



HHS Public Access

Author manuscript

J Autism Dev Disord. Author manuscript; available in PMC 2015 May 10.

Published in final edited form as:

J Autism Dev Disord. 2006 July ; 36(5): 637–642. doi:10.1007/s10803-006-0110-5.

Using Carey Temperament Scales to Assess Behavioral Style in Children with Autism Spectrum Disorders

Susan L. Hepburn and

Department of Psychiatry, University of Colorado Health Sciences Center, 4200 E. Ninth Street, Box C 268-31, Denver, CO 80262-0234, USA, Susan.Hepburn@UCHSC.edu

Wendy L. Stone

Vanderbilt Children's Hospital, Nashville, TN, USA

Abstract

Many researchers have suggested that temperament information could be useful for understanding the behavioral variability within the autism spectrum. The purpose of this brief report is to examine temperament profiles of 110 children with ASD (ages 3–8 years, 61 with Autistic Disorder, 42 with PDD-NOS; and 7 with Asperger Disorder) via a commonly used parent report measure of child temperament. Internal consistency of temperament dimensions, test–retest reliability, descriptions of means and standard deviations are examined, relative to previously published norms. Internal consistency of the dimensions and test–retest reliability were comparable to published norms; however, children with autism were rated as presenting with more extreme scores than typically-developing children on several dimensions. Limitations and implications for future work are discussed.

Keywords

Temperament; Behavioral variability; Assessment; Autism

Within the past two decades, the role of child temperament in the adaptation of children in special populations has received increased attention (Goldberg & Marcovitch, 1989; Hatton, Bailey, Hargett-Beck, Skinner, & Clark, 1999; Rothbart & Jones, 1999). However, few studies have examined the psychometric characteristics of temperament assessments in children within specific diagnostic categories (e.g., autism, fragile × syndrome). One cannot assume that the psychometric properties of a psychological instrument will be consistent across samples with different characteristics (Anastasi, 1989).

To date, the construct validity of temperamental dimensions has not been assessed empirically in children with Autism Spectrum Disorders (ASD), even though many researchers have commented on their potential usefulness for understanding the behavioral variability in children with ASD (Eaves, Ho, & Eaves, 1994; Kasari & Sigman, 1997; Konstantareas & Homalidis, 1989; Konstantareas & Stewart, 2001; Leibowitz, 1991).

Individual differences in temperament of children with ASD have been hypothesized to be related to the development of maladaptive behaviors (Eaves et al., 1994; Konstantareas & Homatidis, 1989), responsiveness of parents (Kasari & Sigman, 1997), and parenting stress (Bristol & Schopler, 1984; Holroyd & MacArthur, 1976). It is also possible that information concerning child temperament could be helpful in describing various manifestations of the behavioral phenotype. For example, a portion of children with autism spectrum disorders present with intense fears, anxiety, and insistence on sameness, leading some researchers to consider an anxious phenotypic subtype of the disorder (Hollander, 2004). Conversely, clinical observations also suggest that a portion of children on the autism spectrum present with a notable lack of fear or nervousness, which may also predispose them to risk of physical harm. This latter phenotypic variation might be described as an “impulsive” subgroup, although it has yet to be empirically validated.

We present data from a study of the use of the Carey Temperament Scales (Carey & McDevitt, 1995) in children with autism. Internal consistency of temperament dimensions, test–retest reliability, descriptions of means and standard deviations are examined, relative to previously published normative data. Implications for research and practice are discussed.

Method

Participants

Participants were 110 children with ASD and their mothers. Eligibility criteria included: (a) a documented diagnosis of Autistic Disorder, Pervasive Developmental Disorder Not Otherwise Specified [PDD-NOS], or Asperger’s Disorder; (b) chronological age (CA) between 36 and 96 months; and (c) the absence of severe sensory, motor, or medical conditions. Participants were recruited either through a data-base of research participants at a regional child development center ($n = 86$) or through advertisements posted in the newsletter of the local chapter of the Autism Society of America and at a local speech and hearing center ($n = 24$). The final sample consisted of 61 children with Autistic Disorder, 42 children with PDD-NOS, and 7 children with Asperger’s Disorder. Eighty children (73%) had been diagnosed at a regional child development center by a multidisciplinary team that included a clinical psychologist, pediatrician, and speech-language pathologist, while 30 children (27%) had been diagnosed by a clinical psychologist or a child psychiatrist in the community. Developmental testing data were available for approximately 60% of participants. Measures used included: Bayley Scales of Infant Development, Kaufman Assessment Battery for Children, Leiter International Performance Scale—Revised, and the Weschler Preschool Scales of Intelligence. Approximately 70% of the tested sample presented with significant cognitive delays (i.e., IQ of less than 70). The sample is predominantly male. Most participants were Caucasian and from well-educated families. Participant characteristics are listed in Table 1.

Test–retest reliability was assessed in a subsample comprising 18 children (10 with Autistic Disorder, 6 with PDD-NOS, 2 with Asperger’s Disorder). There were no significant differences between the test–retest subsample and those in the total sample on age, gender, number of children in the family, or mothers’ level of education, $P > .52$.

Measures

Demographic information—Demographic information was collected through parental report using a one-page form that included questions about the child’s date of birth, race, gender, mother’s level of education, number of children in the family, birth order, diagnosis, and source of diagnosis.

Child temperament—The Behavioral Style Questionnaire (BSQ) of the Carey Temperament Scales (CTS; McDevitt & Carey, 1996) was chosen as the temperament measure for these studies because of its underlying theoretical model, the strength of its psychometric properties, its widespread use in research examining the relation of temperament to problem behaviors (Cameron, 1978), and its ease of administration. The BSQ is a 100-item parental report measure that yields scores on the nine dimensions of temperament first outlined by Thomas, Chess, Birch, Hertzig, and Korn (1963) and refined by Thomas, Chess, and Birch (1968) and Chess and Thomas (1996; see Table 2). Items are phrased as statements about a child’s behavior, and parents rate how often the child behaves in the way described in the statement using a score ranging from 1 (almost never) to 6 (almost always). Items are recoded as necessary so that higher dimension scores are indicative of greater challenge or difficulty. For example, a high score on the Approach/Withdrawal dimension indicates more withdrawal from novelty; and a high score on Rhythmicity indicates more irregular, or less rhythmic daily functions. Summary scores for each dimension were computed by dividing the sum of items on each dimension by the number of ratings available as described by McDevitt and Carey (1996).

Studies of the psychometric properties of the BSQ support its reliability and validity in assessing the temperaments of typically developing children ages 3–8 years (Carey, 1986; Carey & McDevitt, 1989; McDevitt & Carey, 1978). The internal consistency of dimension scores ranged from .47 to .80, with a median of .70 (McDevitt & Carey, 1996). Internal consistency was lowest for Rhythmicity and Threshold of Responsiveness. All other scores were above .60.

Procedures

Participants recruited from the database were sent letters announcing the study and 42% indicated interest in participating. Informed consent was obtained from all parents prior to enrollment in the study. Initial temperament data (using the BSQ) and demographic data had already been collected for this subgroup as part of their involvement in a longitudinal study of autism. Participants recruited through advertisements were asked to complete the temperament questionnaire and the demographic form. They were offered monetary compensation (\$10.00) for their time and received a brief written interpretation of their child’s temperament. The test–retest reliability subsample was sent a second BSQ to complete approximately one month after completing the first. This subsample was recruited exclusively through advertisements and received an extra \$10.00 for completing the second form. All of the mothers recruited for the test–retest study returned both sets of forms.

Results and discussion

Internal consistency of temperament dimensions

Cronbach's coefficient alpha (Cronbach, 1951) was used to measure the internal consistency of the nine temperament dimensions. Coefficient alphas are reported in Table 2. Six of the nine dimensions demonstrated internal consistency that was comparable to results obtained in other psychometric studies of typically-developing children (i.e., over .60) (Baydar, 1995; McDevitt & Carey, 1978, 1996). The normative samples included children ages 3 to 8 years of age, who presented with no known developmental delay or disorder. The normative sample was split evenly by gender and was more ethnically diverse than the sample described in this study (see McDevitt & Carey, 1996).

In the present study, the coefficients of the Rhythmicity, Mood and Threshold of Responsiveness dimensions were quite low, suggesting that perhaps these dimensions measure more than one aspect of functioning. The developers of the BSQ also reported very similar, low alpha coefficients for Rhythmicity and Threshold of Responsiveness (see McDevitt & Carey, 1996).

The alpha coefficient for Mood was higher in their psychometric study ($\alpha = .66$) than in the present study. A possible explanation for this finding is that this dimension has the highest proportion of items that presume functional language (e.g., "the child complains when tired"). Because language delays and disorders are common in children with autism spectrum disorders, these items may have been endorsed by parents as "almost never," a score that would reflect more the child's expressive language abilities than his or her emotional responses. Although the instructions on the BSQ suggest skipping items that do not apply, some parents may have responded to these questions even when their child did not use functional speech.

Although McDevitt and Carey (1996) accepted an alpha level of .60 as determining acceptability, others have suggested that higher alpha levels are required to demonstrate internal consistency of subscales (Huck, 2000). More stringent criteria of .70 or higher would suggest that only Approach/Withdrawal, Activity Level, Adaptability, Intensity and Distractibility are internally consistent in children with ASD. At an alpha of .80, only Adaptability and Distractibility are cohesive subscales in children with ASD in this sample, and only the subscale for Approach exceeds this criterion in the normative sample described by McDevitt and Carey (1996).

Test-retest reliability

Consistent with the methods used by other researchers concerning test-retest reliability or short-term stability of temperament across time (see Baydar, 1995), correlational analysis of temperament dimension scores obtained at two points in time was conducted. The mean length of time between assessment was 26 days, with a range of 22–35 days. Pearson product moment correlations ranged from .68 to .93 and are reported in Table 2, along with the test-retest coefficients reported by McDevitt and Carey (1996) for children with typical development. These results suggest that mothers' reports of child temperament are fairly stable across short periods of time and therefore probably not unduly influenced by transient

factors, such as mothers' mood or a recent behavioral episode with the child. Test-retest reliability coefficients were very similar to those obtained in the normative reference sample.

Description of temperament scores in children with ASD

Scores for the nine temperament dimensions were computed for all subjects ($n = 110$) and are displayed in Fig. 1. For descriptive purposes, normative data reported by McDevitt and Carey (1996) derived from a sample of 350 typically developing children of similar ages (see p. 7 for brief description) are also provided in Fig. 1.

In order for temperament to be a valid construct for understanding individual differences in behavior across children, it is necessary to establish that sufficient variability exists for each dimension within the sample of children with ASD. The standard deviations obtained for the sample of children with ASD ranged from .7 to .9, which is consistent with those reported in the previously published normative studies (McDevitt & Carey, 1996). These findings are also consistent with those reported by Bailey, Hatton, Mesibov, Ament, and Skinner (2000), DiLavore (1991), and Pollack (1998) who found similar variability in dimension scores within smaller samples of children with autism. Examination of the distributions of scores reveals that all of the dimensions were normally distributed within the sample, as evidenced by skewness and kurtosis values: Activity level: skewness = $-.24$, kurtosis = $.12$; Rhythmicity: skewness = $-.03$, kurtosis = $-.39$; Approach/Withdrawal: skewness = $-.18$, kurtosis = $-.08$; Adaptability: skewness = $-.03$, kurtosis = $-.37$; Intensity: skewness = $-.05$, kurtosis = $-.26$; Mood: skewness = $-.17$, kurtosis = $-.18$; Distractibility: skewness = $-.29$, kurtosis = $.24$; Persistence: skewness = $.26$, kurtosis = $-.003$; Threshold of Responsiveness: skewness = $.04$, kurtosis = $-.40$.

Examination of the mean scores reveals that, as a group, children with ASD scored at least one standard deviation above the means established for typically developing children on the dimensions of Adaptability and Persistence, and at least one standard deviation below the mean on the Threshold of Responsiveness dimensions. These ratings indicate that, as a group, children with autism were reported to be less adaptable, less persistent, and to require more intense stimulation from the environment in order to obtain a response. These results are also consistent with prior research on smaller samples (Bailey et al., 2000; DiLavore, 1991; Pollack, 1998). The finding that children with ASD are less adaptable to changes in their routines or environments is not surprising, given that insistence on sameness is part of the behavioral phenotype of the disorder. It is important to note that while many parents report that their child with ASD is extremely persistent in completing his or her favorite activities, this behavioural propensity appears to be context-specific and is not generalized across activities and settings.

Results for Threshold of Responsiveness were some-what unexpected, in that children with autism appeared to be less responsive to small changes in their environment. This finding runs contrary to the reports by parents and teachers that children with autism sometimes react in exaggerated ways to small changes in sensory aspects of their environment (Rogers, Hepburn, & Wehner, 2003). One explanation may be that the items in this scale do not reflect a unitary construct (see internal consistency results), but rather reflect responsiveness

to both social (4 items) and non-social (7 items) events. Clearly, children with ASD will have differential sensitivities to these varying sorts of stimuli. While they may be under-responsive to social events (e.g., “The child responds to mild disapproval by the parent”), many are reported to be highly responsive to non-social changes in stimulation (e.g., “Unusual noises (sirens, thunder) interrupt my child’s behavior”). This finding is consistent with the work of Dawson et al. (2004), who found that children with autism demonstrated a specific impairment in social orienting. Thus, the Threshold of Responsiveness dimension is not a unitary construct for children with ASD and is difficult to interpret.

Temperament data are most useful in the assessment of individual differences; thus one could argue that group comparisons of means and standard deviations are less informative than an examination of individual performance. Over half of the children with ASD were reported to fall within the average range in the following dimensions: Activity Level, Rhythmicity, Approach/Withdrawal, Mood, and Distractibility. Approximately two-thirds of the children were reported to be non-adaptable, one-half of the children were reported to be fairly mild in emotional intensity, and one-third were reported to be primarily negative in mood. Over one-half were described as non-persistent. One-third were reported to be very difficult to distract.

In summary, the psychometric properties of 6 of the 9 dimensions of temperament appear to be adequate. Rhythmicity, Mood and Threshold of Responsiveness were not reliable dimensions within this group of children, primarily due to low internal consistency. Adaptability and Distractibility demonstrated the strongest internal consistency. It is important to note that the strong alpha reliabilities obtained for these two dimensions could be related to the diagnosis of ASD—of all of the temperament dimensions, low distractibility from one’s own activities and difficulty accommodating changes in one’s experiences are most closely associated to the core features of autism.

Test–retest reliability was comparable to previous data reported by McDevitt and Carey (1996), and adequate for all dimensions. The standard deviations of temperament dimensions within the sample of children with ASD were similar to those reported for typically-developing children; however the means for children with ASD were different. Specifically, while the majority of children with ASD were reported to be within the average range on 5 of the 9 dimensions, from one-third to two-thirds of the children demonstrated elevated scores on several dimensions in the following directions: highly active, arrhythmic, withdrawn, non-adaptable, negative in mood, mild in emotional reactions, non-persistent, and less responsive to changes in the environment. These differences from the normative sample do not appear to compromise the reliability or validity of the assessment tool, but might rather appear to reflect differences in overall behavioral style.

Limitations of this study include: (a) lack of a comparison sample specifically recruited to address these research questions, (b) reliance on clinical Judgment for diagnosis of autism, (c) incomplete data on developmental functioning of all participants. However, the data provided herein suggest that children identified as having autism present with different behavioral styles, which may be quite relevant when designing interventions, assisting parents in preventing problem behaviors, and identifying environments that best suits the

child's behavioral style. These results support the use of temperament information in this population; however the BSQ may be useful for obtaining information on some dimensions of behavioral functioning (e.g., Adaptability), but not all dimensions (e.g., Threshold of Responsiveness). Future work examining the reliability and validity of other temperament measures will be helpful.

Possible next steps for temperament research include: (1) longitudinal investigations incorporating multiple measures of child temperament, neuropsychological functioning, symptoms of autism, and problem behaviors across time; (2) evaluating the use of temperament information to determine the intensity, focus, and type of intervention that is appropriate for a particular child; or (3) examining if temperament of siblings and other family members could be useful in defining the boundaries of the broader phenotype.

Acknowledgements

This project was supported in part by NIMH #MH50620 (Stone), the Merck Dissertation Scholars Program (Hepburn) and the Collaborative Programs for Excellence in Autism Research funded by the National Institute of Child Health and Development (#HD35468). The authors express their gratitude to the following people who contributed to this endeavor: Barbara Hepburn, William MacLean, Steven Warren, Linda Ashford, Bahr Weiss, Bob Newbrough.

References

- Anastasi A. Ability testing in the 1980's and beyond: Some major trends. *Public Personnel Management*. 1989; 18:471–485.
- Bailey DB, Hatton DD, Mesibov G, Ament N, Skinner M. Early development, temperament, and functional impairment in autism and fragile × syndrome. *Journal of Autism and Developmental Disorders*. 2000; 30:49–59. [PubMed: 10819120]
- Baydar N. Reliability and validity of temperament scales of the NSLY child assessments. *Journal of Applied Developmental Psychology*. 1995; 16:339–370.
- Bristol, M.; Schopler, E. A developmental perspective on stress and coping in families of autistic children. In: Blacher, J., editor. *Severely handicapped children and their families*. New York: Academic Press; 1984. p. 91-141.
- Cameron J. Parental treatment, children's temperament and risk of childhood behavior problems. II: Initial temperament, parent attitudes, and their incidence and form of behavioral problems. *American Journal of Orthopsychiatry*. 1978; 48:140–147. [PubMed: 623213]
- Carey, WB. Temperament and clinical practice. In: Chess, S.; Thomas, A., editors. *Temperament in clinical practice*. New York: Guilford; 1986. p. 239-246.
- Carey, WB.; McDevitt, SC. *Clinical and educational applications of temperament research*. Berwyn, PA: Swets North America; 1989.
- Carey, WB.; McDevitt, SC. *The Carey temperament scales*. Scottsdale, AZ: Behavioral-Developmental Initiatives; 1995.
- Chess, S.; Thomas, A. *Temperament: Theory and practice*. New York: Bruner-Mazel; 1996.
- Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951; 16:297–334.
- Dawson G, Toth K, Abbot R, Osterling J, Munson J, Estes A, Liaw J. Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology*. 2004; 40(2):271–283. [PubMed: 14979766]
- DiLavore, PC. Unpublished doctoral dissertation. University of North Carolina, Chapel Hill; 1991. *Maternal ratings of temperament in children with autism and children with Down syndrome: A comparative study*.
- Eaves LC, Ho HH, Eaves DM. Subtypes of autism by cluster analysis. *Journal of Autism and Developmental Disorders*. 1994; 24:3–22. [PubMed: 8188572]

- Goldberg, S.; Marcovitch, S. Temperament in developmentally disabled children. In: Kohnstamm, GA.; Bates, JE.; Rothbart, MK., editors. *Temperament in childhood*. New York: John Wiley and Sons; 1989. p. 387-403.
- Hatton D, Bailey DB, Hargett-Beck M, Skinner M, Clark RD. Behavioral style of young boys with fragile × syndrome. *Developmental Medicine & Child Neurology*. 1999; 41:625–632. [PubMed: 10503921]
- Hollander, E. Fluoxetine for repetitive behaviors and Valproate for irritability in autism. In: Volkmar, F., Chair, editor. *Intervention symposium*. Symposium conducted at the annual meeting of the Collaborative Programs of Excellence in Autism; Bethesda, MD. 2004 May.
- Holroyd J, McArthur D. Mental retardation and stress on the parents: A contrast between Down's syndrome and childhood autism. *American Journal of Mental Deficiency*. 1976; 80:431–438. [PubMed: 129006]
- Huck, S. *Reading statistics and research*. Boston: Allyn & Bacon; 2000.
- Kasari C, Sigman M. Linking parental perceptions to interactions in young children with autism. *Journal of Autism and Developmental Disorders*. 1997; 27:39–57. [PubMed: 9018581]
- Konstantareas, MM.; Stewart, K. Affect regulation and temperament in children with pervasive developmental disorder. Paper presented at the Society for Research in Child Development Conference; Minneapolis, MN. 2001 Apr.
- Konstantareas MM, Homatidis S. Assessing symptom severity and stress in parents of autistic children. *Journal of Child Psychology and Psychiatry*. 1989; 30:459–470. [PubMed: 2745596]
- Leibowitz G. Organic and biophysical theories of behavior. *Journal of Developmental and Physical Disabilities*. 1991; 3:201–243.
- McDevitt SC, Carey WB. The measurement of temperament in 3–7 year old children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 1978; 19:245–253.
- McDevitt, SC.; Carey, WB. *Manual for the behavioral style questionnaire*. Scottsdale, AZ: Behavioral-Developmental Initiatives; 1996.
- Pollack CF. An examination of temperamental variation and range among autistic children as reported by caregivers. *Dissertation Abstracts International, Section B: the Sciences & Engineering*. 1998; 59(1-B) (UMI No. 0443).
- Rogers SJ, Hepburn S, Wehner E. Parent reports of sensory symptoms in toddlers with autism and those with other developmental disorders. *Journal of Autism and Developmental Disorders*. 2003; 33(6):631–642. [PubMed: 14714932]
- Rothbart, MK.; Jones, LB. Temperament: Developmental perspectives. In: Gallimore, R.; Bernheimer, LP., editors. *Developmental perspectives on children with high-incidence disabilities*. The LEA series on special education and disability. Mahwah, NJ: Lawrence Erlbaum Associates; 1999. p. 33-53.
- Thomas, A.; Chess, S.; Birch, H. *Temperament and behavior disorders in children*. New York: New York University Press; 1968.
- Thomas, A.; Chess, S.; Birch, HG.; Hertzig, ME.; Korn, S. *Behavioral individuality in early childhood*. New York: New York University Press; 1963.

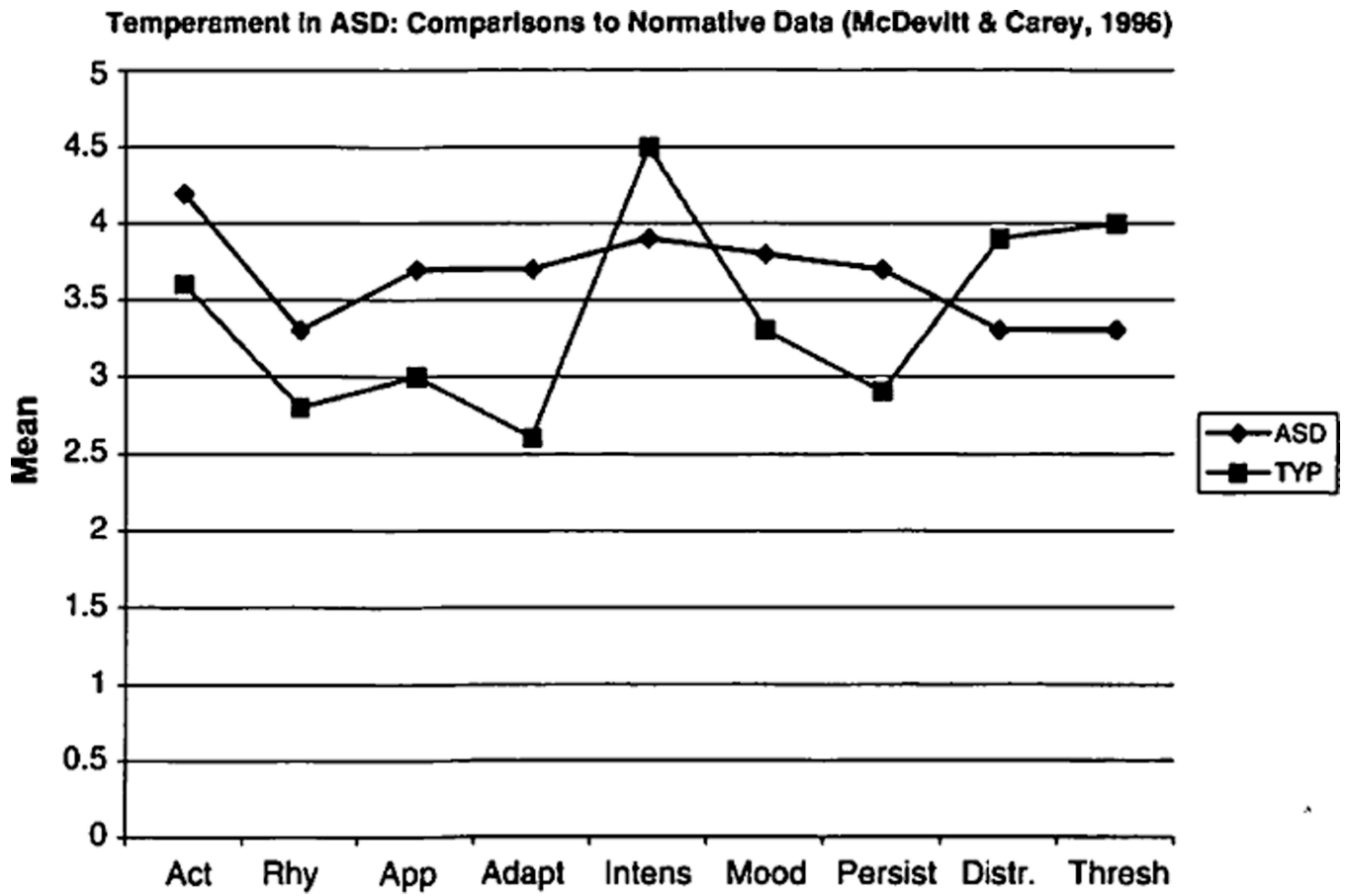


Fig. 1. Temperament in ASD: Comparisons to normative data (McDevitt and Carey, 1996)

Table 1Participant characteristics ($n = 110$)

CA (mos.)	
M (SD)	57.3 (15.4)
Range	23–94
MA (mos.) (n = 66)	
M (SD)	47.5
Range	18–70
Race (%)	
Caucasian	83
African - American	14
Other	3
Gender (%)	
Male	86
Maternal education (%)	
High school	35%
College/graduate work	65 %

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Dimensions of temperament and psychometric properties in children with ASD ($n = 110$) and normative sample (McDevitt & Carey, 1996)

Table 2

Dimension	Definition ¹	Cronbach's Alpha		Test-retest r	
		ASD	Norms ²	ASD	Norms
Activity level	Motor component in a child's functioning	.69	.76	.87	.93
Rhythmicity	Regularity; predictability in daily functions	.48	.48	.93	.80
Approach	Initial response to a novelty	.73	.80	.82	.94
Adaptability	Behavioral flexibility in changing context	.84	.72	.78	.85
Intensity	Energy level of an emotional response	.76	.71	.83	.75
Mood	Tone of overall affect (positive or negative)	.51	.66	.87	.87
Persistence	Continuation of activity in face of obstacles	.64	.60	.77	.70
Distractibility	Effectiveness of extraneous stimuli in altering the direction of ongoing behavior	.79	.70	.68	.82
Threshold of response	The intensity level of stimulation that is necessary to evoke a discernible response	.40	.47	.83	.67

¹Chess and Thomas 1996

²Carey and McDevitt 1995