



ELSEVIER

Infant Behavior & Development 25 (2002) 113–116

**Infant
Behavior &
Development**

Discussion

Measuring infant temperament

Mary K. Rothbart*, Julie Hwang

Department of Psychology, University of Oregon, Eugene, OR 97403-1227, USA

Received 31 December 2001; accepted 31 December 2001

One of the most exciting aspects of infancy research during the past quarter century has been our increasing ability to measure and trace the development of early temperamental dispositions. By temperament we mean individual differences in emotional and motor reactivity and in the attentional capacities that support self-regulation (Rothbart & Bates, 1998; Rothbart & Derryberry, 1981). The 1981 publication of “The assessment of infant temperament: A critique of the Carey Infant Temperament Questionnaire,” by Vaughn, Taraldson, Crichton, and Egeland, provided an important contribution to methodological development in this area. Due in part to Vaughn et al.’s work, recent advances in measurement have allowed early temperament to be empirically linked to later personality (Molfese & Molfese, 2000), the development of conscience and empathy (Kochanska, 1995), and the development of behavior problems (Eisenberg et al., 2001). However, some of the ideas put forward in the paper, e.g., that parental reports are relatively invalid “perceptions,” whereas observer ratings are not, have persisted in the field (Kagan, 1994). We believe that evidence over time has not supported this idea.

Infant temperament has been measured via parent reports, observer ratings, and direct observations, in both home and laboratory. Vaughn et al.’s study made use of one of the first questionnaires designed to assess Thomas and Chess’ nine temperament dimensions and their difficulty construct, Carey’s (1970) Infant Temperament Questionnaire (ITQ). It is important to note that because of psychometric problems with this measure, Carey and McDevitt (1978) had already discarded this measure and put forward a revised and improved version of the ITQ in 1978.

Taking as a starting point the use of the term “perception” to characterize mothers’ reports of temperament in children, and noting that such perceptions may not be accurate, Vaughn et al. sought to examine the validity of the ITQ in relation to observer ratings. They examined

* Corresponding author. Tel.: +1-541-346-4938; fax: +1-541-346-4911.

E-mail address: maryroth@oregon.uoregon.edu (M.K. Rothbart).

mother reports and observer ratings for a sample of one hundred and eighty seven 6-month-old infants from young, low SES, high risk families. “Diagnostic categories” of easy, intermediate low, intermediate high, or difficult derived from scores on the ITQ were related to infants’ behavior during two feedings and a 15-min free play situation.

The diagnostic categories were based on an algorithm for measuring difficulty developed by Carey (1970) and derived from ITQ scale scores of Rhythmicity, Adaptability, Approach, Intensity, and Mood. Vaughan et al.’s observer also rated mothers and infants on 33 variables during feeding and 12 variables during play. Although trained on videotapes, observers made these ratings in vivo. Ratings for feeding items yielded factor-based scores for Mothers’ Caretaking Skills and Affective Behavior as well as Infant Social Responsiveness. Three factor-based scores were derived from the play situation: Maternal Play Skills, Maternal Attitude toward Play, and Infant’s Involvement in Play.

Vaughn et al. found no significant differences between temperament difficult–easy categories and either of the infant factor scores. However, mothers of more difficult infants had higher self-reported aggression, defiance, and anxiety scores, lower scores on Broussard’s measure of knowledge of reciprocity between mother and infant, and lower maternal feelings, in measures taken when their infants were 3 months old. Vaughn et al. concluded that, for their sample, “we have no evidence that the Carey ITQ assesses infant temperament. Rather, it appears to be an assessment of the mother” (Vaughn et al., 1981, p. 15).

Vaughn et al. (1981) has been widely cited in both developmental and psychiatric journals. In our reading of a representative sample of 17 of these studies, we found that approximately half used the study as an example of lack of agreement between mother and observer ratings of infant temperament. In response to this interpretation, Carey (1983) in this journal argued that ratings based on brief home observations could not appropriately be compared with maternal reports derived from quite different contexts. He stated that researchers were interpreting lack of agreement as evidence that the mother is “wrong” or inaccurate, rather than questioning the validity of the observer ratings. Carey’s questions about the appropriateness of the measures are well taken, and Vaughn et al. also raise them in discussing their findings.

Concerns raised about the appropriateness of comparison led to papers like that of Hagekull, Bohlin, and Lindhagen’s (1984) on the validity of parent reports, also published in *IBAD*. Hagekull et al. demonstrated that when both parents and observers kept track of the same infant behaviors over two 4-hr visits, their data converged strongly, ranging from .60 for attentiveness to .83 for sensory sensitivity, using items from the Baby Behavior Questionnaire. They also found modest to moderately strong correlations between mothers’ questionnaires and observers’ reports ranging from .21 to .63.

Other psychometrically appropriate caregiver report measures were also available by 1981, and later research using a number of these measures has provided data in support of their validity (Rothbart, Chew, & Gartstein, 2001). In a recent review of this literature, Rothbart and Bates (1998) concluded that “evidence to date is supportive of the use of parent-report measures of temperament . . . they have established a fair degree of objective validity. In addition, parent report measures have contributed to substantial empirical advances, such as our current understanding of the structure of temperament in relation to the Big Five or Big Three models and their parallels in psychophysiological systems (Bates, Wachs, & Emde, 1994; Rothbart & Bates, 1998, p. 126; Rothbart, Derryberry, & Posner, 1994).”

In the other half of the articles we sampled, authors emphasized Vaughn et al.'s findings of relationships between maternal attributes, characteristics and caregiving styles, and pre- or post-birth temperament measures. More specifically, these studies dealt with the finding that mothers' anxiety is related to their perceptions of "difficulty" in their infants. This very interesting finding is sometimes taken as evidence of biases in mothers' reports based on their own temperamental characteristics, and Vaughn et al. suggested that, "high-anxious mothers are not perceiving their infants' behaviors similarly to the way others would perceive the same behaviors."

Another interpretation, which might operate in addition to bias, as suggested by Diener, Goldstein, and Mangelsdorf (1995), is that the findings may reflect genetic inheritance of infants from their biological mothers. Congruent with this interpretation, research by Matheny, Wilson, and Thoben (1987), using laboratory measures of temperament, found that mothers' personality measures correlated not only with maternal reports about their infants' temperament, but also with their children's behavior as independently evaluated in the laboratory. Matheny et al. followed Bates and Bayles' (1984) strategy of dividing parent-report variance into objective and subjective components, and they found that temperament scale scores based on mothers' reports correlated moderately to strongly with laboratory scores of temperament at 12–24 months. They concluded that "the objective component of maternal ratings was clearly demonstrable and prominent" (Matheny et al., 1987, p. 324).

Vaughn et al.'s study represented one of the first to systematically investigate the validity of parent report in the measurement of temperament. Our assessment of more recent research in this area is that a number of subsequent validation studies inspired by their work have been a good deal more supportive of parent-report measures. Parent report is by no means a perfect measure, but all of the approaches to the assessment of temperament, including laboratory and observer measures, are also subject to flaws and biases (Rothbart & Bates, 1998; Rothbart & Goldsmith, 1985). However, when we use different measures, each with different biases and sources of error, we can look for convergence of findings across methods. Such convergence can allow us, "to reach many reasonably strong conclusions *despite* the manifest flaws in all psychological research . . . many different methodologies and experimental techniques exist in psychology, and . . . this should be viewed as a strength of the field rather than a weakness, as is sometimes done" (Stanovich, 1986, p. 112).

In conclusion, Vaughn et al.'s research showed the way in validation of temperament measures, even though, as a result of subsequent studies, we may now come to a different conclusion about the use of parent report than they did in their evaluation of the 1970 ITQ. They also raised important questions about the relationship between parent personality/temperament, parent reports, and other measures of temperament, important topics for current and future research.

References

- Bates, J. E., & Bayles, K. (1984). Objective and subjective components in mothers' perceptions of their children from age 6 months to 3 years. *Merrill-Palmer Quarterly*, 30, 111–130.
- Bates, J. E., Wachs, T. D., & Emde, R. N. (1994). Toward practical uses for biological concepts of temperament. In J. E. Bates & T. D. Wachs (Eds.), *Temperament: Individual differences at the interface of biology and behavior*. Washington, DC: American Psychological Association.

- Carey, W. B. (1970). A simplified method for measuring infant temperament. *The Journal of Pediatrics*, 77, 188–194.
- Carey, W. B. (1983). Some pitfalls in infant temperament research. *Infant Behavior and Development*, 6, 247–254.
- Carey, W. B., & McDevitt, S. C. (1978). Revision of the Infant Temperament Questionnaire. *Pediatrics*, 61, 735–739.
- Diener, M. L., Goldstein, L. H., & Mangelsdorf, S. C. (1995). The role of prenatal expectations in parents' reports of infant temperament. *Merrill-Palmer Quarterly*, 41, 172–190.
- Eisenberg, N., Cumberland, A., Spinrad, T. L., Fabes, R. A., Shepard, S. A., Reiser, M., Murphy, B. C., Losoya, S. H., & Guthrie, I. K. (2001). The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. *Child Development*, 72, 1112–1134.
- Hagekull, B., Bohlin, G., & Lindhagen, K. (1984). Validity of parental reports. *Infant Behavior and Development*, 7, 77–92.
- Kagan, J. (1994). *Galen's prophecy: Temperament in human nature*. New York: Basic Books.
- Kochanska, G. (1995). Children's temperament, mothers' discipline, and security of attachment: Multiple pathways to emerging internalization. *Child Development*, 66, 597–615.
- Matheny, A. P., Jr., Wilson, R. S., & Thoben, A. S. (1987). Home and mother: Relations with infant temperament. *Developmental Psychology*, 23, 323–331.
- Molfese, V. J., & Molfese, D. L. (2000). *Temperament and personality development across the life span*. Mahwah, NJ: Erlbaum.
- Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional and personality development* (5th ed., pp. 105–176). New York: Wiley.
- 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.
- Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (Vol. 1, pp. 37–86). Hillsdale, NJ: Erlbaum.
- Rothbart, M. K., Derryberry, D., & Posner, M. I. (1994). A psychobiological approach to the development of temperament. In J. E. Bates & T. D. Wachs (Eds.), *Temperament: Individual differences at the interface of biology and behavior* (pp. 83–116). Washington, DC: American Psychological Association.
- Rothbart, M. K., & Goldsmith, H. H. (1985). Three approaches to the study of infant temperament. *Developmental Review*, 5, 237–260.
- Stanovich, K. E. (1986). *How to think straight about psychology*. Glenview, IL: Scott. Forestman.
- Vaughn, B. E., Taraldson, B. J., Crichton, L., & Egeland, B. (1981). The assessment of infant temperament: A critique of the Carey Infant Temperament Questionnaire. *Infant Behavior and Development*, 4, 1–17.