





Check for updates

Super, C. M., Harkness, S., Bonichini, S., et al. (2020). Developmental continuity and change in the cultural construction of the “difficult child”: A study in six western cultures. In S. Harkness & C. M. Super (Eds.), *Cross-Cultural Research on Parents: Applications to the Care and Education of Children*. *New Directions for Child and Adolescent Development*, 170, 43–68.

# 3

## Developmental Continuity and Change in the Cultural Construction of the “Difficult Child”: A Study in Six Western Cultures

Charles M. Super,  Sara Harkness,  Sabrina Bonichini, Barbara Welles, Piotr Olaf Zylicz, Moisés Rios Bermúdez, Jesús Palacios

### Abstract

*This study explores the cultural construction of “difficult” temperament in the first 2 years of life, as well as the logistical and thematic continuity across infancy and childhood in what mothers perceive as difficult. It extends earlier work regarding older children in six cultural sites: Italy, the Netherlands, Poland, Spain, Sweden, and the United States. In order to compare temperament profiles across sites, a “derived etic” version of standard temperament scales is constructed, and then examined in relation to mothers’ global ratings of how “difficult” the child is to manage. Results are compared to the earlier report. Negative Mood and low Adaptability tend to be problematic in most sites in both age groups. High Activity and Intensity increase in their relevance to difficulty from the first 2 years to early childhood. In some sites, dispositions such as low Approach become less difficult to manage. Of particular note are culturally unique patterns of continuity that appear to be related to larger cultural themes. These results have implications for our theoretical understanding of parenting, as well as for educational and clinical practice. © 2020 Wiley Periodicals, Inc.*

The idea of a “difficult temperament” or “difficult child” remains a salient concept in both the scientific literature and the popular arena (e.g., Acar, Ahmetoğlu, Özer, & Yağlı, 2019; Gouge, Dixon, Driggers-Jones, & Price, 2020; Matthews, 2020; Stephens, 2020; Turecki, 1989), even though the term has come under scientific criticism (e.g., Rothbart, 1982), and the researchers who first introduced the term quickly expressed regret over it—“difficult to manage” would have been better, Stella Chess has remarked (S. McDevitt, personal communication, November 2014). By whatever label, it is certainly the case that particular patterns of child behavior can be problematic for parents.

In the New York Longitudinal Study (NYLS) of middle-class White families, parents expressed particular difficulty with infants who displayed frequent negative moods, high intensity of expression, high levels of activity, irregularity in daily rhythms of wakefulness and eating, withdrawal from new situations, and slowness in adapting to changes in schedule, activities, or settings (Chess, Thomas, & Birch, 1959). However, the same researchers, in a related study with working-class Puerto Rican families, found a different pattern (Korn & Gannon, 1983): low biological regularity did not initially pose a problem, because infants’ and preschoolers’ bedtimes often followed the daily life of the family rather than being keyed to the clock. Highly active toddlers and young children, on the other hand, were more problematic for life in the smaller apartments, which had less access to parks and play spaces. When the age of kindergarten arrived, however, shifting to a regular schedule suddenly became a source of stress and sometimes dysfunction for the Puerto Rican families, while negotiations and socialization in this domain had been largely resolved for the middle-class White families.

The concept of “goodness-of-fit” was developed to explain this pair of results, defined by Carey (1997, p. 196): “A good fit occurs when the values and expectations of parents and other caregivers ... are in accord with the capacities and temperament of the child.” This framework has been applied extensively to understand the broader consequences of temperament, not just parental distress. Keogh (1989), for example, explored the role of temperament in social development as well as academic success in elementary school, and several groups have derived applications for preventive intervention in the contexts of pediatric care (Cameron, Rice, Hanse, & Rosen, 1994) and elementary-school classrooms (McClowry, 2014). Others have looked at goodness of fit in diverse cultures, using a variety of methods and with various outcome measures. Observations in Kenya, Zaïre, the United States, and Samoa, for example, indicate that who has moment-to-moment care of the child (specifically, the mother or other caretaker), and the nature of their interaction (such more or less verbal communication), are associated with the child’s temperament (e.g., negative mood), but that this association varies across cultures (deVries & Sameroff, 1984; Odden, 2009; Super & Harkness, 1994; Winn, Tronick, & Morelli, 1989). Studying

older children in the 1990s, Chen and his colleagues found that behavioral inhibition (shyness) in children was positively related to peer acceptance, maternal approval, and academic adjustment in rural China, whereas the opposite held true in Canada and among modern urban Chinese (Chen et al., 1998; Chen, Rubin, & Sun, 1992; Chen, Wang, & Wang, 2009).

Despite the emergence of other perspectives (e.g., Buss & Plomin, 1975; Kagan, Snidman, Arcus, & Reznick, 1994), the Thomas and Chess framework dominated temperament research for several decades, in part due to its implementation by Carey, McDevitt, and colleagues into questionnaires covering the full lifespan (e.g., Carey & McDevitt, 1978; Fullard, McDevitt, & Carey, 1984). A major project to refine and expand this approach was undertaken by Rothbart and her colleagues, who also developed corresponding questionnaires that cover a full range of ages (Rothbart, 2012). These scales have been particularly useful in exploring the developmental links from early temperament characteristics to later outcomes, including personality (Ahadi, Rothbart, Halverson, Kohnstamm, & Martin, 1994). They have also been used extensively by Gartstein and colleagues to document the structure and outcomes of infant and child temperament in a large number of cultural sites (Gartstein & Putnam, 2018; Gartstein, Slobodskaya, Kirchhoff, & Putnam, 2013).

Cross-cultural comparisons in temperament research engage several difficulties, both statistical and semantic. For the former, confirmatory factor analysis of questionnaire items provides an elegant solution to ensuring structural comparability across groups, but it requires a larger number of respondents than is sometimes practical (Van De Schoot, Schmidt, De Beuckelaer, Lek, & Zondervan-Zwijnenburg, 2015). Even when that technique is feasible, however, questions of meaning remain, as behaviors—even by an infant—are construed by others through cultural systems of thought (D’Andrade, 1990).

In prior research, we developed a method of questionnaire modification that speaks to both sets of issues (Super et al., 2008). Working in six cultural settings, we started with careful translations of the McDevitt and Carey (1978) temperament questionnaire for early childhood. Internal consistency (Cronbach’s alpha) was calculated for each of the nine dimensions, separately for each of the six sites, and any negatively correlated (or very poorly correlated) items were dropped for that site in order to maximize the internal cohesion of each scale in that place. This process converted an “etic” instrument (that is, a questionnaire developed in a specific cultural context: Berry, 1989) to an “emic” set of scales for each sample, which should presumably correspond more closely to local understandings of children’s temperament. We then selected (for each scale) items that remained in at least five of the six samples, resulting in a “derived etic” questionnaire that contained collectively meaningful clusters of behavioral descriptions and was thus able to be appropriately applied to all the groups studied. Using these revised scales, we demonstrated both meaningful cultural

differences in mean ratings on four aspects of temperament (Regularity, Intensity, Distractibility, and Threshold), as well as the overall judgment of Difficulty. Most importantly, we demonstrated cultural differences in how the temperament dimensions related to mothers' subjective ratings of Difficulty.

These results point to a further question for studies of culture, temperament, and parents' experience. Such studies, including ours just cited, tend to draw on a general ethnographic understanding to sketch the settings, customs, and parental ethnotheories pertaining to the children in the studied locale—that is, their developmental niche (Super & Harkness, 1986). Temperament research, however, usually focuses on a particular age group, and thus we have little insight into how parents' perception of difficulty might shift from one stage of development to the next. We know that parents in all cultures have a developmental agenda for their children (Harkness, 2000), including age-related expectations for child behaviors and accomplishments (e.g., Edwards, Gandini, & Giovaninni, 1996; Ninio, 1979;). Yet culture's power to shape development depends on the repetition and elaboration of "messages" across childhood (Mead, 1972; Super & Harkness, 2002), "messages" about cultural values, behavior, and identity. Thus, we arrive at an apparent paradox of long-term cultural continuity and short-term developmental change; where on this continuum do we find "difficult temperament"?

In this light, the present study has four related goals. The first is to replicate and compare the earlier analysis of cultural differences in temperament ratings, now focusing on the first 2 years of life: Do additional mothers from the same set of cultures show the same pattern of temperament ratings as each other, and if not, is the pattern of cultural differences the same as with the older children? The second goal is to investigate cultural variations in the relationship of temperament to difficulty ratings during infancy and toddlerhood: Do mothers agree across these cultural groups on what aspects of temperament are difficult to manage in the opening years? Third, do the intercorrelations of the temperament dimensions themselves appear similar across sites, and in comparison with the earlier results? Finally, the fourth goal is to examine the difficulty results in relation to our earlier findings: To what degree is there cultural continuity and/or developmental change, within cultures, in the relationship of temperament with difficulty?

## Methods

**Samples.** The data here are drawn from the International Study of Parents, Children, and Schools (ISPCS), a collaborative study of parental ethnotheories and practices related to children's development and learning, carried out from 1995 to 1998. The present report focuses six cultural sites: 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 or advised by an established, local psychologist; their names are included in the Acknowledgments.

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. In addition, the samples were balanced for age group (see below), sex, and birth order (first vs. later) of the “focal” children. Other sources of variability such as maternal employment or use of non-parental childcare was allowed to vary freely as these tend to be integral aspects of different cultural places. Our purpose was not to recruit nationally representative samples, but rather to study parents’ cultural models and their role in parent behavior, in a particular time and place. The question of how much our results can be generalized to other populations is beyond the scope of the present research, although some insight can be gained from comparing the results to other studies.

With the exception of the United States, samples in each country were recruited in one geographic area, using a variety of methods including assistance from parent-teacher associations, childcare centers, health clinics, and schools, as well as snowball sampling. 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 childcare 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 metropolitan Boston who were recruited through a health maintenance organization, families living in central Pennsylvania, and families in central and eastern Connecticut. The latter two subsamples were recruited through schools, community centers, and health care practices.

Samples for the ISPCS were built on the assumption that parental ethnotheories are constructed and elaborated in the context of actual practices with specific children, and that the age of the focal 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 focal ages

that do not fall at known periods of particularly 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 focal age groups: 6 months, 18 months, 3 years, 4.5 years, and 7–8 years. At each age, there are at least twelve children, approximately evenly divided by sex and birth order (first-born vs. later-born). The present report uses data from mothers of children in the youngest two age groups (Super et al., 2008, presents results from the school-transition groups). Statistical analysis of demographic measures 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., fathers in our Italian sample average 2 years older than the other groups; US and Polish families are slightly larger; Italian mothers work longer hours outside the house). As demonstrated below, however, these demographic differences have no bearing on our conclusions.

**Instruments.** Mothers completed either the Revised Infant Temperament Questionnaire (RITQ: Carey & McDevitt, 1978) or the Toddler Temperament Scales (TTS: Fullard et al., 1984). These questionnaires are built around the nine temperament dimensions specified by Thomas and Chess (1977), and have several points to recommend them. First, the questionnaire items are behaviorally specific in a way that discourages overall judgments; rather, they focus on particular, observable behaviors. Second, the Carey and McDevitt questionnaires tap the same concepts through age-appropriate behaviors over the wide span of ages needed for the ISPCS research (they were the only such set available at the time). Finally, and still uniquely, the scales (“dimensions”) and their labels are easily understood by parents of all educational levels. This is in part, no doubt, because they were devised from semi-structured, “clinical,” or even ethnographic, interviews with parents. The results presented here, and earlier (Super et al., 2008), ultimately demonstrate the coherence of items belonging to each of the dimensions in the minds of mothers from a broad sample of western cultures.

The Infant Temperament Questionnaire (Carey, 1970) was the first questionnaire implementation of the Thomas and Chess scheme, subsequently improved through psychometric analysis to become the RITQ (Carey & McDevitt, 1978), used here; at the same time, the research team developed the Behavioral Style Questionnaire for children 3 to 7 years (McDevitt & Carey, 1978), used in Super et al. (2008). The TTS, used here, selected items from both the RITQ and the BSQ, implementing “changes that seemed necessary to make the items relevant to the developmental level

of the intended age period” (Fullard et al., 1984, p. 207). Thus many of the items in the two questionnaires used here are similar (e.g., RITQ: “The infant moves much (squirms, bounces, kicks) while lying awake in crib,” and TTS: “The child fidgets during quiet activities (story-telling, looking at pictures”).

Designed to be filled out by parents, the questionnaires contain ninety-five (RITQ) or ninety-seven (TTS) 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). Relevant items are then averaged to construct scales for the nine temperament dimensions. In addition, the last page of each questionnaire asks the respondent for global impressions of the child along each of the dimensions, as well as how “difficult” the child is. The present analysis uses only the questionnaire data provided by mothers.

**The Derived Etic Questionnaire.** In order to replicate the procedure used by Super et al. (2008), internal consistencies (Cronbach’s alpha) for the RITQ and the TTS were calculated for each of the nine dimensions, separately for each of the six sites, and any negatively correlated (or very poorly correlated) items were dropped for that age/site in order to maximize (within a change in alphas of less than .03) the internal cohesion of each scale. Items that remained in at least five of the six samples constituted the “derived etic” questionnaire used for all analyses in the present report.

In this process, the RITQ required least change for the Dutch sample (82 of the original 95 questions were retained), and the most for Italy (43 questions, but 9 of the 52 omitted items for Italy were previously removed following a standardization of the instrument used in other research; Axia, 1993). The TTS required least change for the United States (90 of the original 97 retained), and the most for Spain (65 questions retained). For both questionnaires, a few items were dropped from only one site, suggesting a specific ecological issue, but more frequently, they were dropped from several sites, suggesting a more general issue of applicability. In the end, the derived etic RITQ consisted of 39 items, and the new TTS 72 items. Alphas within samples had a median of .80. The resulting derived etic scales constitute a common core of items with high face validity and, in general, acceptable inter-item correlations, suggesting an operational definition of the nine temperament dimensions generally agreed to by mothers in all the study sites.<sup>1</sup> Illustrative items from each of the nine scales are as follows:

Active infants and toddlers wiggle, squirm, or kick during diapering, dressing, and feeding, and while playing by self. The active toddler will run or jump when exploring a new place, and runs ahead when walking with parents.

Infants and toddlers who score high on Regularity wake up, nap, and become hungry and sleepy at about the same time (within half-an-hour) every day. They are regular in periods of physical activity and fussiness.

Infants and toddlers who are low in Approach are fretful (or very quiet) in a new place or with a stranger. They are likely to object to a new babysitter, a new procedure (such as haircut), or a new food.

Adaptable infants and toddlers adjust to new surroundings within ten minutes, and adjust in two or three days to a change in sleeping arrangements or the introduction of new foods. An adaptable toddler will quickly accept new prohibitions and can be coaxed out of forbidden activities.

Intensity is evident in infants and toddlers when they react strongly to foods, greet a new toy with a loud voice, and laugh or cry at the sight of strangers. Toddlers with intense reactions cry or stamp feet to failure, and show large body movement when upset.

An infant or toddler rated high in negative Mood is fussy when waking up or going to sleep, and does not make happy sounds when diapering, dressing, brushing hair, and other such procedures. Such a toddler will cry after a fall or bump, and has “off” days.

A Persistent infant will play with a toy, or watch other children for more than a minute without looking elsewhere, and will persist for many minutes when working on a new skill such as picking up an object. A Persistent toddler will stay interested in a favorite toy for over an hour, and will go back to an activity after a brief interruption.

The Distractible infant’s attention can be diverted from a soiled diaper for over a minute by picking her up. The Distractible toddler stops eating or playing and looks up when someone walks by, or at a sudden noise such as the telephone or doorbell.

Finally, the infant or toddler with a low sensory Threshold shows discomfort with a soiled diaper or wet clothing, and reacts to changes in temperature or consistency of foods.

To facilitate the core analyses, temperament scores were standardized on the full, six-site corpus (pooled mean = 0, standard deviation = 1, equally weighted for the samples), separately for each age group, and then pooled into a single dataset for all remaining analyses. Combining results from these two questionnaires is justified by their common theoretical underpinnings, overlapping author teams, and similarity in face value of questionnaire items. The resulting combined z-scores, representing deviations within a hypothetical six-community population, are used for the primary analyses, but at key points, we also test for possible differences between the 6- and 18-month groups.

**Procedure.** Parents who expressed interest in participating were telephoned 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 focal child. The nature and purpose of the study were then explained. On a subsequent visit, the team member reviewed the forms left for the parents to complete (including the questionnaires used here), obtained informed consent, and made a date to return for a full interview. The interviews were carried out in the family’s home, often with both parents present, usually



in the evening and lasting about one hour. Questionnaires were computer scored according to the original design, then modified as described above. The interviews were audio-recorded and later transcribed (in the original language), and subsequently examined for further insight into parental thinking about “difficulty.”

**Plan of Analysis.** The analysis of the derived etic scores involves three steps: (1) examination of variation among the samples’ mean scores on the derived dimensions, exploration of the possible influence of background variables, and comparison to the results for the older, school-transition children; (2) correlation of the derived etic temperament dimensions with mothers’ ratings of their child’s Difficulty, within each sample, and relatedly the pattern of intercorrelation among all the temperament dimensions; and (3) comparison of the Difficulty correlations for the infants and toddlers to the previously reported results for older children, with particular attention to patterns of continuity and change.

## Results

**Variation in Mean Temperament Scores.** Table 3.1 presents the mean standardized rating for each dimension in each of the samples (balanced for the distributions of child age and sex), first (per sample) for the present data (6–18 months, marked “Y”) and then for comparison for the 3- to 7-year sample, marked “O” (Super et al., 2008). The raw (not standardized) average rating of Difficulty is also given for each cultural site, and toward the bottom the average unstandardized (raw) temperament scores for the entire, pooled sample are given. The following row shows the probability and effect size (partial omega squared) associated with the main effect of Cultural Site for the present data, derived from a three-way Analysis of Variance (Site  $\times$  Age  $\times$  Sex) of the raw scores. (There is no main effect of Age—that is, between infants and toddlers—in that analysis).

Finally, in order to examine the degree to which group differences in socioeconomic status might bias the cultural comparisons presented here—a question commonly asked of cross-cultural research—we repeated the primary Analysis of Variance reported above after controlling for parental education (average of mother’s and father’s years of schooling). We relied on this education measure, rather than combining it with a rating of occupational prestige (Hollingshead, 1975), as the latter is problematic when comparing across cultures. The results are reported in the last line of Table 3.1. Comparison of the simple and adjusted results suggests that SES plays little role in the comparative results reported in Table 3.1: the differences are trivial and do not weaken the conclusions.

The comparison of means and their statistical analysis in Table 3.1 should be taken only as indicators of potential interest, because (given our sample sizes) we are not able to establish metric invariance through multigroup confirmatory factor analysis. Nevertheless, at least some of the

**Table 3.1. Means/Standard Deviations of Global Difficulty Rating and Standardized Derived Etic Temperament Scores for Younger (6–18 months) and Older (3–7 years) Children**

Site	Age	n	Difficulty (not std)	Activity (high)	Regularity (low)	Approach (neg)	Adaptability (low)	Intensity (high)	Mood (neg)	Persistence (low)	Distractibility (high)	Threshold (low)
Italy	Y	24	2.29/.81	.00/.48	.06/.34	.13/.45	.08/.41	-.07/.48	.16/.48	.07/.31	.28/.45	.14/.25
	O	24	2.80/.82	.09/.48	.11/.36	.09/.33	.16/.38	-.24/.41	.07/.34	.09/.13	-.02/.36	-
The Netherlands	Y	29	2.26/1.06	-.10/.37	-.02/.42	-.17/.31	-.11/.40	.03/.34	-.19/.34	-.11/.32	.00/.41	-.02/.42
	O	64	2.71/0.92	-.07/.33	-.21/.33	-.05/.33	.04/.42	.04/.34	.01/.38	.01/.35	.04/.34	.09/.31
Poland	Y	19	2.28/1.09	.05/.48	-.01/.42	.07/.42	.06/.31	-.10/.34	.11/.30	.12/.43	.07/.35	-.08/.43
	O	45	3.01/1.07	.04/.37	.09/.35	.11/.41	-.04/.40	.13/.37	.01/.42	.10/.40	.11/.39	.09/.42
Spain	Y	28	2.25/1.12	.15/.41	-.20/.44	-.03/.09	-.04/.51	.27/.39	.01/.49	-.14/.54	-.14/.39	.18/.44
	O	41	2.43/1.22	.09/.43	-.09/.48	-.05/.47	.06/.42	.13/.48	.03/.46	-.07/.56	.10/.48	.18/.38
Sweden	Y	24	1.67/0.64	.02/.38	.09/.43	-.08/.35	.01/.39	-.17/.42	-.16/.30	.09/.42	-.12/.32	-.11/.46
	O	35	1.90/0.96	.07/.41	.03/.43	-.17/.40	-.10/.35	-.04/.25	-.13/.36	-.05/.32	-.14/.38	-.22/.45
USA	Y	30	1.89/0.90	-.05/.34	-.01/.37	-.00/.38	-.02/.43	.01/.37	-.04/.37	-.06/.38	-.10/.39	-.06/.39
	O	68	2.84/1.30	.02/.37	.02/.37	.01/.38	.01/.36	-.02/.37	.06/.40	.09/.33	-.05/.42	-.08/.35
Average raw score	Y	154	2.15/0.97	4.29/.95	2.70/.93	2.80/1.22	3.19/1.14	4.11/1.15	2.96/1.06	3.52/1.08	3.60/1.33	3.80/1.19
	O	277	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
Site effect (Y) $p/\omega^2$ simple			.06/.04	.32/.01	.15/.02	.07/.04	.57/.00	<.001/.11	-.007/.07	.24/.01	.001/.10	.11/.03
SES controlled			.06/.04	.36/.00	.21/.02	.05/.05	.23/.01	.03/.06	.003/.09	.22/.02	.001/.11	.09/.03

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

comparisons are of potential interest, particularly the three moderate effects ( $\omega^2 \geq .06$ ) of Cultural site. They suggest that the Spanish infants and toddlers are likely to be rated as more Intense than others; and the Italians as more negative in Mood than others, particularly the Dutch and Swedish; these are largely parallel to the results for school-transition age children (Super et al., 2008), shown in the bottom line of each site's row in Table 3.1. More generally, we computed the rank-order correlation for each dimension (plus Difficulty) between the younger results and the school-transition age results. Despite the very small  $n$  (6 sites), two of the results appear significant, indicating that in cultural sites where mothers rate their infants and toddlers as Approaching or Active relative to the other sites, other mothers in the same sites also rate their older children (these are cross-sectional samples) as relatively Approaching or Active ( $r = .88, p = .02$ , and  $.84, p = .04$ , respectively). Results for two other dimensions point in the same direction: Regularity ( $r = .72, p = .10$ ) and Threshold ( $r = .82, p = .09$ ). In contrast, the correlations for Mood, Persistence, and Intensity are smaller ( $r = .61, p = .20$ ;  $r = .64, p = .17$ ; and  $r = .43, p = .39$ , respectively), and for Adaptability and Distractibility even more so ( $r = -.03, p = .96$ , and  $.20, p = .70$ , respectively).

**Temperament and Difficulty.** The correlation of temperament scores to ratings of Difficulty addresses the second core question of this investigation: To what degree do perceptions of difficulty vary systematically among mothers of different cultural groups? The Pearson correlations by site (top line, marked "Y") are shown in Table 3.2. Using  $r > .30$  (medium effect size) as a standard, there is one near universal: negative Mood is related to Difficulty in all sites except Italy. Less consistently, both negative Approach and low Adaptability are each related to Difficulty in three of the six sites, while high Activity, low Persistence, high Distractibility, and low Threshold are each related to Difficulty in only one. The second line for each site (marked "O") presents for comparison the same correlations for the 3—7-year old sample from Super et al. (2008).

To confirm the rationale of combining the two age-based questionnaires for this analysis, data were separated for the 6- and 18-month-olds, and then pooled across sites, and the nine temperament scales were again correlated with Difficulty ratings. The nine resulting correlations from the two age groups were compared; no significant differences were found, except regarding Regularity: low Regularity is more highly correlated with Difficulty for the infants than for the toddlers. The pooling across age, used in all other analyses, necessarily loses this distinction to an "average" result, but the doubling of the  $n$  results in a more meaningful analysis overall.

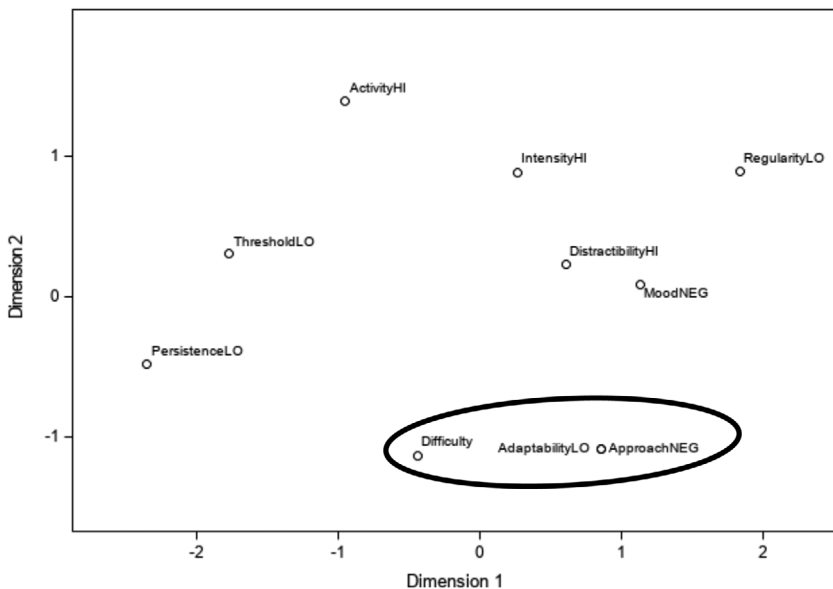
**Intercorrelations Among the Temperament Dimensions.** The divergence among sites in the relationship of Difficulty to the temperament dimensions leads to the third question, regarding the intercorrelations among the nine dimensions. As in the report for the school-transition age children, these correlations vary from site to site, indicating patterns which

**Table 3.2. Correlations of Temperament Dimensions with Difficulty Rating for Younger (6–18 months) and Older (3–7 years) Children**

Site	Age	n	Activity (high)	Regularity (low)	Approach (neg)	Adaptability (low)	Intensity (high)	Mood (neg)	Persistence (low)	Distraction (high)	Threshold (low)
Italy	Y	24	.10	-.17	.32	.10	.27	.22	.07	.10	.03
	O	29	.09	.15	.42*	.43*	.19	-.09	.12	-.05	-
The Netherlands	Y	29	.35+	.17	.19	.68***	.12	.53*	.40*	.31+	.08
	O	63	.38**	.04	.06	.29*	.33**	.30**	.49**	.27+	.10
Poland	Y	19	.17	.16	.53*	.26	-.19	.54*	.15	-.15	.23
	O	44	.32+	.17	.14	.42*	.14	.49**	.24	.13	.13
Spain	Y	28	.14	.21	.22	.30	-.14	.32+	.08	.29	-.05
	O	41	.28	-.10	.11	.50***	.12	.32*	.27	-.05	.26
Sweden	Y	24	.26	.27	.19	-.02	.10	.33	-.17	.22	-.16
	O	33	.06	.02	.14	.27	-.02	.59	.44*	.15	.30
USA	Y	30	.14	.21	.40*	.25	-.15	.73***	.16	-.03	.69***
	O	66	.35***	.19	.01	.51***	.38**	.60***	.18	.06	.00
Pooled	Y	154	.17*	.12	.31***	.27**	.03	.45***	.11	.17	.19
	O	276	.25	.09	.18**	.41***	.21***	.44***	.29***	.11+	.15*

Note: Original scoring used, thus high scores imply "difficult" behavior.  
+ indicates  $p \leq .10$ , \* indicates  $p \leq .05$ , \*\* indicates  $p \leq .01$ , \*\*\* indicates  $p \leq .001$ .

**Figure 3.1. Multidimensional scaling of correlations among nine temperament dimensions and difficulty rating: Italy.**

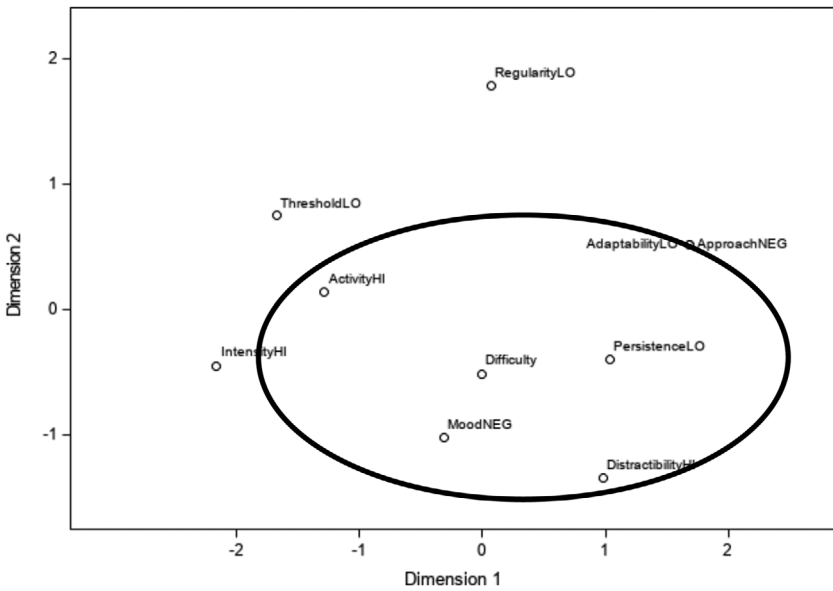


Note: Dark oval encloses items correlating .30 or greater with Difficulty.

help define the meaning of “difficult” in each cultural group. To represent these configurations visually, we applied multidimensional scaling (Weller & Romney, 1990) to the 10 × 10 correlation matrix (nine temperament dimensions plus Difficulty) for each cultural site. Satisfactory solutions were found in three dimensions for all sites (.02 < stress < .07). Plots in the first two dimensions are shown in Figures 3.1–3.6 (6- and 18-month combined). In these figures, the qualities of temperament that correlate .30 (medium effect size) or greater with Difficulty are circled for visual delineation. In addition, a suffix has been added to the name of each temperament dimension to clarify the meaning of a high score, presumably marking difficulty (thus the Activity is noted as ActivityHI and Regularity as RegularityLO).

There are several informative aspects of these figures. The Italian scaling (Figure 3.1) is distinctive in that only low Approach and low Adaptability are placed close to Difficulty. Unique to this site, Mood is excluded from the close correlates for Italy, as it is for the school-transition age children. The consistently close proximity of Approach and Adaptability in all sites indicates substantial conceptual overlap, even though in most cases only one meets the .30 criterion. In the Dutch scaling (Figure 3.2), there is a uniquely strong association of Persistence with Difficulty, and it is visually evident that Distractibility, Activity, Mood, and Adaptability are nearby,

**Figure 3.2. Multidimensional scaling of correlations among nine temperament dimensions and difficulty rating: The Netherlands.**



Note: Dark oval encloses items correlating .30 or greater with Difficulty.

providing further insight into the construct for Dutch parents. In the Spanish scaling, by contrast, Persistence is nearly irrelevant and, like Regularity, off to one side.

Table 3.2 and the six scaling figures illustrate how aspects of temperament contribute to a judgment of difficulty, but not their collective power. Multiple linear regression, using all nine temperament dimensions as predictors of Difficulty, yields an  $R^2$  of .22 ( $p < .0001$ ) and Cohen's  $f^2$  of .28, a medium effect size (Ferguson, 2009); within sites,  $R^2$  varies from .25 in Spain to .67 in the Netherlands. The temperament correlations with Difficulty shown in Table 3.2 appear meaningful in cultural context, and the regression analysis indicates more generally that temperament is a powerful contribution to mothers' perception of difficulty in managing their children.

**Continuity and Change in Difficulty.** Having established similarities and differences among the six cultural samples in the correlates of mothers' difficulty ratings, we can now address the final core question of this study: How consistent across development periods are the temperament–difficulty relationships (keeping in mind that the two age periods involve different mothers and children)? The answer proves to be complex and involves four different patterns.

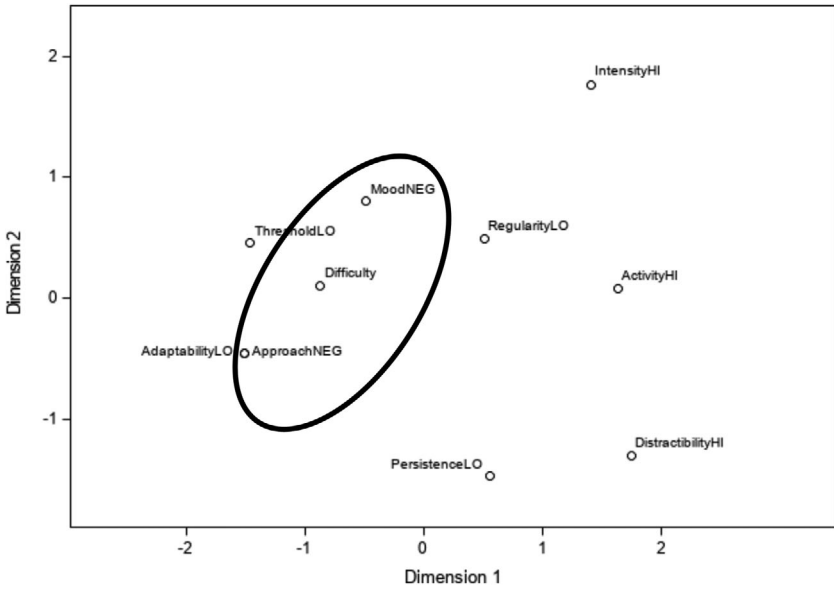
Table 3.3 presents for each cultural site the presence (+) or absence (0) of a moderate relationship ( $r \geq .30$ ) between Difficulty and each of the

**Table 3.3. Comparison of Correlations Between Temperament Dimensions and Difficulty Rating (Younger and Older)**

Site	Activity (high)	Regularity (low)	Approach (neg)	Adaptability (low)	Intensity (high)	Mood (neg)	Persistence (low)	Distractibility (high)	Threshold (low)
Italy	00	00	++	0+	00	00	00	00	00
The Netherlands	++	00	00	++	0+	++	++	++	00
Poland	0+	00	+0	0+	00	++	00	00	00
Spain	00	00	00	++	00	++	00	00	00
Sweden	00	00	00	00	00	++	0+	00	00
USA	0+	00	+0	0+	0+	++	0+	00	+0
Pooled	00	00	+0	++	00	++	00	00	00

Note: The two entries in each cell indicate the degree of correlation in the younger and older age groups. 0 indicates correlation < .30, + indicates correlation  $\geq$  .30

**Figure 3.3. Multidimensional scaling of correlations among nine temperament dimensions and difficulty rating: Poland.**



Note: Dark oval encloses items correlating .30 or greater with Difficulty.

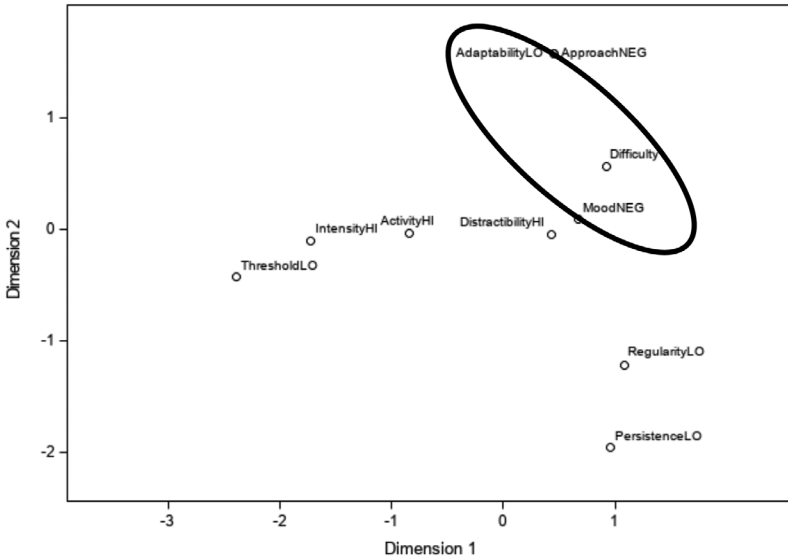
nine temperament dimensions, for the present data and again for the earlier report concerning older children. The patterns in this table illustrate both cross-age continuities and discontinuities in what makes a child difficult to manage.

*Not Difficult.* As shown in Table 3.3, a majority (31, or 57%) of the 54 comparisons (9 dimensions times 6 sites), show temperament ratings within each site unrelated to Difficulty at both preschool and school-transition ages. Distractibility, for example is not related to difficulty at either age in five of the six cultural sites (the exception being the Netherlands), nor is low Approach in three sites (exceptions are Italy, Poland, and the United States).

*Consistently Difficult.* In eleven of the comparisons, there is a significant association with Difficulty in both age groups. Seven of these eleven involve negative Mood and low Adaptability, which tend to be problematic in both age groups. In Spain, they are the only two correlates of Difficulty at both ages. Three of the other four instances of continued association occur in the Netherlands, where high Activity, low Persistence, and high Distractibility are all rated as difficult at both age groups. This last configuration is evident in Figure 3.2, and also in Figure 3.2 of Super et al. (2008) for the school-transition age children. The final case of difficulty at both ages is in Italy, where low Approach is a challenge to manage.



**Figure 3.4. Multidimensional scaling of correlations among nine temperament dimensions and difficulty rating: Spain.**



Note: Dark oval encloses items correlating .30 or greater with Difficulty.

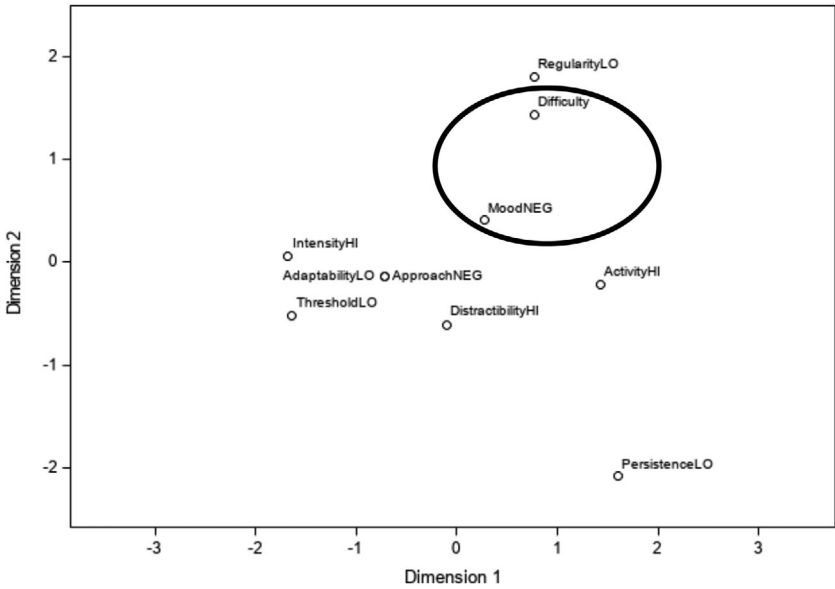
*Older Children, More Difficulty With Activity and Intensity.* There are nine instances in which a temperament characteristic increases in its relevance to Difficulty from the 6- to 18-month period to 3–7 years. High Activity becomes important in this regard in Poland and the US Low Adaptability becomes difficult to manage in three sites (Italy, Poland, and the United States) High Intensity is correlated with Difficulty only in the school-transition age group in the Netherlands and the United States. Finally, low Persistence in becomes problematic in Sweden and the United States.

*Maturity and Declining Difficulty.* Finally, there are three cases where the association of a particular temperament trait with Difficulty decreases with age. Low Approach in both Poland and the United States, and low Threshold in the United States, are correlated with Difficulty for infants and toddlers, but not at the older ages.

## Discussion

Some behavioral dispositions, such as Distractibility and Intensity, may be difficult for particular families to manage, but in general, they appear not to be systematically related to parents' experience of difficulty in day-to-day life with their child. In contrast, other temperamental dispositions, such as

**Figure 3.5. Multidimensional scaling of correlations among nine temperament dimensions and difficulty rating: Sweden.**



Note: Dark oval encloses items correlating .30 or greater with Difficulty.

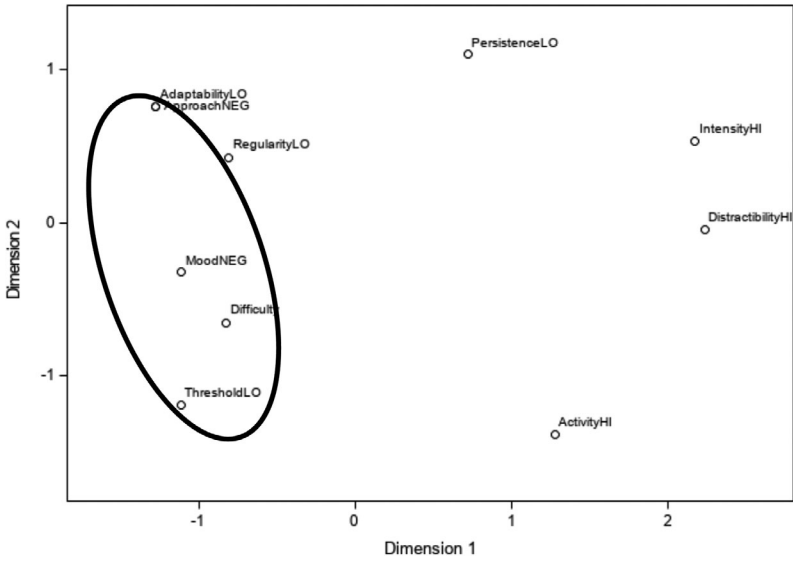
negative Mood, are likely to be challenging for parents in many cultural contexts. In neither of these cases is there any systematic connection with age.

It is not surprising, although infrequently discussed, that some aspects of difficulty do change with age. In part, this is because children grow bigger and stronger, with louder voices and a shifting set of needs; they also, with time, develop more sturdy self-regulation skills. In addition, parents have new expectations and the developmental agenda changes. Children encounter and come to occupy new and often more complex settings, and customary ways of soothing, feeding, entertaining, and supporting shift. In short, the developmental niche changes along with the child. As difficulty is fundamentally a question of fit between the child and the niche, it is not surprising that parents encounter new and different joys and challenges of child management as their child matures.

Most enlightening in the present analyses, however, and most relevant to the understanding of parenting and culture, are the unique thematic continuities that tap into larger networks of cultural meaning. Such thematic threads are most evident in the present study in Italy, Spain, and the Netherlands.

Only in Italy is negative Mood *not* related to the mother's judgment that the child is "difficult." This is true in both age groups—infants and

**Figure 3.6. Multidimensional scaling of correlations among nine temperament dimensions and difficulty rating: United States.**



Note: Dark oval encloses items correlating .30 or greater with Difficulty.

toddlers here, and for 3- to 7-year-olds in our earlier publication. Further, Italy is also the only site where withdrawal from novel situations (high negative Approach) contributes to Difficulty at both age periods. Thus in Italy, unlike other sites, concerns about sociability and emotional closeness seem to present a particular difficulty for parents, but overall negative Mood does not. Other findings in this project also point to this conclusion. In talking about their (older) children, for example, the Italian parents were unique in focusing on social qualities: “easy, even-tempered, well-balanced, and *simpatico*” (Harkness & Super, 2006, p. 73). Likewise, an emphasis on socioemotional learning in the context of close emotional relationships was found for a different sample of Italian parents in a cross-cultural study of parental ethnotheories for infant development (Harkness et al., 2007). Evidently, throughout infancy and childhood, Italian parents’ observations of their child’s social and emotional life are particularly linked with larger cultural models of development and successful functioning (Axia, 1999).

The Spanish concerns for children at both the younger and older ages often engage a theme of courtesy and disobedience. We have shown elsewhere a particular Spanish focus on “good character” and social skills; this is evident in parents’ descriptions of their children (Harkness & Super, 2006), their understanding of qualities that lead to school success (Feng et al., 2020), and an emphasis on developing an

infant's social awareness and emotional closeness within a web of attentive caretakers (Harkness et al., 2007). The respectful, socially graceful child is *bien educado*. Such children are easy to manage and parents are alert to the precursors of those traits.

Dutch parents—and the professionals who advise them—put a strong emphasis on rest and regularity (Harkness et al., 2007; Super et al., 1996); in part, they see this as setting the foundation of a calm, pleasant, and adaptable child, but one who is also strong, “knows exactly what she wants” and has a “strong will.” Even in the first years of life, persistence is a quality that appears to evoke these valued characteristics. In describing their children, the Dutch parents in our study speak of “entertaining oneself” four times more frequently in than do US parents (Harkness & Super, 2006); it is part of the Dutch cultural model that includes independent functioning. One Dutch mother explained: “Independence. I think that everyone must be independent. And I have the idea myself that you don't have to push them to be that way. Playing alone outside is independent. Doing a puzzle by yourself is independent. Making a mess and cleaning it up yourself is independent. So, it all belongs together” (Harkness, Super, & van Tijen, 2000, p. 35). Asked to focus on success in school, these parents imagine a child who is positive in mood and sociable, but not overly dependent on the group (Feng et al., 2020). Children need to learn to “stand on their own feet” (be Persistent) and make their own choices, even in a small country and dense social world—a particularly Dutch framing of “independence” (Harkness et al., 2000). An infant who can play with a toy for more than 10 minutes (Carey & McDevitt, 1978), or a 3-year-old who will look at pictures for over an hour (McDevitt & Carey, 1978), appeals to the Dutch model of independence and *zelf-standigheid*.

Also noteworthy in our findings is some consistency across age and questionnaire in the relative standing of cultural sites in how mothers typically rate their children's temperament (Table 3.2). Infants, toddlers, and school-transition-age children in one site may all be rated, on average, as more or less active and approaching than in another, for example. A number of factors might yield this pattern: genuine genetic or epigenetic differences among the populations (Minkov & Bond, 2017); socialization practices that influence the frequency of behaviors included in the questionnaire items (Friedlmeier, Corapci, & Cole, 2011); subtle differences in the connotation of translated words (Allen, 2007); and finally, a general rater bias (Hoyt, 2000). We cannot evaluate these possibilities with the present data.

## Conclusions

The central finding of this study is that mothers in these six cultural samples, despite many similarities, also see somewhat different patterns of temperament as difficult in their children; some of these cultural patterns remain consistent over the opening years of life. In addition to the

logistical problems posed for many parents by some aspects of temperament, such as high Activity or negative Mood, these results suggest that there is in some cases a more symbolic or thematic challenge to important values. Close analysis of the individual cultural sites provides some insight into this matter, which is of considerable importance for the process of enculturation as the stability of parental concerns over time—when it occurs—suggests a particular continuity in the development niche (Harkness, Super, Barry, Zeitlin, & Long, 2009; Super & Harkness, 1986). This is noteworthy, given the important changes over these years in the settings of daily life and the socially regulated customs of child rearing. Much of the continuity, we suggest, lies in the parents' ethnotheories, that is, the conceptual understanding of children, their development, and what it means to be a good parent (Harkness & Super, 1996). These cultural models are interwoven with more abstract networks of cultural meaning (D'Andrade, 1990), and this complex of interconnections, through contemporary redundancy and thematic repetition of "messages," lies behind the power of enculturation (Super & Harkness, 2002).

This study also highlights, once again, a key tenet of modern temperament theory: that "fit" with the environment of daily life is a more accurate and useful focus of analysis than is an absolutist notion of "difficulty" as a trait. Further, recognition of cultural variability in both logistical and thematic aspects of what parents find difficult has implications for both parent education and clinical practice. Continuities of "difficulty" across infancy and childhood can be revealing of underlying cultural models that may warrant attention in the educational or clinical context, beyond the particulars of a given set of behaviors.

## Acknowledgments

The work reported 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 as well as data: 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, Ellen van der Vlugt, and Saskia van Schaik, 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 and Caroline

Tovatt; in the United States, Sara Harkness, Charles Super, Xin Feng, Marcia Hughes, Archana Khattar, Amy Miller, Beth Muller, and Parminder Parmar.

## Note

1. Items used for the derived etic scales (“r” indicates reverse scoring): Activity (RITQ: 17, 33, 51, 64, 95; TTS: 2, 18r, 24, 47, 52, 58, 64r, 73r), Regularity (RITQ: 67, 28r, 50r, 60r, 73r, 85r; TTS: 1r, 30r, 33, 42, 54, 68, 77r), Approach (RITQ: 24r, 31, 36, 45r, 91; TTS: 5r, 9, 26r, 38r, 45r, 53, 57r, 76r, 87, 96r), Adaptability (RITQ: 6r, 34, 49r, 56r, 94r; TTS: 32r, 56r, 60r, 65r, 78, 88), Intensity (RITQ: 38, 42, 54, 70; TTS: 14, 19, 29, 35, 46, 55r, 74), Mood (RITQ: 2, 23r, 39r, 63r, 81r; TTS: 4r, 8, 20, 25, 43r, 49, 69, 84, 89), Persistence (RITQ: 3, 8r, 25, 32r, 58r, 88; TTS: 15r, 28r, 34, 51, 62r, 71r, 86), Distractibility (RITQ: 7, 44r, 53r, 59r, 80r; TTS: 22, 27, 37r, 41r, 61, 66r, 72, 81, 97), Threshold (RITQ: 11, 37, 47r, 57, 72r, 69; TTS: 16r, 48, 67r).

## References

- Acar, I. H., Ahmetoğlu, E., Özer, İ. B., & Yağlı, Ş. N. (2019). Direct and indirect contributions of child difficult temperament and power assertive parental discipline to Turkish children's behaviour problems. *Early Child Development and Care*, 189, 1–14. <https://doi.org/10.1080/03004430.2019.1698561>
- Ahadi, S. A., Rothbart, M. K., Halverson, C. F., Jr., Kohnstamm, G. A., & Martin, R. P. (1994). Temperament, development, and the Big Five. In *The developing structure of temperament and personality from infancy to adulthood* (pp. 189–207). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Allan, K. (2007). The pragmatics of connotation. *Journal of Pragmatics*, 39(6), 1047–1057. <https://doi.org/10.1016/j.pragma.2006.08.004>
- Axia, G. (1993). *Questionari per la misurazione del temperamento nei primi 3 anni di vita: Traduzione e taratura delle scale: Early Infancy Temperament Questionnaire, per 1–4 mesi; Revised Infant Temperament Questionnaire, per 4–8 mesi; Toddler Temperament Scale, per 12–30 mesi*. Dipartimento di Psicologia dello Sviluppo e della Socializzazione, University of Padova. Padova, Italy.
- Axia, G. (1999). *La timidezza [Shyness]*. Bologna: Il Mulino.
- Berry, J. W. (1989). Imposed etics-emics-derived etics: The operationalization of a compelling idea. *International Journal of Psychology*, 24, 721–735.
- Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality development*. Oxford, England: Wiley-Interscience.
- Cameron, J. R., Rice, D., Hansen, R., & Rosen, D. (1994). Developing temperament guidance programs within pediatric practice. In W. B. Carey & S. C. 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. 226–234). Philadelphia, PA: Brunner/Mazel.
- Carey, W. B. (1970). A simplified method for measuring infant temperament. *Journal of Pediatrics*, 77, 188–194.
- Carey, W. B. (1997). *Understanding your child's temperament*. New York: Simon & Schuster-MacMillan.
- Carey, W. B., & McDevitt, S. C. (1978). Revision of the infant temperament questionnaire. *Pediatrics*, 61(5), 735–739.
- 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. <https://doi.org/10.1037/0012-1649.34.4.677>

- 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. <https://doi.org/10.2307/1131559>
- Chen, X., Wang, L., & Wang, Z. (2009). Shyness-sensitivity and social, school, and psychological adjustment in rural migrant and urban children in China. *Child Development*, 80(5), 1499–1513. <https://doi.org/10.1111/j.1467-8624.2009.01347.x>
- Chess, S., Thomas, A., & Birch, H. (1959). Characteristics of the individual child's behavioral responses to the environment. *American Journal of Orthopsychiatry*, 29, 791–802.
- D'Andrade, R. G. (1990). Some propositions about the relations between culture and human cognition. In J. W. Stigler, R. A. Shweder, & G. Herdt (Eds.), *Cultural Psychology: Essays on comparative human development* (pp. 65–129). Cambridge, MA: Cambridge University Press.
- deVries, M. W., & Sameroff, A. J. (1984). Culture and temperament: Influence on infant temperament in three East African societies. *American Journal of Orthopsychiatry*, 54(1), 83–96.
- Edwards, C. P., Gandini, L., & Giovaninni, D. (1996). The contrasting developmental timetables of parents and preschool teachers in two cultural communities. In S. Harkness & C. M. Super (Eds.), *Parents' cultural belief systems: Their origins, expressions, and consequences* (pp. 270–288). New York: Guilford Press.
- Feng, X., Harkness, S., Super, C. M., Welles, B., Rios Bermudez, M., & Bonichini, S., ... Zylcz, P. O. (2020). Parents' concepts of the successful school child in seven western cultures. In S. Harkness & C. M Super (Eds.), *Cross-Cultural Research on Parents: Applications to the Care and Education of Children. New Directions for Child and Adolescent Development*, 170, 143–170.
- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40, 532–538.
- Friedlmeier, W., Corapci, F., & Cole, P. M. (2011). Emotion socialization in cross-cultural perspective. *Social and Personality Psychology Compass*, 5(7), 410–427. <https://doi.org/10.1111/j.1751-9004.2011.00362.x>
- Fullard, W., McDevitt, S. C., & Carey, W. B. (1984). Assessing temperament in one- to three-year-old children. *Journal of Pediatric Psychology*, 9(2), 205–217. <https://doi.org/10.1093/jpepsy/9.2.205>
- Garstein, M., & Putnam, S. P. (2018). *Toddlers, parents, and culture: Findings from the Joint Effort Toddler Temperament Consortium*. London, UK: Routledge.
- Gartstein, M. A., Slobodskaya, H. R., Kirchhoff, C., & Putnam, S. P. (2013). Cross-cultural differences in the development of behavior problems: Contributions of infant temperament in Russia and US. *International Journal of Developmental Science*, 7(2), 95–104.
- Gouge, N., Dixon, W. E., Jr., Driggers-Jones, L. P., & Price, J. S. (2020). Cumulative sociodemographic risk indicators for difficult child temperament. *The Journal of Genetic Psychology: Research and Theory on Human Development*, 181(1), 32–37. <https://doi.org/10.1080/00221325.2019.1699012>
- Harkness, S. (2000). Developmental agenda. In Alan E. Kazdin (Ed.), *Encyclopedia of psychology* (Vol. 3, pp. 1–4). New York: Oxford University Press.
- Harkness, S., & Super, C. M. (1996). Introduction. In S. Harkness & C. M. Super (Eds.), *Parents' cultural belief systems: Their origins, expressions, and consequences* (pp. 1–23). New York: Guilford.
- Harkness, S., & Super, C. M. (2006). Themes and variations: Parental ethnotheories in western cultures. In K. H. Rubin & O.-B. Chung (Eds.), *Parental beliefs, parenting, and child development in cross-cultural perspective* (pp. 61–79). New York: Psychology Press.
- Harkness, S., Super, C. M., Barry, O., Zeitlin, M., & Long, J. (2009). Assessing the environment of children's learning: The developmental niche in Africa. In E.

- Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 133–155). New York: Springer.
- Harkness, S., Super, C. M., Moscardino, U., Rha, J.-H., Blom, M. J. M., Huitrón, B., ... 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. Raeff, & C. M. Super (Eds.), *Variability in the social construction of the child (New Directions for Child and Adolescent Development)* (Vol. 87, pp. 23–39). San Francisco, CA: Jossey-Bass.
- Hollingshead, A. (1975). *Four-factor index of social status*. Unpublished manuscript. Department of Sociology, Yale University, New Haven, CT.
- Hoyt, W. T. (2000). Rater bias in psychological research: When is it a problem and what can we do about it? *Psychological Methods, 5*(1), 64–86. <https://doi.org/10.1037/1082-989X.5.1.64>
- Kagan, J., Snidman, N., Arcus, D., & Reznick, J. S. (1994). *Galen’s prophecy: Temperament in human nature*. Cambridge, MA: Harvard University Press.
- Keogh, B. K. (1989). Applying temperament research to school. In G. A. Kohnstamm, J. E. Bates & M. K. Rothbart (Eds.). *Temperament in childhood* (pp. 437–450). Oxford, England: Wiley.
- Korn, S. J., & Gannon, S. (1983). Temperament, cultural variation and behavior disorder in preschool children. *Child Psychiatry and Human Development, 13*(4), 203–212.
- Matthews, D. (2020). *Raising a difficult child? Try a “spirited” spin*. Retrieved from <https://www.psychologytoday.com/us/blog/going-beyond-intelligence/201901/raising-difficult-child-try-spirited-spin>
- McClowry, S. G. (2014). *Temperament-based elementary classroom management*. Lanham, MD: Rowman & Littlefield.
- 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.
- Mead, M. (1972). *Blackberry winter: My earlier years*. New York: Simon & Schuster.
- Minkov, M., & Bond, M. H. (2017). A genetic component to national differences in happiness. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being, 18*(2), 321–340. <https://doi.org/10.1007/s10902-015-9712-y>
- Ninio, A. (1979). The naive theory of the infant and other maternal attitudes in two subgroups of Israel. *Child Development, 50*, 976–980.
- Odden, H. L. (2009). Interactions of temperament and culture: The organization of diversity in Samoan infancy. *Ethos, 37*(2), 161–180. <https://doi.org/10.1111/j.1548-1352.2009.01036.x>
- Rothbart, M. K. (1982). The concept of difficult temperament: A critical analysis of Thomas, Chess, and Korn. *Merrill-Palmer Quarterly, 28*(1), 35–40.
- Rothbart, M. K. (2012). *Becoming who we are: Temperament and personality in development*. New York: Guilford.
- Stephens, K. (2020). *Strategies for parenting children with difficult temperament* [Parenting Exchange]. Retrieved from [www.parentingexchange.com](http://www.parentingexchange.com)
- Super, C. M., Axia, G., Harkness, S., Welles-Nyström, B., Zylicz, P. O., Rios Bermudez, M., ... McGurk, H. (2008). Culture, temperament, and the “difficult child” in seven Western cultures. *European Journal of Developmental Science, 2*(1–2), 136–157.
- 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. (1994). The cultural regulation of temperament-environment interactions. *Researching Early Childhood*, 2(1), 59–84.
- Super, C. M., & Harkness, S. (2002). Culture structures the environment for development. *Human Development*, 45(4), 270–274.
- 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.
- Turecki, S. (1989). *The difficult child*. New York: Bantam Books.
- Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology*, 6, Article 1064.
- Weller, S. C., & Romney, A. K. (1990). *Metric scaling: Correspondence analysis*. Newbury Park, CA: Sage.
- Winn, S., Tronick, E. Z., & Morelli, G. A. (1989). The infant and the group: A look at Efe caretaking practices in Zaire. In J. K. Nugent, B. M. Lester, & T. B. Brazelton (Eds.), *The cultural context of infancy, Vol.: Biology, culture, and infant development* (Vol. 1, pp. 87–109). Westport, CT: Ablex Publishing.

*CHARLES M. SUPER is Professor of Human Development and Pediatrics, and Co-Director of the Center for the Study of Culture, Health, and Human Development, at the University of Connecticut, U.S.A. He is interested in cultural variation in early development, and advancing theoretical approaches – in particular the Developmental Niche - to accommodate global findings and facilitate their application to programmatic interventions.*

*SARA HARKNESS is Professor of Human Development and Pediatrics, and Director of the Center for the Study of Culture, Health, and Human Development, at the University of Connecticut, U.S.A. Her research focuses on the cultural construction of parenting, especially parental ethnotheories and their influence on children's development in varied social contexts.*

*SABRINA BONICHINI is Associate Professor of Developmental Psychology at the University of Padua, Italy. Her research interests include cultural differences in parenting, assessment of development, and pediatric psychology.*

*BARBARA WELLES is Emeritus Professor at Fairfield University, U.S.A; previously, she taught at the Karolinska Institute in Stockholm. Her research interests include the effects of swaddling on infant health and development and maternal attachment in Russia, and cross-cultural variation in co-sleeping.*

*PIOTR OLAF ZYLICZ is Associate Professor in Organizational Psychology and Director of the International Center for Research on Leadership at SWPS University, Warsaw Poland. His research interests include leadership, and the cultural and ethical context of adult and children behavior.*

*MOISÉS RIOS BERMÚDEZ is an Adjunct Professor in the Department of Developmental and Educational Psychology at the University of Seville, Spain. His research interests focus on cultural regulation of human development during infancy, childhood, and adolescence, and interventions with families and schools to promote positive development and mental health at the household and community levels.*

*JESÚS PALACIOS is Professor of Developmental and Educational Psychology at the University of Seville, Spain. His research interests focus on children and families, especially children developing in out-of-home contexts, and foster care and adoptive families.*