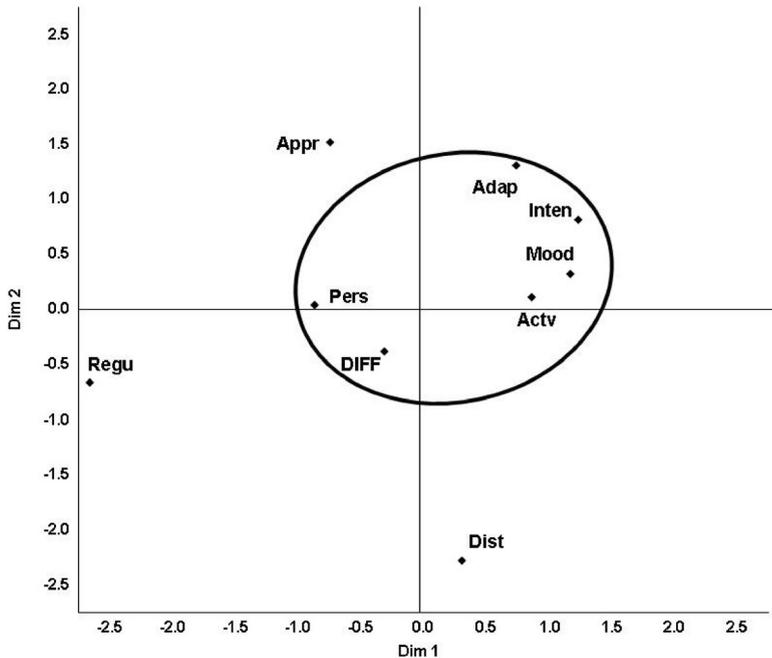


Figure 2. Multidimensional scaling of correlations among eight temperament dimensions and difficulty rating: Netherlands.

Looking at the larger distribution of temperament terms in the seven scaling solutions, it is apparent that there is one dimensionality that carries across all samples (except for Australia), with Approach, Regularity, Adaptability, and Persistence tending to appear on one side, contrasted with Distractibility, Intensity, Mood, and Activity on the other. This contrast appears on Dimension 1 (horizontal) in the US, Polish, and Swedish scalings, and on Dimension 2 (vertical) in the Italian, Spanish and Dutch scalings. In either case, this dimension seems to indicate that parents differentiate between active, disorganized, negative energy on the one hand and friendly, adaptable stability on the other.

The Representation of Cultural Models in Parental Discourse

The culturally distinctive patterns of relationships among temperament dimensions and global Difficulty ratings, as indicated in their intercorrelations and portrayed through multidimensional scaling, suggest that mothers in these communities have

both shared and culturally distinctive conceptualizations of children’s temperament, and in particular of what makes a child “difficult.” In order to explore this interpretation further, we examined portions of interviews in which parents were asked to describe their child “to someone like me who doesn’t know your child very well,” focusing in particular on interviews of parents who had rated their child at the top of the “difficulty” distribution. (Transcriptions were not available for the Australian sample.) As the following examples illustrate, the dimensions of temperament identified through analysis of the questionnaire data also find expression in the ways that both mothers and fathers actually talked about their “difficult” children.

US: Jason, an Intense and Emotional Five-year-old Boy

In the US sample, global Difficulty ratings are correlated significantly with Activity, Adaptability, Intensity, and Mood. These qualities—especially intensity and mood — are highlighted by Jason’s parents (all names are fictitious), who rated him as quite difficult. When asked to describe their son, his parents’ first response is to note his intensity of response and emotionality: As his mother says, “He’s probably hypersensitive.” His father adds, “He’s sensitive, thoughtful, impulsive and compassionate on either side. Any spectrum which you could imagine, he’s on the edges, he’s very rarely in the middle.” Jason’s parents go on to relate his intensity to being “hard on himself”—a perfectionist who won’t demonstrate new knowledge to his parents until he has mastered it completely on his own. This is related in turn to his frequently negative mood: As his mother says, “Things weigh heavy on him.” They also describe him as being unable at times to modulate his behavior to rules at home - an aspect of low adaptability.

Italy: Francesca, a “Selectively” Shy Three-year-old Girl

Based on the Italian questionnaire results, we would expect parents in this sample to be especially preoccupied with children’s ability to approach new social situations with ease. Francesca’s parents, who have rated her as somewhat difficult, open their description with a comment on this very topic, with her father characterizing her as “positive, extroverted...she is responsive to the needs of those around her. I’d say she’s not afraid of...she’s stimulated by novelty, let’s say she’s a bit stubborn...” Francesca’s mother agrees with her husband’s description, but then introduces an aspect of Francesca’s behavior that worries her:

“I almost agree, especially about her sociability, her independence and her stubbornness. I’m in agreement with him about her sociability, but there’s one thing we don’t agree on - he says she’s “selective” and that’s what he finds difficult. I still don’t really understand this, but it seems that in reality even if she wants to be sociable and although she feels secure, she knows that mainly with adults, with all adults, even the ones she doesn’t know. But in relationships with other children she’s not so easy... For example, two weeks ago we went to a birthday party for some twins that she has known

for a while, and there were so many other children she didn't know, and she stayed beside me for half an hour, almost without even moving. So, this selectivity, she has trouble with entering, but if she gets to know them bit by bit...it's the (initial) impact."

Francesca's parents both then describe her as "difficult," even while noting that "Everyone says, "She's so sociable, you won't have any difficulties."

The Netherlands: Marinka, an Active and Intense Five-year-old Girl

Analysis of the Dutch questionnaire data produced significant correlations between Difficulty and six of the temperament dimensions: Activity, Adaptability, Intensity, Mood, Persistence, and Distractibility. These last two are culturally distinctive, and seem consistent with findings from our research on parents' ideas on children's independent and dependent behavior (Harkness, Super, & Pai, 2000). For Marinka's parents, however, the first four dimensions are the focus of attention. According to her parents, Marinka lives from one exciting event to the next, whether it's decorating their house with orange balloons for the Queen's Birthday, or going back to school after a short break. As her mother says, "Ya, her day is really busy. She's very happy, enthusiastic, hopping and running. Constantly on the go." Her father agrees, "Ya, she's always been like that." Going to school, though, presents new problems related to adaptability, as it seems to for many Dutch children. Her father elaborates: "Ya, it's really since she turned four, and the reason is really that now she goes to school. And they naturally get more self-willed. The teacher teaches her things, and then she comes home and she has to cope with other rules. So, she is a lot more difficult." Marinka's mother concurs: "Ya, I find her very difficult...But it also depends. If it's nice weather and she can play outside, then you don't have any trouble with her. Then she listens nicely, and she's sweet. But if she has to sit inside for two days because it's raining and there's no little friend to play with, then she gets fussy, and becomes a pest."

Poland: Theodor, a Moody, Sensitive, and Unadaptive 4-year-old Boy

From the questionnaire results, we would expect parents in the Polish sample to focus on Activity level, Adaptability, and Mood in talking about "difficult" children. These last two qualities are highlighted by Theodor's parents. As his mother describes him, "First and foremost, Theodor is very sensitive. For example, his sister Magda never cried when listening to fairy-tales. Theodor reacts in an unusual way, he gets so stirred-up. So, first of all he is emotional, sensitive, and you have to talk to him in a quiet way, you have to explain everything. You really need a lot of patience for him. Even though you keep repeating things to him, he just won't comply." Theodor's older sister, who has been listening to the interview, chimes in: "On Sundays, instead of getting up and leaving [with us], he stays in bed, he doesn't want to have breakfast or even get dressed. We are always late because of him." His mother agrees, adding, "Theodor has a problem with getting up and eating. And when we argue, he reacts very intensely. I do believe we need a good psychologist because there are so many things I am not able to cope with!"

Spain: Carmen, a "Difficult to Raise" Four-year-old Girl

Adaptability and Mood are the only two dimensions of temperament that are significantly related to Difficulty in the Spanish sample. Four-year-old Carmen, according to her parents, illustrates how both qualities can present challenges to parents. After describing her as "normal" because "she behaves herself well," her mother goes on to characterize her as "a difficult little girl to bring up." Both parents describe Carmen as "nervous" and "capricious," a child who demands a special diet and who is able to get her own way with her parents: "She does whatever she wants with us." The mother qualifies this description, however, by stating that Carmen is "only difficult for me." When asked why by the interviewer, the mother explains: "Because she is very shy, so with me she expresses herself as she really is. Of course, with me she expresses whatever she wants." In contrast, the father asserts, "When she's with me, she does whatever little thing I ask. For example, if I ask her to sing me a Sevillana [song], because she's beginning to sing and dance... until she can't do it and she says, 'Come, papa, I'm going to sing for you,' then she tries again." Her mother adds, "With me, she also sings and dances," and the father returns to the earlier theme by adding, "She's quite timid." When the interviewer seeks clarification by asking, "In front of other people, right?" both parents reply together, "Right, right."

Sweden: Kerstin, a Happy but "Crabby" 3-year-old

In the Swedish questionnaire results, the only two dimensions of temperament significantly related to Difficulty were Mood and Persistence; Adaptability is also modestly correlated with Difficulty (at .27) but as noted above, the relationship does not reach statistical significance in this sample. Very few of the Swedish children were given high ratings on the Difficulty scale, as illustrated by the following excerpt from an interview with parents of three-year-old Kerstin, who sum up their comments by calling her "quite easy," corresponding well to her rating of four on a scale from one to seven - one of only two children who were rated that high. Nevertheless, Kerstin's parents do find her behavior difficult in some circumstances, especially with regard to maintaining a positive mood through the daily routines that may require adaptations to changes in social setting or schedule. Kerstin's persistence in demands for her parents' attention is also annoying to her parents. As her mother recounts, "If at daycare you want to chat with another parent, you shouldn't try to move away too quickly." Her father adds, "She will stand there and pull on your pant leg and so forth, but as soon as our attention is turned back to her, then..." The mother finishes the sentence, "Then she is happy. But she has a short fuse and is a little crabby, quite crabby and [insistent]...She is, not sad, but a bit whiney, she'll make a fuss and whine. But otherwise, she is quite happy, she sings often and is cheerful, can zip around a bit."

As these examples of parental discourse suggest, there seems to be substantial convergence between the temperament questionnaire results and how parents actually talked about their children. All the children described here were rated at or near the top of the “difficulty” range for parents in each cultural sample, and all the parents did express feelings about the child as difficult for them in some ways. The particular behaviors that caused difficulty for parents in these examples showed some of the same trends evident in the questionnaire results: negative mood and low adaptability show up repeatedly in parents’ talk about their children. “Sensitivity” as described by parents—a component of Intensity in the questionnaires—also seems to be an important part of parents’ perceptions of their child’s difficultness in the US and Polish examples, as exemplified by boys whose intense negative moods posed challenges for other members of their families. Intensity of emotional expression can also be positive, however, as in the Dutch example—and in this case parents can find this quality difficult simply because it demands their attention.

Beyond these general trends, it also seems clear that the particular aspects of temperament that parents found “difficult” in each sample reflect cultural models of the child. In this regard, the uniqueness of the Italian sample in the association between Approach and Difficulty (see Table 2) is illustrated by the ways that 3-year-old Francesca’s parents describe their problems with her in social situations. Interestingly, Francesca is seen as “difficult” even though she is not generally shy, but rather “selectively shy” when confronted with many other children at once. The relationship between temperament and a cultural model of the child that places a strong emphasis on ease in establishing social relationships gains further support from the parents’ comment from a friend that “She’s so sociable, you won’t have any difficulties,” as if the only source of difficulty in a child would be found in this dimension of temperament. The implicit model used by the Spanish parents is similar to that of the Italian parents in its focus on shyness in social situations, but Carmen’s parents seem to be more concerned about the performance aspects of social behavior—being well-behaved in public, and showing future promise in singing and dancing for an audience. In this context, Carmen’s parents emphasize the difference between the “social” Carmen who is quite shy, and the “real” Carmen who can express herself openly and assert her own desires, even if this makes her somewhat difficult at home. Implicitly, Carmen’s parents seem to assume that she will naturally outgrow her shyness and do well in social occasions. We suggest that this may relate, in turn, to a Spanish cultural model of the child that emphasizes qualities of good citizenship and social decorum (Harkness & Super, 2005). In the Swedish example, we find yet another parental perspective on the quality of shyness, one that seems to emphasize parents’ preference for a child who is relatively undemanding of parental attention in transitional situations such as being dropped off at daycare in the morning.

Discussion and Conclusions

The patterns of cultural variability in this study are generally consistent with findings from other research with these populations. Parents' ratings of their children's temperaments, as noted above, were generally quite similar, on average, with the primary exception of Regularity, such that the Dutch parents rating their children as significantly more regular than parents in any other sample. The importance of establishing and maintaining a regular and calm schedule for young children has emerged from other research with this population (Harkness, Moscardino, et al., 2007; Super, Harkness, van Tijen, van der Vlugt, Dykstra, & Fintelman, 1996), and is an example of how cultural practices may actually lead to differences in children's temperaments. The uniquely Italian pattern of association between the temperament dimensions of Approach and Adaptability with global Difficulty is consistent with a cultural model of parenting that includes introducing the child to a variety of social situations and encouraging the development of emotionally close relationships starting in early infancy (Harkness, Blom, Oliva, Moscardino, Zylicz, Rios Bermudez, Feng, Axia, & Super, 2007). In this context, a child who is temperamentally shy or withdrawn in social encounters may truly represent a difficult challenge to its parents (Axia, 1999). The unique lack of association between Mood and Difficulty in the Italian sample, likewise, may reflect a cultural perspective in which a child who shows more frequent negative moods may be perceived simply as one whose emotional expression is more accessible and thus more available for forming close bonds (Axia & Weisner, 2002).

Another distinctive pattern emerges with regard to Persistence, which is associated with Difficulty only in the Australian, Dutch, and Swedish samples. This may be related to an emphasis shared among these three groups on the early development of independence in the sense of being able to entertain oneself for periods of time. The unique association between Distractibility and Difficulty in the Dutch sample is also suggestive in this regard and is consistent with other research on Dutch parents' perceptions of their children (Harkness, Blom, et al., 2007; Harkness, Super, & Pai, 2000).

Along with cultural variability in some of these results, there is also evidence for universality across the samples. First, it is notable that the "derived etic" version of the Behavioral Style Questionnaire (based on a reduced set of items with improved internal reliability compared to the full set) seems to capture groupings of specific behaviors that parents in all sites perceive as related to each other, and which are well summarized by the Thomas and Chess dimensions. Second, in addition to their agreement on the content of the Thomas and Chess temperament dimensions, the parents in our samples also seem to perceive their own children's temperaments as remarkably similar on average; that is, the individual variability in children's temperaments according to their parents generally converged around similar means (with the particular exception of Regularity as mentioned above). Thirdly, the association of low Adaptability with Difficulty among parents in all our samples suggests that, at least for parents in communities like the ones we studied, young children who have a hard time adjusting to changes in environment or routines, or who tend to battle with rules imposed by adults, are likely to be perceived

as more difficult than their “laid-back” peers. The general association of Adaptability and Mood (except for the Italian sample) with Difficulty lends partial support to the temperamental profile of the “difficult child” originally proposed by Thomas and Chess. Notably, however, the other three dimensions in the “difficult” group - Regularity, Approach, and Intensity - do not seem to relate to maternal perceptions of Difficulty in most samples; and another dimension that is not part of the “difficult” construct, Activity, is significantly or almost significantly related to Difficulty in five of the samples.

The cross-cultural patterning of temperamental correlates of Difficulty, as seen in the present study, supports a key tenet of modern temperament theory regarding the importance of “fit” with the environment of daily life, rather than an absolutist position of “difficulty” as a trait. As further illustrated by the interview excerpts presented here, cultural variability in what parents find difficult has important implications for the shaping of developmental outcomes. By extension, there are also implications for parent education and clinical practice. What is appropriate or healthy in one cultural context may not be in another, due to differences in the meaning and functionality that are constructed around specific behaviors.

Limitations

The present study has several limitations that should be kept in mind when interpreting the results. First, we chose to work with Thomas and Chess’ nine-dimension conceptualization of temperament rather than any of several other excellent frameworks. Using a different questionnaire might have produced a different pattern of cultural similarities and differences. Second, this study is limited to mothers, and we know that fathers’ perceptions of their children’s temperaments are rarely identical to those of their partners. Third, although our interpretations of cultural models in parental discourse reported here are consistent with other related research, they remain to be examined systematically in relation to parents’ perceptions of temperament both among these parents and in other samples. Finally, our samples are quite small (especially in Australia, Italy, and Sweden) and in addition they are limited to broadly middle-class families. Our results, therefore, should be taken as more suggestive than conclusive, indicating possible future directions for research on children’s temperament in cultural context.

Notes

¹ Details of demographic variables are available from the first author.

² The following BSQ items were retained in the final, derived-etic version (-R indicates a reversed item). For Activity: 4, 9, 14, 24, 32, 58, 70, 87, 26-R, 44-R; Regularity: 23-R, 36-R, 47-R, 84-R; Approach: 21, 50, 68, 86, 98, 12-R, 25-R, 31-R, 43-R, 54-R, 67-R; Adaptability: 8, 10, 55, 65, 56-R, 61-R, 63-R; Intensity: 7, 20, 41, 42, 45, 46, 53, 82, 92, 99; Mood: 1, 34, 38, 64, 69, 72, 78, 96, 100; Persistence: 39, 71, 27-R, 35-R, 40-R, 93-R; Distractibility: 51, 66, 77, 81, 85, 89; Threshold: 57. Psychometric details are available from the first author.

³ The Australian, Polish, Swedish, Spanish, and US scalings are available from the first author.

References

- Attili, G. (1989). The psychology of character and temperament in Italy: A historical review and recent trends. In G. A. Kohnstamm & J. E. Bates (Eds.), *Temperament in childhood* (pp. 581-596). Oxford, England: John Wiley & Sons.
- Axia, G. (1999). *La timidezza [Shyness]*. Bologna: Il Mulino.
- Axia, G., Prior, M., & Carelli, M. G. (1992). Cultural influences on temperament: A comparison of Italian, Italo-Australian, and Anglo-Australian toddlers. *Australian Psychologist*, 27(1), 52-56.
- Axia, G., & Weisner, T. S. (2002). Infant stress reactivity and home cultural ecology of Italian infants and families. *Infant Behavior and Development*, 25(3), 255-268.
- Bates, J. (1998). Temperament. In N. Eisenberg (Ed.), *Handbook of Child Psychology*, Vol. 3: Social, Emotional, and Personality Development, 5th Ed. (pp. 105-176). New York: Wiley.
- Berry, J. W. (1989). Imposed etics - emics - derived etics: The operationalization of a compelling idea. *International Journal of Psychology*, 24, 721-735.
- Carey, W. B., & McDevitt, S. C. (1989). *Clinical and educational applications of temperament research*. Bristol, PA: Swets and Zeitlinger.
- Chen, X., Hastings, P. D., Rubin, K. H., Chen, H., Cen, G., & Stewart, S. L. (1998). Child rearing attitudes and behavioral inhibition in Chinese and Canadian toddlers: A cross-cultural study. *Developmental Psychology*, 34(4), 677-686.
- Chen, X., Rubin, K. H., & Li, Z.-Y. (1995). Social functioning and adjustment in Chinese children: A longitudinal study. *Developmental Psychology*, 31(4), 531-539.
- Chen, X., Rubin, K. H., & Sun, Y. (1992). Social reputation and peer relationships in Chinese and Canadian children: A cross-cultural study. *Child Development*, 63(6), 1336-1343.
- Chess, S., & Thomas, A. (1984). *Origins and evolution of behavior disorders*. New York: Brunner/Mazel.
- D'Andrade, R., & Strauss, C. (1992). *Human motives and cultural models*. Cambridge: Cambridge University Press.
- Harkness, S., Blom, M. J. M., Oliva, A., Moscardino, U., Zylicz, P. O., Rios Bermudez, M., Feng, X., Axia, G., & Super, C. M. (2007). Teachers' ethnotheories of the "ideal student" in five western cultures. In J. Elliott & E. Grigorenko (Eds.), *Western psychological and educational theory in diverse contexts*, Special Issue of *Comparative Education*, 43(1), 113-135.
- Harkness, S., Moscardino, U., Ríos Bermúdez, M., Zylicz, P. O., Welles-Nyström, B., Blom, M., Parmar, P. Axia, G., & Super, C. M. (2006). Mixed methods in international collaborative research: The experiences of the International Study of Parents, Children, and Schools. *Cross-cultural Research*, 40(1), 65-82.
- Harkness, S., & Super, C. M. (1996). Introduction. In S. Harkness & C. M. Super (Eds.), *Parents' cultural belief systems: Their origins, expressions, and consequences* (pp. 1-24). New York: Guilford Press.
- Harkness, S., & Super, C. M. (2005). Themes and variations: Parental ethnotheories in western cultures. In Rubin, K. (Ed.), *Parental beliefs, parenting, and child development in cross-cultural perspective* (pp. 61-79). New York: Psychology Press.
- Harkness, S., Super, C. M., Moscardino, U., Rha, J.-H., Blom, M. J. M., Huitrón, B., Johnston, C., Sutherland, M., Hyun, O.-K., Axia, G., & Palacios, J. (2007). Cultural models and developmental agendas: Implications for arousal and self-regulation in early infancy. *Journal of Developmental Processes*, 1(2), 5-39.

- Harkness, S., Super, C. M., & van Tijen, N. (2000). Individualism and the "Western mind" reconsidered: American and Dutch parents' ethnotheories of children and family. In S. Harkness, C. Raef, & C. M. Super (Eds.), *The social construction of the child: Understanding variability within and across contexts* (Vol. 87, pp. 23-39). *New Directions in Child Development*. San Francisco: Jossey-Bass.
- Hsu, C.-C. (1981). The temperamental characteristics of Chinese babies. *Child Development*, 52(4), 1337-1340.
- Kagan, J., Snidman, N., Arcus, D., & Reznick, J. S. (1994). *Galen's prophecy: Temperament in human nature*. Cambridge, MA: Harvard University Press.
- Korn, S. J., & Gannon, S. (1983). Temperament, cultural variation and behavior disorder in preschool children. *Child Psychiatry and Human Development*, 13(4), 203-212.
- Kramer, C. Y. (1956). Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics*, 12, 307-310.
- Kyrios, M., Prior, M., Oberklaid, F., & Demetriou, A. (1989). Cross-cultural studies of temperament in Greek infants. *International Journal of Psychology*, 24(5), 585-603.
- Martin, R. P., Wisenbaker, J., & Huttunen, M. (1994). Review of factor analytic studies of temperament measures based on the Thomas-Chess structural model: Implications for the Big Five. In C. F. J. Halverson, G. A. Kohnstamm & Martin, R. P. (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 157-172). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Maziade, M., Caron, C., Côté, R., Mérette, C., Bernier, H., Laplante, B., Boutin, P., & Thivierge J. (1990). Psychiatric status of adolescents who had extreme temperaments at age 7. *American Journal of Psychiatry*, 147(11), 1531-1536.
- McDevitt, S. D., & Carey, W. B. (1978). The measurement of temperament in 3 to 7 year old children. *Journal of Child Psychology and Psychiatry*, 19(3), 245-253.
- Nakagawa, A., & Sukigara, M. (2005). How are cultural differences in the interpretation of infant behavior reflected in the Japanese Revised Infant Behavior Questionnaire? *Japanese Journal of Educational Psychology*, 53(4), 491-503.
- Prior, M. R., Kyrios, M., & Oberklaid, F. (1986). Temperament in Australian, American, Chinese, and Greek infants: Some issues and directions for future research. *Journal of Cross Cultural Psychology*, 17(4), 455-474.
- Prior, M., Garino, E., Sanson, A., & Oberklaid, F. (1987). Ethnic influences on "difficult" temperament and behavioural problems in infants. *Australian Journal of Psychology*, 39(2), 163-171.
- Rogoff, B., & Angelillo, C. (2002). Investigating the coordinated functioning of multifaceted cultural practices in human development. *Human Development*, 45(4), 211-225.
- Rothbart, M. K., Ellis, L. K., & Posner, M. I. (2004). Temperament and self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 357-370). New York: Guilford Press.
- Rothbart, M. K., Chew, K. H., & Gartstein, M. A. (2001). Assessment of temperament in early development. In L. T. Singer & P. S. Zeskind (Eds.), *Biobehavioral assessment of the infant* (pp. 190-208). New York: Guilford Press.
- Rutter, M. (1989). Pathways from childhood to adult life. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 30(1), 23-51.
- Shwalb, D. W., Shwalb, B. J., & Shoji, J. (1996). Japanese mothers' ideas about infants and temperament. In S. Harkness & C. M. Super (Eds.), *Parents' cultural belief systems: Their origins, expressions, and consequences* (pp. 169-191). New York: Guilford Press.

- Super, C. M., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioral Development*, 9, 545-569.
- Super, C. M., & Harkness, S. (1993). Temperament and the developmental niche. In W. B. Carey & S. A. McDevitt (Eds.), *Prevention and early intervention: Individual differences as risk factors for the mental health of children - a Festschrift for Stella Chess and Alexander Thomas* (pp. 15-25). New York: Brunner/Mazel.
- Super, C. M., & Harkness, S. (1994). The cultural regulation of temperament-environment interactions. *Researching Early Childhood*, 2(1), 59-84.
- Super, C. M., Harkness, S., van Tijen, N., van der Vlugt, E., Dykstra, J., & Fintelman, M. (1996). The three R's of Dutch child rearing and the socialization of infant arousal. In S. Harkness & C. M. Super (Eds.), *Parents' cultural belief systems: Their origins, expressions, and consequences*. (pp. 447-466). New York: Guilford Press.
- Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.
- van den Boom, D. C., & Hoeksma, J. B. (1994). The effect of infant irritability on mother-infant interaction: A growth-curve analysis. *Developmental Psychology*, 30, 581-590.

Address for correspondence: Charles M. Super, Center for the Study of Culture, Health, and Human Development, University of Connecticut, 348 Mansfield Road, Storrs, CT 06269-2058 USA. E-mail: charles.super@uconn.edu

Charles M. Super is Professor of Human Development and Pediatrics, and Co-Director of the Center for Study of Culture, Health, and Human Development, at the University of Connecticut. His research interests focus on the cultural regulation of early development.

Giovanna Axia was a Professor in the Department of Developmental and Social Psychology at the University of Padua. Her research focused on the psychological and ecocultural factors involved in child and family health across different contexts, especially temperament, pediatric psycho-oncology, and emergency psychology.

Sara Harkness is Professor of Human Development, Pediatrics, and Anthropology, and Director of the Center for the Study of Culture, Health and Human Development at the University of Connecticut. Her research focuses on the cultural structuring of parenting and children's development, education and health.

Barbara Welles-Nyström is an Assistant Professor in the Department of Women and Child Health, at the Karolinska Institute in Stockholm. She has conducted cross-cultural research on mothers and infants, and families in Sweden, the USA, and Russia.