

Is the Discrepancy Criterion for Defining Developmental Disorders Valid?

Murray J. Dyck¹, David Hay¹, Mike Anderson², Leigh M. Smith¹, Jan Piek¹
and Joachim Hallmayer³

¹ School of Psychology, Curtin University of Technology

² School of Psychology, University of Western Australia

³ Department of Psychiatry and Behavioral Sciences, Stanford University

running head: DEFINING DEVELOPMENTAL DISORDERS

Correspondence to:

Murray J. Dyck, School of Psychology, Curtin University of Technology, GPO Box U1987, Perth, Western Australia, 6845.

Tel: 61 8 9266 3442

Fax: 61 8 9266 2464

Email: m.dyck@curtin.edu.au

Abstract

Background: Most developmental disorders are defined by an achievement discrepancy in which achievement on one or more specific abilities is substantially less than a person's measured intelligence. We evaluated the validity of this discrepancy criterion by assessing parameters that determine variability across abilities and by assessing relationships between achievement discrepancies and behavioral disturbances. **Methods:** Measures of intelligence, language, motor coordination, empathic ability, and attentional control were administered to a representative sample of 390 children aged 3 to 12 years. Parent ratings of child behavior were obtained. **Results:** Results indicate that achievement discrepancies are a function of the correlation between ability measures, the shape of the ability distributions, and position on an index ability dimension. Discrepancies in achievement were not related to behavioral disturbance, but underachievement relative to age peers was invariably related to behavioral disturbance. **Conclusions:** We conclude that developmental disorders need to be redefined in ways that are consistent with how Mental Retardation is now defined, by (a) underachievements, (b) of defined magnitude, (c) using standardized measures, (d) with known relations to normal development, and (e) concurrent deficits on standardized measures of impaired function.

Key Words: Intelligence, Motor Coordination, Language, Empathy, Attention, Developmental Disorders, Normal Development

Abbreviations: MAND, McCarron Assessment of Neuromuscular Development; ERS, Emotion Recognition Scales; CELF, Clinical Evaluation of Language Fundamentals;

Is the Discrepancy Criterion for Defining Developmental Disorders Valid?

The DSM-IV (American Psychiatric Association, 1994) and ICD-10 (World Health Organisation, 1992) classify psychopathology on the basis of phenomenology rather than on the supposed etiology of disorders (Cantwell, 1996). This approach, suggested by Feighner, Robins, Guze et al. (1972) and adopted since DSM-III, recognizes that it is easier to achieve consensus on phenomenological descriptions of disorder than on theoretical conceptions of disorder. But this pragmatism forces us to question “the possibility of defining deviations without defining a norm, or describing psychopathology without first proposing a psychology” (Sedler, 1994). The problem is most acute in the classification of “disorders usually first diagnosed in infancy, childhood or adolescence” (DSM-IV), which includes the so-called developmental disorders.

The term “developmental” implies that problems start early in life and interfere with “normal” development (Rispen & van Yperen, 1997). The hallmark of a developmental disorder is deviation and/or delay from normal development, and each developmental disorder is characterized by a specific pattern of developmental delay. For example, in DSM-IV the syndrome of Mental Retardation is defined by relatively even delays in multiple aspects of development, the specific developmental disorders by deviations in a single or related aspects of development, and the group of pervasive developmental disorders by uneven patterns of development across domains (Cohen, Paul, & Volkmar, 1986).

Diagnosis of developmental disorders requires the identification of some pattern of deviant development, which means that they cannot be validly defined unless what is normal has first been defined. It may be possible to know that something is wrong with a child’s development, but we can’t know what that something is unless the child’s characteristics can be compared with relevant norms. We may dispute how great

deviations from a norm must be before a child's development is regarded as abnormal, but this dispute should be based on knowledge of how much variability in development is to be expected.

Few of the developmental disorders listed in the DSM are defined by measurable deviations, exceeding a threshold, from normal development. For example, neither the most commonly diagnosed developmental disorder, Attention Deficit Hyperactivity Disorder (ADHD), nor the most severe developmental disorders, the Pervasive Developmental Disorders, are defined by measurable deviations from normal development. Where disorders like Expressive Language Disorder are defined by measured deviations from the norm, the threshold deviation required to make the diagnosis appears to derive from convention rather than knowledge of the likelihood of the observed discrepancy (Aram, Morris, & Hall, 1992; Plante, 1998).

How Big is a Substantial Discrepancy?

All developmental disorders are defined as a deviant pattern of development, and in most cases, the deviant pattern is a discrepancy between development in one or more specific domains and a baseline. For example, learning (Reading, Mathematics, Written Expression) and two communication (Expressive Language, Mixed Receptive-Expressive Language) disorders are defined by a discrepancy in achievement. In these disorders, a child underachieves in a specific ability domain (e.g., reading, expressive language) compared with the child's own intelligence, not with the child's age peers (as in Mental Retardation). The assumption that specific abilities do not differ from general intelligence in normally developing children is implicit in how these disorders are defined.

Because specific developmental disorders are defined by unequal achievement, information on normal variability in achievement is required to identify children whose achievement is abnormally variable. However, DSM criteria do not require that an observed discrepancy exceed some threshold of normal variability. Instead, the most

rigorously stated criteria (Learning Disorders) require that achievement in the specific ability is substantially below what is expected and “substantially below is usually defined as a discrepancy of more than 2 standard deviations between achievement [in the specific ability] and IQ” (American Psychiatric Association, 1994, p. 46).

This criterion is unlikely to be meaningful in terms of the distribution of discrepancy scores in the population. The likely result of applying this criterion is that the children selected would have above average intelligence (Braden & Weiss, 1988). For any two abilities which are normally distributed, have the same mean and standard deviation, and are independent of each other (as intelligence and specific ability measures must partly be for discrepancies to be observed), the probability that one ability will exceed the other by 2 sd or more is .079. For any given child, the probability will vary as a function of the child’s position on the index or comparison dimension (i.e., IQ). An ‘average’ child with an IQ of 100 would obtain a specific ability score < 70 (-2 sd) in about 2% of cases, and an ‘underachieving’ child with an IQ of 70 would have a trivially small chance of obtaining a specific ability score less than 40 (2 sd lower than the IQ), but an ‘overachieving’ child with an IQ of 130 has a 50:50 chance of obtaining a specific ability score < 100 . The probability of observing an achievement discrepancy increases as a function of distance from the mean on either measured ability. If, by definition, the only discrepancies of interest are ones in which an intelligence score exceeds a specific ability score, then the probability of a discrepancy increases solely as a function of increasing scores on the intelligence measure (see Hawkins & Tulky, 2001; Iverson, Woodward, & Green, 2001).

The preceding probabilities are based on the assumption that the specific ability and IQ are normally distributed and independent. However, we know that intelligence and specific abilities are not independent and that the DSM assumes that all abilities covary strongly in normally developing children. In fact, when abilities are not independent, the

greater the correlation between them, the less likely it is that a discrepancy of one or two or three standard deviations will be observed. For example, the probability of observing a discrepancy of at least 2 sd drops from .079 to .024 as the correlation between two abilities increases from 0 to 0.5, decreases to .005 as the correlation increases to 0.7, and drops to .0001 as the correlation increases to 0.85 (cf. Braden & Weiss, 1988; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992).

Finally, the meaning of the discrepancy criterion varies across disorders because of differences in how the index ability, intelligence, is defined. In particular, the definition of Communication Disorders requires that achievement on the language measure is compared with non-verbal intelligence rather than with general intelligence. This redefinition of the index ability increases the chance of discovering a “substantial” discrepancy between the language and intelligence measures. Language measures are more strongly related to verbal than to nonverbal intelligence scales, which means that the probability of observing a discrepancy between language and non-verbal IQ is greater than the probability of observing a discrepancy between language and verbal IQ.

When a child’s language ability is less discrepant from his/her verbal than non-verbal intelligence, the child’s score on the language measure represents but one point on an ability gradient. If this child’s scores were ranked from lowest to highest, the difference between the child’s achievement on any two adjacently ranked scales need not be “substantial” even if the difference between the highest and lowest scores is substantial (Humphries & Bone, 1993). For example, a child may have a standardized expressive language score of 65, a receptive language score of 75, a verbal IQ of 85, and a performance IQ of 95. Performance IQ exceeds expressive language by 30 points or 2 sd even though adjacently ranked scores never differ by more than 10 points. In cases like this, we need to consider if it is appropriate to define the child as having a specific developmental disorder.

It is clear that when deviance is defined by an achievement discrepancy of 2 sd or more, no single criterion has been defined at all. What has been defined is a continuum of criteria that necessarily mean different things for different people, and different things for different pairs of abilities that are sometimes defined in different ways. The varying meaning of this criterion may explain, in part, why children who are diagnosed with one specific developmental disorder are so often also diagnosed with a second or third specific developmental disorder (e.g., Hazell, Carr, Lewin, Dewis, Heathcote, & Brucki, 1999; Piek, Pitcher, & Hay, 1999; Pierre, Nolan, Gadow, Sverd, & Sprafkin, 1999).

How Do We Know if Performance is Relatively Poor?

The 2 sd discrepancy criterion may be meaningless, but it does have the advantage of requiring the assessment of specific and general abilities with standardized, normed, individually-administered ability tests. These tests, which are used in the diagnosis of Mental Retardation, Communication Disorders, and Learning Disorders, provide good estimates of a child's ability relative to age peers and allow the clinician to conclude that a child is a poor reader or has limited receptive language ability even if these abilities are not "substantially" less than the child's intelligence. In the case of Learning Disorders, there is increasing evidence that low achievement within a domain, not whether or not one also has a low or high IQ, is what best defines whether a child has a learning problem (Gonzalez & Espinel, 1999; Stanovich & Stanovich, 1997; Sternberg & Grigorenko, 2002; Vellutino, Scanlon, & Lyon, 2000).

Whether or not low achievement better defines a problem than does discrepant achievement, in many cases (Attention Deficit Hyperactivity Disorder, Developmental Coordination Disorder, the Pervasive Developmental Disorders) a diagnostician is not required to use standardized, normed tests to estimate a child's ability. Unless such tests are used, it is not possible objectively to estimate a child's ability, or to compare the

performance of the child with his/her peers, or to gage the size of any achievement discrepancy.

In the case of Developmental Coordination Disorder, clinicians have a choice of several reliable and valid tests of motor coordination that could be used to measure a child's motor ability. But in the case of Attention Deficit Hyperactivity Disorder and the Pervasive Developmental Disorders, appropriate measures of the respective ability domains have not been developed, which precludes the objective estimation of ability, relative ability, and achievement discrepancy. For these disorders, we have no reliable way of knowing whether a child's achievement is poor.

Quality Versus Quantity in Defining Pervasive Developmental Disorders

The impairments in social reciprocity skills, communication skills, and behavioral repertoire in children with pervasive developmental disorders are defined as “qualitative” deviations from age and IQ-referenced norms (American Psychiatric Association, 1994, p. 65). The reference to qualitative impairments is consistent with traditional definitions of autism (Kanner, 1944) and, when defined in this way, autism appears to be qualitatively distinct from normality (Rutter & Sroufe, 2000). However, the concept of “autism spectrum disorders”—which includes Autistic Disorder, Pervasive Developmental Disorders Not Otherwise Specified (PDD-NOS), and Asperger's Disorder—implies variability across a dimension. Many children have autistic-like features which are not severe or wide-ranging enough to merit a diagnosis of autism. Asperger's Disorder (Asperger, 1968) refers to high-functioning cases where language milestones were normal but there is social impairment and stereotyped behavior and interests, and PDD-NOS is diagnosed when a child meets fewer than six of the criteria needed for a diagnosis of Autistic Disorder. Rather than focus on a qualitative distinction between normal and deviant, it may be more appropriate to focus on identifying the point along the severity

spectrum where deviant development is distinguished from normal development (Berney, 2000).

If we wished to regard the deviations that define pervasive developmental disorders in the same way that we regard the age-referenced deviations that define Mental Retardation, then two of the domains which define these disorders—communication skills and behavioral stereotypies—are related to the normal development of language and motor skills and can be distinguished from the normal by substantial underachievement in these ability domains compared with the person's age and general intelligence baseline. This is not to suggest that pragmatic language impairments or behavioral stereotypes are adequately described or explained only in terms of delayed acquisition of language (cf. Bishop, 2002) and motor coordination (cf. Gal, Dyck, & Passmore, 2002; Turner, 1997); rather, qualitative impairments are typically accompanied by quantitative impairments.

Unlike the communication or motor skills problems evident in PDDs, there is no domain comparable to “social reciprocity skills” among the specific developmental disorders or in the developmental literature more generally. The point of greatest convergence between how social reciprocity skills are defined in the DSM and how normal development is described is what Gillberg (1992) defined as empathic ability and Baron-Cohen (1994) described as mind-reading ability, but which is elsewhere described as theory of mind (Wimmer & Perner, 1983), emotion recognition ability (Eisenberg, Murphy, & Shepard, 1996), emotion understanding ability (Dunn, 1995), empathic accuracy (Ickes, 1993), and perspective-taking ability (Flavell, 1992). Empathic ability refers to the “ability to conceptualize other people's inner worlds and to reflect on their thoughts and feelings” (Gillberg, 1992, p. 835). An adequate conceptualization of another person's thoughts and feelings is a guide for engaging appropriately with that person, especially for “developing peer relationships,” and “sharing enjoyments, interests, or achievements with other people” (American Psychiatric Association, 1994, p. 70).

Having a defined set of normal ‘empathic abilities’ does not mean that these abilities, or discrepancies between these abilities and intelligence, can be measured. To date, measures of relevant developmental constructs have been used to assess normally developing children across only a restricted range of ages and ability. For example, theory of mind tasks were designed to assess change across a categorical divide that separates possessing, from not possessing, a theory of mind; children typically begin acquiring a theory of mind by age three years, and have acquired a theory of mind before age five years. It is also the case that these constructs have rarely been assessed concurrently, which means that even if they form one nomothetic network, their actual relations to each other remain unknown. Therefore, the normal developmental course of empathic ability, as opposed to distinct theory of mind or emotion recognition abilities, remains unknown and it is not possible to define deviations from the norm.

Capacity Versus Behavior in Defining ADHD

By definition, ADHD is a developmental disorder because a child’s symptoms of inattentiveness and/or hyperactivity/impulsivity are “inconsistent with developmental level” (American Psychiatric Association, 1994, pp. 83-84). However, the domains of ‘inattentiveness’ or ‘hyperactivity/impulsiveness’ have not been defined in the normal developmental literature and it is not clear that these domains can be effectively measured. The behaviors which define the inattentiveness or hyperactivity/impulsivity syndromes of ADHD are behaviors that allow a clinician to infer that a capacity to maintain attention or to inhibit behavior is impaired. The need to rely on such inferences is understandable—and necessary—if the ability domain has not been well-defined and if standard measures of the ability have not been constructed. In fact, although the development of the capacity for attentional control and behavioral inhibition has not been well documented in the developmental literature (Anderson, 2001; Dempster & Corkill,

1999), recent research is making substantial progress toward standardizing measures and reporting norms (e.g., Manly et al., 2001).

This progress is based on operationally defining the capacity for attentional control and behavioral inhibition with measures of executive functioning. Recent research on executive functioning deficits in children with ADHD (e.g., Barkley, 1997; Catellanos & Tannock, 2002; Oosterlaan, Logan, & Sergeant, 1998; Ozonoff, Pennington, & Rogers, 1991; Pennington & Ozonoff, 1996; Weirs, Gunning, & Sergeant, 1998) suggests that the behavioral syndromes which define ADHD may also reflect an impaired capacity to maintain attentional control and/or to inhibit behavior.

From our perspective, the observation that executive functions are impaired in children with ADHD means that measures of executive functioning may represent a way directly to assess the ability to maintain attention and/or to inhibit behavior. Direct assessment of ‘attentional control ability’ would eliminate the need to infer impaired attentional capacity from parent or teacher ratings, and would reduce the need to discriminate ‘apparent inattention’ from oppositional behavior (American Psychiatric Association, 1994, p. 84). Direct assessment of attentional control would also allow ADHD to be defined as a specific ability deficit consistent with how other developmental disorders are defined. Whether this potential can be actualized depends on knowledge of how executive functions develop and how their development covaries with the development of other abilities.

Research Aims

If developmental disorders are to be identified and distinguished from each other on the basis of patterns of unequal achievement, we need to know how discrepancy scores are distributed among normally developing children. Observing the relevant distributions requires concurrent assessment of the set of abilities which define the developmental disorders. In the case of empathic and attentional control abilities, such observations have

not been possible because tasks suitable for administration across childhood were not available. Suitable tasks appear now to be available, and their relations to other measures in a representative sample of normally developing children need to be observed.

This study has two main aims. The first aim is to estimate the distribution of discrepancy scores between intelligence and four specific abilities in a representative sample of children aged 3 to 12 years. The second aim is to assess the relationship between unequal ability and behavioral disturbance. If discrepancies between a specific ability and intelligence are a precondition for developmental disorders, then children who show substantial discrepancies (i.e., $> 2 \text{ sd}$) between the specific ability and intelligence should obtain significantly higher scores on measures of behavioral disturbance. Alternatively, if norm-referenced underachievement in a specific ability rather than unequal ability more validly defines disorders, then children who underachieve relative to age peers should obtain higher scores on measures of behavioral disturbance.

Method

Participants

Participants were 390 children aged 3 to 12 years, approximately balanced for age in years and for sex (see Table 1). Participants were recruited from 42 schools/preschools in the Perth metropolitan region. Schools were targeted for recruitment on the basis of their position on a state-wide index of average student achievement, i.e., because they represented the distribution of academic achievement within Western Australia.

Table 1 also reports the standard scores achieved by our sample on measures of verbal and performance IQ (for children aged 6 to 12 years, for whom relevant age norms were available), and on the neuromuscular development index (for all children; see Measures below). The results imply that our sample has ‘above-average’ intelligence and ‘below-average’ motor coordination compared to the standardization samples. If taken at face value, these results would imply that our sample is characterized by ‘discrepant

achievers.’ An alternative explanation is that American norms are inappropriate for use with Australian samples, both because of population differences and because of post-standardization increases in intelligence in all populations (Kamieniecki & Lynd-Stevenson, 2002). Previous research has shown that Perth-based samples routinely achieve ‘above-average’ scores on American-normed tests such as the Wechsler scales. Piek and Edwards (1997) observed that their sample of 171 children had a mean verbal IQ of 111, Piek, Dworcan, Barrett and Coleman (2000) observed that their sample of 72 children had a mean verbal IQ of 108, and Pitcher, Piek and Barrett (2002) observed that their control samples of 39 and 31 children had mean verbal IQs of 108 and 111, respectively. In our sample, the mean verbal IQ was 110, consistent with what is typically observed among Perth children. This pattern of results means that it was necessary to rescale measures of each ability construct to ensure that each standard ability score has the same mean and standard deviation in our sample.

Insert Table 1 about here

Once a school had agreed to participate, participants were recruited in one of two ways. Parents of children aged 7 to 12 years received letters seeking permission to enrol their child in ‘Project KIDS.’ Project KIDS is conducted through a child study center during school holidays, and involves intensive data collection, for one day per child, with small groups of children (see Procedure below). This method resulted in the recruitment of 234 children, including all participants aged 7 to 11 years, 9 6-year-olds (typically siblings of target-age children), and 20 12-year-olds. Parents of children aged 3 to 6 years or 12 years received letters seeking permission to assess their child at the preschool/school in which the child was enrolled. This method resulted in the recruitment of 156 children, including all children aged 3 to 5 years, and the balance of children aged 6 (n =

26) or 12 years ($n = 18$). Across both forms of recruitment, approximately one sixth of the total number of parents contacted consented to their child's participation in this study.

Measures

We administered measures of five ability constructs and a measure of nine dimensions of behavioral disturbance. The ability constructs were intelligence, language, motor coordination, empathic ability, and attentional control ability.

Intelligence. Intelligence was estimated with four subscales from the third edition of the Wechsler Intelligence Scale for Children (WISC; Wechsler, 1991)—Vocabulary, Information, Block Design, and Picture Completion. We used the WISC rather than the Wechsler Preschool and Primary Scale of Intelligence with the youngest participants to ensure that all participants was measured on the same scale. We judged that the value of using one scale outweighed the risk of floor effects from administering 'too difficult' items to the youngest children. The WISC subtests were selected because they represent the verbal and performance components of intelligence and because they provide a good estimate of fullscale IQ. Each test has excellent split-half and test-retest reliability, and both criterion and concurrent validity are well-established (Wechsler, 1991).

Language. Language ability was estimated with four subscales from the third edition of the Clinical Evaluation of Language Fundamentals (CELF; Semel, Wiig, & Secord, 1995)—Concepts and Directions, Word Classes, Recalling Sentences, and Formulated Sentences. We selected the CELF partly because it has been standardized across a wide range of ages, but like the WISC, not for the youngest children (aged 3 to 5 years) in our study. Again, we judged that the value of using one scale outweighed the risk of floor effects from administering 'too difficult' items to the youngest children. Specific scales were selected because they are the only CELF scales which are administered to all children and because they sample receptive (Concepts and Directions, Word Classes) and expressive (Recalling Sentences, Formulated Sentences) language.

These subscales have acceptable internal consistency (alphas from .54 to .91), test-retest reliability (.69 to .87), and concurrent validity (Semel et al., 1995).

Motor Coordination. Motor coordination was assessed with the McCarron Assessment of Neuromuscular Development (MAND; McCarron, 1982, 1997). The MAND comprises 10 tasks, of which five assess fine motor skills (e.g., putting beads in a box) and five assess gross motor skills (e.g., heel to toe walking). The 10 MAND tasks have acceptable test-retest reliability (.67 to .98), criterion validity (prediction of work performance), and concurrent validity (McCarron, 1982, 1997).

Empathic Ability. Empathic ability was estimated with a combination of four first order theory of mind tasks, an advanced theory of mind task, and six subscales from the Emotion Recognition Scales (Dyck, Ferguson, & Shochet, 2001). These tasks represent developmental constructs implicated in autism spectrum and other developmental disorders and which are cognate with the social reciprocity skills named in the DSM.

First order theory of mind tasks are false belief tasks commonly used to assess differences between children with/without some disorder, and included the “Sally Ann” (Baron-Cohen, Leslie & Frith, 1985), “Smarties” (Perner, Frith, Leslie & Leekam, 1989; Wimmer & Perner, 1983), “John and Mary icecream story” (Perner & Wimmer, 1985), and “Ella the Elephant” tasks (Harris, Johnson, Hutton, Andrews & Cooke, 1989). In each task, a child is asked whether a protagonist will act consistently with the protagonist’s beliefs, known to be false, or consistently with what the test-taker knows to be the true state of the world. Responses which indicate action consistent with the protagonist’s false beliefs are scored correct (1) and indicate that the test-taker has acquired a theory of mind. We treated these tasks as separate items on a 4-point theory of mind scale.

The Strange Stories Test (Happe, 1994) assesses the ability to provide context-appropriate mental state explanations for non-literal (irony, sarcasm, lies) statements. Deficits in this ability may account for interpersonal deficits in children with autism who

pass so-called first and second order theory of mind tests. Happe (1994) has used this test with autistic, mentally retarded, and normal children and adults; results indicate that children with autism perform less well on this task than mentally retarded and normal children (cf. Dyck et al., 2001). The test consists of 12 stories (one for each form of non-literal statement), each accompanied by a picture. Subjects indicate whether a statement made by the protagonist is true or false (to establish that the story was understood), and then explain why the statement was made. In this study, responses were scored correct if an explanation was both adequate and relied on references to mental states.

The Emotion Recognition Scales (ERS) include measures of emotion recognition (Fluid Emotions Test, Vocal Cues Test) and emotion understanding ability (Emotion Vocabulary Test, Comprehension Test, Unexpected Outcomes Test).

The Fluid Emotions Test (FET; Dyck, Farrugia, Shochet, & Holmes-Brown, 2002) measures the ability to recognize static and changed/changing facial expressions of emotion. This is a computer-presented test and items are drawn from Matsumoto and Ekman's (1995) color slides of adults expressing one of seven emotions (anger, contempt, disgust, fear, happiness, sadness, surprise) or a neutral expression. Each item consists of two head and shoulders pictures of a Japanese or Caucasian male or female expressing one of the 7 emotions or a neutral expression. The test-taker is asked what emotion is being expressed in the first picture. After responding, the image is gradually (over 4 seconds) transformed to another person expressing a different emotion. Subjects identify, as quickly as they can, the second emotion. Speed of response is measured with a stop-watch. Two FET scales were used: initial accuracy (ACC-1; initial emotions correct); and speed given accuracy (SGA). The SGA is based on the speed of accurate post-morph responses. Response latencies greater than 12 seconds are scored 0 whether the response is accurate or not. Latencies of 9 – 12 seconds are scored 1, and each subsequent 1 second decrease in latency results in an incremental score of 1. Latencies less than 4 seconds are

scored 7. These scales are internally consistent (ACC-1, $\alpha = .90$; SGA, $\alpha = .94$), have good concurrent validity (Dyck et al., 2002), and are useful in identifying empathic ability deficits in children with autism spectrum and non-spectrum disorders (Dyck et al., 2001).

The Vocal Cues Test (VCT; Dyck et al., 2002) measures the ability to recognize vocal intonations specific to different emotions. We used VCT “Unreal” scale, which consists of 43 items in which emotions are expressed using non-semantic content: numerals, letters, nonsense syllables. The emotions sampled are identical to those in the FET. Items are approximately balanced for gender of the speaker and for emotion category. Responses are scored correct (1) or incorrect (0). This scale is internally consistent ($\alpha = .93$) and has good concurrent validity (Dyck et al., 2002).

The Emotion Vocabulary Test (EVT) measures the ability to define emotion words (e.g., what does the word “angry” mean?). The response format of the EVT is open-ended and initial responses may be queried in order to resolve ambiguities in the initial response. The EVT is internally consistent ($\alpha = .82 - .89$), moderately related to other ERS, and strongly related to other measures of vocabulary (Dyck et al., 2001, 2002).

The Comprehension Test (CT) measures the ability to understand the emotional consequences of exposure to an emotion-eliciting context (e.g., Susan is given a new bicycle for her birthday. What will Susan feel?). CT items sample the 7 emotions in the FET, ‘social variants’ of emotions (e.g., pride, embarrassment, shame, pity) and variations in the intensity of emotions (e.g., terror versus fear). Emotion causes include ‘material causes’ (e.g., loss/gain of an object), ‘social causes’ (e.g., interpersonal rejection), and ‘intrapsychic causes’ (e.g., failure to achieve one’s goals). The CT has acceptable reliability ($\alpha = .64 - .79$) and is moderately related to other ERS and to measures of intelligence (Dyck et al., 2001, 2002).

The Unexpected Outcomes Test (UOT) measures the ability to apply reasoning skills and knowledge of the causes of emotions to explaining apparent incongruities

between an emotion-eliciting context and the emotion elicited by the context. UOT items provide information about a situation that is likely to cause an emotional response in a protagonist (e.g., “John likes a girl called Susan, and he wants her to go to the movies with him. When he asks her, she says yes”). Items then indicate what emotion has been experienced (e.g., “On their way to the movies, he is very angry”). In each case, the emotion differs from what is usually expected to occur in the situation. The test-taker is asked to explain the apparent incongruity. The UOT has adequate reliability ($\alpha = .73 - .81$) and is moderately to strongly related to other ERS and to measures of intelligence (Dyck et al., 2001, 2002).

Attentional Control/Behavioral Inhibition. The capacity to control attention and inhibit responses was assessed with a set of four computer-administered executive functioning tasks. The Goal Neglect Task (Duncan, Emslie, & Williams, 1996), Go/No-Go Task (Fox, Michie, Wynne, & Maybery, 2000; Shue & Duncan, 1992), Compatibility Switch Task (Anderson, 1988), and Trailmaking/Updating Memory Task (Rabbitt, 1997) were selected because of their known relations to ADHD (Pennington & Ozonoff, 1996) and/or their clear face validity.

The Goal Neglect Task measures the ability to formulate and respond to goal-directed plans (Duncan et al., 1996). It requires a test-taker to disregard a task requirement which has been understood and remembered in order to achieve another goal (Duncan et al., 1996). This is a typical executive functioning task involving novel behavior, concurrent tasks, and the absence of verbal feedback (Duncan et al., 1996; Rabbitt, 1997).

In this task, letters and numbers are presented to the left or right of a fixation point. Test-takers are asked to read the stimuli on either the left or right of the screen, and then switch to the opposite side if a + sign is presented or stay on the same side if a – sign is presented. The task has 12 trials (6 ‘switch’ and 6 ‘stay’ trials) in which presentation of 10 sets of stimuli is followed by the switch or stay cue, and then 3 additional sets of

stimuli. A trial is passed if, before and after a cue, more letters are called from the correct than the incorrect side. Performance on the Goal Neglect Task is positively related to age and IQ (Duncan et al., 1996).

We modified the Go/No-Go task of Shue and Douglas (1992) to assess simple motor inhibition. In this task, letters are designated either as 'go' (respond) or 'no-go' (do not respond) stimuli, and are presented at one second intervals. When a go stimulus is presented, the test-taker is required to press a response key as quickly as possible, and when a no-go stimulus is presented, the test-taker is required not to press a response key. Test-takers complete two trials of the task. Each trial consists of 120 stimuli, of which 60 are 'go' and 60 are 'no-go' stimuli. Responses to the 'no-go' stimulus are scored as commission errors, and a failure to respond to the 'go' stimulus are scored as omission errors. This test consistently discriminates between ADHD and non-ADHD children (Pennington & Ozonoff, 1996; Shue & Douglas, 1992).

The Compatibility Switch Task is a line-length discrimination task designed to assess visual inspection time (Anderson, 1988). The task requires test-takers to press, as quickly as possible, a blue key if two lines (antennae on 'aliens') are the same length, and to press a red key if they differ in length. The task comprises 120 stimulus presentations.

The task was modified to assess simple motor inhibition by adding a second trial of 120 stimulus presentations in which the required responses are reversed (i.e., blue for lines of differing length, red for lines of equal length). This second trial requires test-takers to override previous learning. This task yields two scores: response time and number of correct responses.

The Trailmaking/Memory Updating Task is a simplification of a more complex task (Rabbitt, 1997) and is designed to assess working memory and behavioral inhibition. In this task, the first four letters of the alphabet are designated as the 'target set,' and within this target set, the actual target changes with successive stimulus presentations

(i.e., from A to B to C to D to A). Test-takers are required to discriminate whether a letter, presented on screen, is part of the target set, and if it is, whether the letter is the current target. Test-takers complete two trials, and each trial consists of 120 stimulus presentations, of which 20 presentations are the target stimulus. For each presentation, test-takers press a blue key if the stimulus is the target stimulus and press a red key otherwise.

Behavioral Disturbance. We estimated the presence and severity of behavioral and emotional disturbances with the age-appropriate parent forms of the Child Behavior Checklist (CBCL; Achenbach, 1988, 1991). The CBCL lists 113 symptoms that parents rate as ‘not at all true’ (0), ‘sometimes true’ (1), or ‘mostly true’ (2) of their child. Items are combined to form eight primary scales (six primary scales for 2-3 year olds). The common subscales include ‘aggressive behavior,’ ‘delinquent’ or ‘destructive behavior,’ ‘anxious/depressed,’ ‘somatic complaints,’ ‘withdrawn.’ The ‘social problems,’ ‘thought problems,’ and ‘attention problems’ scales are specific to the form for 4 to 18-year-olds, and the ‘sleep problems’ scale is specific to the 2 to 3-year-olds. Both forms also include a list of ‘other problems,’ which we treated as a non-specific problems scale.

In the context of a study assessing the validity of DSM criteria (cf. Hartman et al., 2001), it is important to note that our use of the CBCL does not imply that we regard it as a ‘superior’ or even as an appropriate way of defining developmental disorders. Rather, while we acknowledge that the CBCL has excellent psychometric characteristics, is frequently used, and has well-established norms (Sawyer, Arney, Baghurst et al., 2001), its scales do not correspond to clinical syndromes and the scale was not designed as a substitute for clinical diagnosis (Achenbach, 1988, 1991; Hartman et al., 1999).

Procedure

Consenting parents were sent a CBCL parent form for each participating child and were asked to return the forms in reply-paid envelopes. Procedure then varied depending on whether the child was or was not assessed as part of Project KIDS.

For children participating in Project KIDS, groups of up to 12 children were scheduled for a full day (8:45 a.m. to 4:30 p.m.) of activities. Upon arrival at the child study center, children participated in a “getting to know you” activity. Testing was then conducted in three 90-minute sessions, each of which was divided into three 30-minute testing blocks. The first test session was followed by a 30-minute recess, and the second by a 60-minute lunch break. Testing was administered by a team of researchers. During breaks, children were provided with coloring books, pencil puzzles, and age-appropriate movies; they were also given access to an outdoor playground.

The order of test administration was not uniform. Rather, each child had his/her own schedule. Adherence to the test schedule was essential to the smooth running of the program; if scheduled activities could not be completed, they were deferred to the end of the day where one hour of unallocated time was available to administer deferred tasks. Testing was usually completed within 4.5 hours, but sometimes required up to 5.5 hours. All tests except the attentional control tasks were individually administered according to the instructions in the relevant manuals. Children were given individual instructions for the attentional control tasks, but performed the tasks in a room with up to four children.

For children not in Project KIDS, testing was done at the school of recruitment. For these children, testing was less rigidly scheduled in order to accommodate the shorter attention span of younger children and to minimize disruption to school activities. Because of test discontinuation rules, younger children usually completed fewer test items, which reduced the total time required. In most cases, testing was completed in a single day; otherwise, testing was completed within two days.

Finally, we discovered that our attentional control tasks were too difficult for the youngest children who could not understand the requirements of the task. For this reason, we discontinued administering these tasks to children aged 3 to 5 years.

Derived Variables

For each ability construct, we constructed two composite scores: an absolute ability score and an age-referenced ability score or specific ability quotient.

Absolute ability scores were constructed by transforming raw scores for each variable into a z-score (based on the means/standard deviations for the whole sample), and then averaging the z-scores of the variables that defined the construct. Absolute intelligence was defined by adding each child's z-scores on the four Wechsler scales and dividing the sum by four. Absolute motor ability was defined by adding each child's z-scores on the McCarron scales and dividing by ten. Absolute empathic ability was defined by adding each child's z-scores on the theory of mind test, Strange Stories Test, and six Emotion Recognition/Understanding scales and dividing by eight. Absolute language ability was defined by adding each child's z-scores on the four CELF scales and dividing by four. Absolute attentional control ability was calculated by: (a) averaging the standardized 'number correct' scores from the second trial of the compatibility switch task, and the two trials of the trailmaking task; (b) averaging the standardized 'response time' scores from these same tasks and then inverting the resultant scale; (c) averaging the standardized response time standard deviations from these same tasks and then inverting the resultant scale; (d) inverting the standardized score for the second trial of the go/no-go task; and then averaging the results of steps 'a' to 'd' with the standardized score of the goal neglect task.

Age-referenced specific ability quotients were calculated as follows. For each absolute ability score, means and standard deviations for each age group were calculated. For each child and for each ability, the appropriate mean was subtracted from the child's

absolute ability score and the difference was divided by the appropriate standard deviation. Each of these age-normalized ability scores was then multiplied by 15, and 100 was added to the product. Thus, each specific ability quotient has a mean of 100 and a standard deviation of 15. This renorming procedure ensured that any observed ability discrepancies could not be attributed to systematic differences (e.g., higher on an IQ measure, lower on a motor coordination measure) between our sample and the sample used to provide the general test norms. Age-referenced discrepancy scores were calculated by subtracting each other specific ability quotient from age-referenced intelligence.

Results

How Independent Are Specific and General Abilities?

We assessed the independence of the five ability constructs by calculating correlations between specific ability quotients. Table 2 shows that all correlations are significantly greater than zero except in the case of language and attentional control abilities. Differences in the size of correlations between different pairs of abilities indicate that the probability of observing substantial discrepancies between abilities will vary markedly across ability pairs. In particular, substantial (greater than 2 sd) discrepancies between intelligence and language abilities should be seldom evident while substantial discrepancies between intelligence and motor coordination or between intelligence and attentional control may approximate what is expected of independent variables.

 Insert Table 2 about here

The observed correlations between ability variables are reflected in the distributions of discrepancy scores for pairs of abilities across the sample. Table 2 shows that for the more highly correlated abilities (intelligence/language, intelligence/empathy), the range of discrepancy scores is narrower and the standard deviation is smaller. For the

more strongly correlated abilities, a discrepancy of 30 IQ-equivalent points is less likely because it is a larger (2 sd for intelligence/empathy, or 2.5 sd for intelligence/language) deviation from the mean than in the case of less strongly correlated variables (about 1.4 sd for intelligence/motor ability and 1.5 sd for intelligence/attention).

We calculated the actual frequency of substantial discrepancies between intelligence and each of the specific abilities. We observed that intelligence exceeds language ability by 30 or more points in 4 cases (1%), motor coordination ability in 28 cases (7.2%), empathic ability in 10 cases (2.5%), and attentional control ability in 18 cases (7.3% of 245). In other words, the distribution of discrepancy scores is consistent with what is expected of pairs of correlated variables, and varies according to the magnitude of the correlation between the paired variables.

Insert Table 3 about here

We also expected that applying a discrepancy criterion of 2 or more standard deviations would result in the selection of children with above average ability on the index dimension (intelligence). We tested this hypothesis by calculating the average intelligence and specific ability scores of children with discrepant abilities, and then comparing them with children whose abilities are not discrepant (do not differ by more than 1/3 sd). Table 3 shows that for each pair of discrepant abilities, the average IQ of children who show discrepant ability is approximately 1 sd above the mean and is significantly higher than that of comparison children. Each pair of discrepant abilities also define extremes between which all other abilities fall, i.e., a gradient across which abilities vary, not a discrepancy between one ability and all other abilities.

Are Discrepancies in Ability Related to Behavioral Disturbance?

We used t-tests to assess whether children with ability discrepancies have more symptoms of behavioral disturbance. For each kind of discrepancy, we compared children with the discrepancy with children whose specific ability quotients did not differ by more than 5 points. The dependent variables were the eight CBCL primary scales and the Other Problems scale. With nominal alpha set at .05, one to two spuriously ‘significant’ results were expected; no significant group difference was observed. The results were that children whose intelligence exceeded their language ability (n = 4), or their empathic ability (n = 10), or their attentional control ability (n = 18), or their motor coordination ability (n = 28) did not differ from comparison children (for language ability, n = 151; for empathic ability, n = 126; for attentional control ability, n = 47; for motor coordination ability, n = 101) on any CBCL scale.

Are Low Scores on Specific Ability Measures Related to Behavioral Disturbance?

Table 3 shows that when children are selected on the basis of discrepancies between intelligence and a specific ability, scores on the specific ability are invariably significantly lower than are achieved by the sample as a whole. We questioned whether children selected on the basis of low specific ability quotients (< 85), rather than ability discrepancies, would differ from comparison children (all other children) in parent-reported symptoms of behavioral disturbance and in achievement on the range of ability measures. We tested these possibilities with one-way analyses of variance. With nominal alpha set at .05, approximately two spuriously ‘significant’ results could be expected; 17 significant group differences were observed. Groups of children with specific ability quotients less than 85 on any ability measure obtained significantly higher scores on between one (in the case of motor coordination) and five (in the case of empathic ability) of the CBCL scales (see Table 4). Elevated scores were most common on the Social, Other, Somatic, and Thought problem scales.

Insert Table 4 and Table 5 about here

We next assessed whether children selected based on low scores on one ability measure also achieve low scores on other ability measures. With nominal alpha set at .05, one spuriously ‘significant’ result was expected; 19 significant group differences were observed. The results, summarized in Table 5, show that children with low ability scores on any one specific ability also obtain significantly lower scores on other ability measures. Children with low scores on attention also have low scores on motor coordination, children with low scores on language, motor coordination or empathic ability also have low scores on three other ability measures, and children with low scores on intelligence also achieve low scores on each other ability measure.

This pattern of results suggests that children who would be classified as having a ‘specific impairment’ on the basis of low scores on one ability dimension would also be classified as having ‘comorbid specific impairments’ on the basis of also having low scores on other ability dimensions. We observed that of the children who would be classified as having an intelligence deficit, about 25% would also be classified as having an attentional control deficit, about 40% an empathic ability deficit, 25% a motor coordination deficit, and 50% a language deficit. Of children who would be classified as having a language deficit, 20% would also be classified as having an attentional control deficit, 55% an empathic ability deficit, and 25% a motor coordination deficit. These results indicate that deficits exceeding one standard deviation co-occur at rates well beyond chance (probability is less than .05, Fisher’s Exact Test, in each case).

What are the Ability Scores of Children with Clinically Significant Behavioral Disturbances?

The children who were given significantly higher ratings of behavioral disturbance in the preceding section would not necessarily receive scores high enough to be in the

‘clinical range.’ We assessed how children classified as having a clinically significant disturbance based on their CBCL scores achieve on ability measures. Children who were given scores in the clinical range on each CBCL scale (using age and sex norms for each scale) were compared with children who obtained a score of ‘0’ on the relevant clinical scale. With nominal alpha set at .05, two spuriously ‘significant’ results could be expected; 13 significant differences were observed. As Table 6 shows, children with clinically significant disturbances on the Attention, Delinquency, Social, Somatic, and Other Problems scales achieved lower scores on one or more ability measures.

Insert Table 6 about here

In contrast to this pattern of results, in no case did a group of children with a clinically significant behavioral disturbance differ from comparison children in any discrepancy score (Intelligence/Language, Intelligence/Motor Coordination, Intelligence/Empathy, Intelligence/Attentional Control).

Discussion

This study had two main aims. The first aim was to estimate the distribution of discrepancy scores across five main ability dimensions in a representative sample of children aged 3 to 12 years. We observed that the distribution of discrepancy scores varies predictably as a function of the correlation between two abilities and as a function of position on the index ability dimension. The second aim was to assess whether achievement discrepancies are related to behavioral disturbances. We observed that discrepancies between abilities were not related to behavioral disturbances in normally developing children. Rather, we observed that underachievement relative to age peers on any ability was associated with behavioral disturbance, and children with clinically

significant behavioral disturbances typically underachieved relative to age peers on more than one ability measure.

The Distribution of Unequal Achievement and Behavioral Disturbance

The DSM defines specific and pervasive developmental disorders by discrepancies in achievement. By defining developmental disorders in this way, the DSM assumes that among normally developing children, achievement in different abilities ought to be uniform. It makes this assumption without regard to the probability of observing discrepancies of a given size between correlated abilities with given distributions (e.g., normal or skewed) and without regard to the role that position on one ability distribution plays in determining the probability that a discrepancy of a given size can be observed. Our results are consistent with the small number of studies assessing discrepancies in achievement (Braden & Weiss, 1988; Hawkins & Tulky, 2001; Iverson et al., 2001; Wilkinson, 1993) and show that the assumption of uniform achievement is false. The probability of observing unequal achievement is entirely consistent with the expected distribution of discrepancy scores given the parameters of the different ability distributions and the correlations between them.

In saying that the distribution of discrepancy scores is consistent with expectation, we are also saying that discrepancy scores cannot be valid markers of developmental disorders. When a distribution is entirely predictable from some set of parameters, there is no 'information' contained by that distribution which is unique to that distribution, that is, information not already contained in the previously known parameters. In the case of developmental disorders, this means two things. First, because the probability of observing a discrepancy of a given magnitude varies according to the parameters of the two distributions from which the discrepancy derives, the meaning of any given discrepancy score will not be consistent across ability pairs. In our sample, a 30-point discrepancy, ostensibly a 2 sd discrepancy between two specific ability quotients,

represented a range of deviations from 1.5 sd (intelligence/attention) to 2.5 sd (intelligence/language) above the average discrepancy score. Second, the presence of a discrepancy or the magnitude of a discrepancy can provide no information about any pathological process that is not already contained in what is known about achievement on individual abilities and the general relationship between the individual abilities. Because we know that the probability of observing a substantial discrepancy between two abilities with a known correlation is mainly a function of superior performance on the index ability, the only source of information on any individual's impaired performance must come from what is known about the person's necessarily inferior performance on the non-index ability. If discrepancy scores provided information about a disordered process additional to what was provided by underachievement relative to age peers on some ability, the disordered process would cause variability in discrepancy scores that could not be accounted for by the usual parameters. Because the observed distribution of discrepancy scores does not deviate from expected values, there is no evidence of any process that independently affects discrepancy scores.

Our findings on the relationship between discrepancy scores and parent ratings of behavioral disturbance are consistent with this conclusion. Children with index ability (intelligence) scores significantly above the sample mean and specific ability scores significantly below the mean did not differ from comparison children in rated symptoms of behavioral disturbance.

Our results do not exclude the possibility that children with some developmental disorders may have characteristic achievement discrepancies or uneven cognitive profiles of the kind associated with autism (Dennis, Lockyer, Lazenby et al., 1999; Happe, 1994b). Rather, because discrepancy scores do not predict dysfunction across the population, the use of discrepancy scores to define the disorder is likely to result in the overidentification of high achievers and the failure to identify children who are, in fact,

impaired, as has been found in the case of learning and communication disorders (Braden & Weiss, 1988; Plante, 1998).

Underachievement versus Unequal Achievement

In marked contrast to the null results obtained with discrepancy scores is the consistently significant association between underachievement relative to age peers and behavioral disturbance. Low scores on each ability dimension are associated with significant elevations on between one and five of the nine CBCL scales. Groups of children who underachieve relative to age peers on any ability measure, not groups of children whose abilities are unequal, show symptoms of behavioral disturbance.

The pattern of associations between categories of ability and categories of behavioral disturbance is mainly inconsistent with the idea that specific ability deficits will be associated with specific behavioral disturbances. All ability deficits are associated with 'somatic' problems, and most ability deficits are associated with 'social' and 'thought' problems. The simplest explanation for the lack of specific associations between ability and behavioral disturbance categories is that ability categories are not independent of each other and behavioral disturbance categories are not independent of each other (cf. Hartman et al., 1999, 2001). Just as we observed that children with low scores on intelligence or language are likely also to have low scores on motor coordination, empathic ability, and attentional control, children who have elevated scores on one CBCL scale are likely also to have elevated scores on other CBCL scales. For example, of the 12 children who exceed the clinical cut-off for 'Anxious/Depressed,' 3 also have attention problems, 4 are delinquent, 8 have other problems, 5 have social problems, 7 have somatic problems, 5 have thought problems, 4 are withdrawn, and 5 are aggressive. Only one of the 12 children with excessive anxiety/depression had this as their only clinically significant problem. If ability categories or behavioral disturbance categories are not themselves independent, then specific associations between ability and behavioral

disturbance categories are unlikely to be observed (cf. Prior, Smart, Sanson, & Oberklaid, 1999). What is important is that if CBCL scores can be taken as an index of adaptive functioning (Pearson & Lachar, 1994), we now know that low scores on all ability measures are associated with impaired adaptive functioning.

An association between underachievement relative to age peers and impaired adaptive functioning is important because it suggests an objective way of assessing whether a child's underachievement should be regarded as 'clinically significant.' The analogy is with the diagnosis of Mental Retardation, where a child must not only underachieve on intelligence tests, but also demonstrate an impaired capacity for adaptive functioning on standardized tests of adaptive functioning. The recommended CBCL cut-offs that we used represent the 98th percentile on each scale, that is, 2 sd above the mean on each index of behavioral disturbance (Achenbach, 1988, 1991).

Our results do not exclude the possibility that underachievement relative to age peers on one or more ability dimensions may be associated with a specific pattern of behavioral disturbance. In order to assess this possibility, it will be necessary to use measures of behavioral disturbance that are less strongly intercorrelated than are the CBCL scales. Alternatively, research could focus on the processes that may cause children with different ability profiles to develop similar patterns of behavioral disturbance.

Normal and Abnormal Variability in Ability

There is no evidence from this study to suggest that unequal achievement across abilities is a marker of developmental disorder. This result ought not to be surprising because it is entirely consistent with what is observed in clinical samples. Despite the fact that developmental disorders are ostensibly defined by discrepancies in achievement, clinical research consistently indicates that children with defined underachievement relative to age peers in one domain also underachieve in other domains. Whether one

samples on the basis of language impairments (Beitchman, Douglas, Wilson, Johnson, Young, et al., 1999; Bishop, 2000; Courtright & Courtright, 1983; Larson & McKinley, 1995; Wiig & Harris, 1974), empathic ability impairments (Baron-Cohen, Scahill, Izaguirre, Hornsey, & Robertson, 1999; Cook & Leventhal, 1995; Ghaziuddin, Tsai, & Ghaziuddin, 1992), attentional control impairments (Hazell et al., 1999; Piek et al., 1999; Pierre et al., 1999), motor coordination impairments (Kaplan et al., 1998; Sheppard, Bradshaw, Pucell, & Pantelis, 1999), or intelligence impairments (Kent, Evans, Paul, & Sharp, 1999; Verhoeven & Tuinier, 1997), impairments in other ability domains will typically be observed. Patterns of variability across ability domains do not appear to differ in normal and clinical samples. What differs is the magnitude of underachievement across a set of abilities.

Experimental Measures

Research on intelligence, language, and motor coordination has a long history, and in each of these domains there now exists a set of standard ability tests from which a researcher or clinician may choose. Because comparable tests do not exist in the empathic and attentional control ability domains, we have had to select sets of tasks to sample the range of abilities that may be thought to comprise each of these domains. Our results support the combination of measures that we selected to measure empathic and attentional control abilities.

Empathic ability scores were observed to have among the strongest association of any ability with behavioral disturbance. Children who achieved low scores (but not discrepant scores) on empathic ability measures obtained significantly higher scores than comparison children on five CBCL scales, more than any other ability domain and including the Aggressive Behavior scale, which was associated with low scores on no other ability (see Table 4). Similarly, children who exceeded clinical cut-offs on five CBCL scales achieved significantly lower scores on empathic ability measures than did

comparison children, more than any other ability measure. In the context of ability constructs that are important to understanding developmental disorders, these results provide strong support for our operational definition of the empathic ability construct.

The most obvious and most important problem in our operational definition of attentional control ability is the fact that our measures were not suitable for administration to children aged 3 to 5 years. Because we were not able to measure increases in attentional control ability over this age range, the range of achievement of attentional control ability that we did assess was restricted compared with other abilities. Both this restriction of range effect and a loss of statistical power for all tests using attentional control measures mean that our estimates of effects associated with this variable are less reliable than for other ability measures.

Despite this problem, children who achieved low scores on attentional control ability measures obtained significantly higher scores than comparison children on four CBCL scales, including the Attention Problems scale, which was associated with low scores on no other ability (see Table 4). Similarly, children who exceeded clinical cut-offs on two CBCL scales achieved significantly lower scores on attentional control ability measures than did comparison children. Associations between attentional control ability and behavioral disturbance are no less strong than they are for other ability measures, and the specific association between underachievement on attentional control measures and parent-rated attention problems indicates that our approach to measuring this domain merits further work.

Implications and Conclusions

The major assumption on which the DSM relies in defining developmental disorders, that discrepancies in achievement are a necessary condition for defining a disorder, is incorrect. Discrepancies in achievement between two or more abilities are a function of the correlation between the abilities, the shape of the respective distributions,

and position on the index distribution. Discrepancies in achievement are unrelated to behavioral disturbances, and contain no information about pathological processes.

Underachievement on one or more abilities, not unequal achievement across abilities, is invariably associated with behavioral disturbance. These results imply that we need to redefine developmental disorders along the same pattern as Mental Retardation is presently defined, that is, by: (a) underachievements, (b) of defined magnitude, (c) using standardized measures, with (d) known relations to normal development, and (e) concurrent deficits on standardized measures of impaired function or behavioral disturbance. If we can subsequently define how underachievement on one ability is functionally related to underachievement on one or more other abilities, then it may become possible to develop diagnostic categories and criteria that describe meaningful inequalities among abilities.

References

- Achenbach, T. (1988). Manual for the Child Behaviour Checklist for ages 2-3 and 1988 Profile. Burlington, VT: University of Vermont, Department of Psychiatry
- Achenbach, T. (1991). Manual for the Child Behaviour Checklist: 4-18 and 1991 Profile. Burlington, VT: University of Vermont, Department of Psychiatry.
- American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: American Psychiatric Association.
- Anderson, M. (1988). Inspection time, information processing and the development of intelligence. British Journal of Developmental Psychology, 6, 43-57.
- Anderson, M. (2001). Annotation: Conceptions of intelligence. Journal of Child Psychology and Psychiatry, 42, 287-298.
- Applegate, B., Lahey, B., Hart, E., Biederman, J., Hynd, G., Barkley, R., Ollendick, T., Frick, P., Greenhill, L., McBurnett, K., Newcorn, J., Kerdyk, L., Garfinkel, B., Waldman, I., & Shaffer, D. (1997). Validity of the age-of-onset criterion for ADHD: A report from the DSM-IV field trials. Journal of the American Academy for Child and Adolescent Psychiatry, 36, 1211-1221.
- Aram, D., Morris, R., & Hall, N. (1992). The validity of discrepancy criteria for identifying children with developmental language disorders. Journal of Learning Disabilities, 25, 549-554.
- Asperger, H. (1968). On the differential diagnosis of early infantile autism. Acta Paedopsychiatrica: International Journal of Child and Adolescent Psychiatry, 35, 136-145.
- Barkley, R. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. Psychological Bulletin, 121, 65-94.

- Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mind-reading. Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 13, 513-552.
- Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a “theory of mind?” Cognition, 21, 37-46
- Baron-Cohen, S., Scahill, V., Izaguirre, J., Hornsey, H., & Robertson, M. (1999). The prevalence of Gilles de la Tourette syndrome in children and adolescents with autism: A large scale study. Psychological Medicine, 29, 1151-1159.
- Beitchman, J., Douglas, L., Wilson, B., Johnson, C., Young, A., Atkinson, L., Escobar, M., & Taback, N. (1999). Adolescent substance use disorders: Findings from a 14-year follow-up of speech/language-impaired and control children. Journal of Clinical Child Psychology, 28, 312-321.
- Berney, T. P. (2000). Autism—an evolving concept. British Journal of Psychiatry, 176, 20-25.
- Bishop, D. (2000). Pragmatic language impairment: A correlate of SLI, a distinct subgroup, or part of the autistic continuum? In D. Bishop & L. Leonard (Eds.), Speech and language impairments in children (pp. 99-114). East Sussex, UK: Psychology Press.
- Braden, J. P. & Weiss, L. (1988). Effects of simple-difference versus regression discrepancy methods: An empirical study. Journal of School Psychology, 26, 133-142.
- Cantwell, D. P. (1996). Classification of child and adolescent psychopathology. Journal of Child Psychology & Psychiatry, 37, 3-12.
- Castellanos, F. X., & Tannock, R. (2002). Neuroscience of Attention Deficit/Hyperactivity Disorder: The search for endophenotypes. Nature Reviews: Neuroscience, 3, 617-628.

- Cohen, D. J., Paul, R., Volkmar, F. R. (1986). Issues in the classification of pervasive and other developmental disorders: Toward DSM-IV. Journal of the American Academy of Child & Adolescent Psychiatry, 25, 213-220.
- Conners, C. K., Sitarenios, G., Parker, J., & Epstein, J. (1999). The revised Conners' Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity. Journal of Abnormal Child Psychology, 26, 257-268.
- Cook, E., & Leventhal, B. (1995). Autistic disorder and other pervasive developmental disorders. Child and Adolescent Psychiatric Clinics of North America, 4, 381-399.
- Courtright, J., & Courtright, I. (1983). The perception of nonverbal vocal cues of emotional meaning by language-disordered and normal children. Journal of Speech & Hearing Research, 26, 412-417.
- Dempster, F., & Corkill, A. (1999). Neo-interference research and the development of intelligence. In M. Anderson (Ed.), The development of intelligence (pp. 215-243). Hove, UK: Psychology Press.
- Dennis, M., Lockyer, L., Lazenby, A., Donnelly, R., Wilkinson, M., & Schoonheydt, W. (1999). Intelligence patterns among children with high-functioning autism, phenylketonuria, and childhood head injury. Journal of Autism & Developmental Disorders, 29, 5-17.
- Duncan, J., Emslie, H., & Williams, P. (1996). Intelligence and the frontal lobes: The organization of goal-directed behavior. Cognitive Psychology, 30, 257-303.
- Dunn, J. (1995). Children as psychologists: The later correlates of individual differences in understanding of emotions and other minds. Cognition & Emotion, 9, 187-201.
- DuPaul, G., Anastopoulos, A., Power, T., Reid, R., Ikeda, M., & McGoe, K. (1998). Parent ratings of attention deficit/hyperactivity disorder symptoms: Factor structure and normative data. Journal of Psychopathology and Behavioral Assessment, 20, 83-102.

- Dyck, M., Farrugia, C., Shochet, I., & Holmes-Brown, M. (2002). Emotion recognition ability in deaf or blind children: Do sounds, sights, or words make the difference? Manuscript submitted for publication.
- Dyck, M., Ferguson, K., & Shochet, I. (2001). Do autism spectrum disorders differ from each other and from non-spectrum disorders on emotion recognition tests? European Child and Adolescent Psychiatry, 10, 105-116.
- Eisenberg, N., Murphy, B., & Shepard, S. (1996). The development of empathic accuracy. In W. Ickes (Ed.), Empathic accuracy (pp. 73-115). New York, NY: Guilford Press.
- Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A. Jr., Winokur, G., Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. Archives of General Psychiatry, 26, 57-63.
- Flavell, J. (1992). Perspectives on perspective-taking. In H. Beilin & P. Pufall (Eds.), Piaget's theory: Prospects and possibilities (pp. 107-139). Hillsdale, NJ: Erlbaum.
- Fox, A., Michie, P., Wynne, C., & Maybery, M. (2000) ERP correlates of response inhibition to elemental and configural stimuli in a negative patterning task. Clinical Neurophysiology, 111, 1045-1053.
- Gal, E., Dyck, M., & Passmore, A. (2002). Sensory differences and stereotyped movements in children with autism. Behaviour Change, 19, ??-??.
- Ghaziuddin, M., Tsai, L., & Ghaziuddin, N. (1992). Comorbidity of autistic disorder in children and adolescents. European Child and Adolescent Psychiatry, 1, 209-213.
- Gillberg, C. (1992). Autism and autistic-like conditions: Subclasses among disorders of empathy. Journal of Child Psychology and Psychiatry, 33, 813-842.
- Gonzalez, J., & Espinel, A. (1999). Is IQ-achievement discrepancy relevant in the definition of arithmetic learning disabilities? Learning Disability Quarterly, 22, 291-301.

- Happe, F. (1994a). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. Journal of Autism and Developmental Disorders, 24, 129-154.
- Happe, F. (1994b). Wechsler IQ profile and theory of mind in autism: A research note. Journal of Child Psychology and Psychiatry, 35, 1461-1471.
- Harris, P., Johnson, C., Hutton, D., Andrews, G., & Cooke, T. (1989). Young children's theory of mind and emotion. Cognition and Emotion, 3, 379-400.
- Hartman, C., Hox, J., Auerbach, J., Erol, N., Fonseca, A., Mellenbergh, G., Novik, T., Oosterlaan, J., Roussos, A., Shalev, R., Zilber, N., & Sergeant, J. (1999). Syndrome dimensions of the Child Behavior Checklist and the Teacher Report Form: A critical empirical evaluation. Journal of Child Psychology and Psychiatry, 40, 1095-1116.
- Hartman, C., Hox, J., Mellenbergh, G., Boyle, M., Offord, D., Racine, Y., McNamee, J., Gadow, K., Sprafkin, J., Kelly, K., Nolan, E., Tannock, R., Schachar, R., Schut, H., Postma, I., Drost, R., & Sergeant, J. (2001). DSM-IV internal construct validity: When a taxonomy meets data. Journal of Child Psychology and Psychiatry, 42, 817-836.
- Hawkins, K., & Tulky, D. (2001). The influence of IQ stratification on WAIS-III/WMS-III FSIQ-General Memory Index discrepancy base-rates in the standardization sample. Journal of the International Neuropsychological Society, 7, 875-880.
- Hazell, P., Carr, V., Lewin, T., Dewis, S., Heathcote, D., & Brucki, B. (1999). Effortful and automatic information processing in boys with ADHD and specific learning disorders. Journal of Child Psychology and Psychiatry, 40, 275-286.

- Humphries, T., & Bone, J. (1993). Validity of IQ-achievement discrepancy criteria for identifying learning disabilities. Canadian Journal of School Psychology, 9, 181-191.
- Ickes, W. (1993). Empathic accuracy. Journal of Personality, 61, 587-610.
- Iverson, G., Woodward, T., & Green, P. (2001). Base rates of WAIS-R VIQ-PIQ differences in 1593 psychiatric inpatients. Journal of Clinical Psychology, 57, 1579-1587.
- Kamieniecki, G., & Lynd-Stevenson, R. (2002). Is it appropriate to use United States norms to assess the “intelligence” of Australian children? Australian Journal of Psychology, 54, 67-78.
- Kanner, L. (1944). Early infantile autism. Journal of Pediatrics, 25, 211-217.
- Kaplan, B., Wilson, B., Dewey, D., & Crawford, S. (1998). DCD may not be a discrete disorder. Human Movement Science, 17, 471-490.
- Kent, L., Evans, J., Paul, M., & Sharp, M. (1999). Comorbidity of autistic spectrum disorders in children with Down syndrome. Developmental Medicine and Child Neurology, 41, 153-158.
- Lahey, B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G., Barkley, R., Newcorn, J., Jensen, P., Richters, J., Garfinkel, B., Kerdyk, L., Frick, P., Ollendick, T., Perez, D., Hart, E., Waldman, I., & Shaffer, D. (1994). DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents. American Journal of Psychiatry, 151, 1673-1685.
- Larson, V., & McKinley, N. (1995). Language disorders in older students. Eau Claire, WN: Thinking Publications.
- Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Watson, P., & Robertson, I. (2001). The differential assessment of children’s attention: The test of everyday

- attention for children (TEA-Ch), Normative sample and ADHD performance. Journal of Child Psychology and Psychiatry, 42, 1065-1081.
- Matsumoto, D., & Ekman, P. (1995). Japanese And Caucasian Facial Expressions Of Emotion (JACFEE) And Neutral Faces (JACNeuF). San Francisco State University, San Francisco.
- McCarron, L. (1982). McCarron assessment of neuromuscular development: Fine and gross motor abilities. Texas: McCarron-Dial Systems.
- McCarron, L. (1997). McCarron assessment of neuromuscular development: fine and gross motor abilities (revised). Texas: McCarron-Dial Systems, Inc.
- Neuman, R., Heath, A., Reich, W., Bucholz, K., Madden, P., Sun, L., Todd, R., & Hudziak, J. (2001). Latent class analysis of ADHD and comorbid symptoms in a population sample of adolescent female twins. Journal of Child Psychology and Psychiatry, 42, 933-942.
- Oosterlaan, J., Logan, G., & Sergeant, J. (1998). Response inhibition in ADHD, CD, comorbid ADHD + CD, anxious and normal children: A meta-analysis of studies with the stop task. Journal of Child Psychology and Psychiatry, 39, 411-426.
- Ozonoff, S., Pennington, B., & Rogers, S. (1991). Executive function deficits in high functioning autistic children: Relationship to theory of mind. Journal of Child Psychology and Psychiatry, 31, 1081-1105.
- Pearson, D., & Lachar, D. (1994). Using behavioral questionnaires to identify adaptive deficits in elementary-school children. Journal of School Psychology, 32, 33-52.
- Pennington, B., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. Journal of Child Psychology and Psychiatry, 37, 51-87.
- Perner, J., Frith, U., Leslie, A., & Leekam, S. (1989). Exploration of the autistic child's theory of mind: Knowledge belief and communication. Child Development, 60, 689-700.

- Perner, J. & Wimmer, H. (1985). "John thinks that Mary thinks that..." Attribution of second-order beliefs by 5-to 10-year-old children. Journal of Experimental Child Psychology, 39, 437-471.
- Piek, J., Dworcan, M., Barrett, N., & Coleman, R. (2000). Determinants of self worth in children with and without Developmental Coordination Disorder. International Journal of Disability, Development and Education, 47, 259-272.
- Piek, J., & Edwards, K. (1997). The identification of children with developmental coordination disorder by class and physical education teachers. British Journal of Educational Psychology, 67, 55-67.
- Piek, J., Pitcher, T., & Hay, D. (1999). Motor coordination and kinaesthesia in boys with attention deficit-hyperactivity disorder. Developmental Medicine and Child Neurology, 41, 159-165.
- Pierre, C., Nolan, E., Gadow, K., Sverd, J., & Sprafkin, J. (1999). Comparison of internalizing and externalizing symptoms in children with attention-deficit hyperactivity disorder with and without comorbid tic disorder. Journal of Developmental and Behavioral Pediatrics, 20, 170-176.
- Pitcher, T., Piek, J., & Barrett, N. (2002). Timing and force variables in children with and without movement and attention deficits. Human Movement Science, 21, 919-945.
- Plante, E. (1998). Criteria for SLI: The Stark and Tallal legacy and beyond. Journal of Speech Language & Hearing Research, 41, 951-957.
- Prior, M., Smart, D., Sanson, A., & Oberklaid, F. (1999). Relationships between learning difficulties and psychological problems in preadolescent children from a longitudinal sample. Journal of the American Academy of Child & Adolescent Psychiatry, 38, 429-436.
- Rabbitt, P. (1997). Methodology of frontal and executive function. East Sussex, UK: Psychology Press.

- Rispens, J., van Yperen, T. A. (1997). How specific are "specific developmental disorders"? The relevance of the concept of specific developmental disorders for the classification of childhood developmental disorders. Journal of Child Psychology & Psychiatry, 38, 351-363.
- Rutter, M., Sroufe, L. A. (2000). Developmental psychopathology: concepts and challenges. Developmental Psychopathology, 12, 265-296.
- Sawyer, M., Arney, F., Baghurst, P., Clarke, J., Graetz, B., Kosky, R., Nurcombe, B., Patton, G., Prior, M., Raphael, B., Rey, M., Whaites, L., & Zubrick, S. (2001). The mental health of young people in Australia: Key findings from the child and adolescent component of the National Survey of Mental Health and Well-being. Australian & New Zealand Journal of Psychiatry, 35, 806-814.
- Sedler, M. (1994). Foundations of the new nosology. Journal of Medicine & Philosophy, 19, 219-238.
- Semel, E., Wiig, E., & Secord, W. (1995). Clinical evaluation of language fundamentals (3rd ed.): Technical manual. San Antonio, TX: The Psychological Corporation.
- Shaywitz, S., Escobar, M., Shaywitz, B., Fletcher, J., & Makuch, R. (1992). Evidence that dyslexia may represents the lower tail of a normal distribution of reading ability. New England Journal of Medicine, 326(3), 145-50.
- Sheppard, D., Bradshaw, J., Pucell, R., & Pantelis, C. (1999). Tourette's and comorbid syndromes: Obsessive compulsive and attention deficit hyperactivity disorder. A common etiology? Clinical Psychology Review, 19, 531-552.
- Shue, K., & Douglas, V. (1992). Attention deficit hyperactivity disorder and the frontal lobe syndrome. Brain and Cognition, 20, 104-124.
- Stanovich, K., & Stanovich, P. (1997). Further thoughts on aptitude/achievement discrepancy. Educational Psychology in Practice, 13, 3-8.

- Sternberg, R., & Grigorenko, E. (2002). Difference scores in the identification of children with learning disabilities: It's time to use a different method. Journal of School Psychology, 40, 65-83.
- Swanson, J., Schuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J., Clevenger, W., Wasdell, M., & McCleary, R. (2002). Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and the SWAN ratings scales. <http://www.adhd.net>.
- Turner, M. A. (1997). Towards an executive dysfunction account of repetitive behaviour in autism. In J. Russell (Ed.), Autism as an executive disorder (pp. 57-100). Oxford: Oxford University Press.
- Vellutino, F., Scanlon, D., & Lyon, G. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. Journal of Learning Disabilities, 33, 223-238.
- Verhoeven, W., & Tuinier, S. (1997). Neuropsychiatric consultation in mentally retarded patients: A clinical report. European Psychiatry, 12, 242-248.
- Wechsler, D. (1991). Wechsler intelligence scale for children – Third edition: Manual. San Antonio, TX: The Psychological Corporation.
- Weirs, R., Gunning, W., & Sergeant, J. (1998). Is a mild deficit in executive functions in boys related to childhood ADHD or to parental multigenerational alcoholism? Journal of Abnormal Child Psychology, 26, 415-430.
- Wiig, E., & Harris, S. (1974). Perception and interpretation of nonverbally expressed emotions by adolescents with learning disabilities. Perceptual & Motor Skills, 38, 239-245.
- Wilkinson, S. (1993). WISC-R profiles of children with superior intellectual ability. Gifted Child Quarterly, 37, 84-91.

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception.

Cognition, 13, 103-127.

World Health Organisation (1992). The ICD-10 classification of mental and behavioural disorders. Geneva: World Health Organisation.

Acknowledgements

This research was supported by grants from the National Health and Medical Research Council (Project Grant #141107) and the Research Centre for Applied Psychology, Curtin University of Technology. We wish to thank the principals and staff of participating schools for their cooperation, and especially to thank the participating children and parents who made this study possible.

Table 1

Age & Sex of Participants

Age	Female	Male	n	NDI	EstPIQ	EstVIQ
3.0 – 3.99	13	13	26	89		
4.0 – 4.99	23	19	42	88		
5.0 – 5.99	20	24	44	91		
6.0 – 6.99	16	19	35	101	127	122
7.0 – 7.99	21	20	41	96	117	107
8.0 – 8.99	20	20	40	99	118	108
9.0 – 9.99	21	20	41	95	113	113
10.0 – 10.99	24	20	44	95	113	105
11.0 – 11.99	18	21	39	99	113	106
12.0 – 12.99	22	16	38	93	116	109
	198	192	390	95	117	110

EstVIQ = Estimated Verbal IQ

EstPIQ = Estimated Performance IQ

NDI = Neuromuscular Development Index

Table 2

Parameters of Discrepancy Score Distributions and Correlations Between Specific Ability Quotients (SAQ)

						Discrepancy with IQ		
	IQ	LA	MC	EA	N	Min/Max	Mean	Sd
LA	.66***				390	-34.4/51.9	-0.09	12.22
MC	.21***	.22***			385	-44.4/49.6	-0.57	21.16
EA	.49***	.59***	.23***		390	-73.6/79.2	-0.12	15.05
AC	.16*	.12	.22***	.19**	245	-53.5/61.2	-0.38	19.46

* $p < .05$, two-tailed

** $p < .01$, two-tailed

*** $p < .001$, two-tailed

Abbreviations: LA = Language Ability, MC = Motor Coordination, EA = Empathic Ability, AC = Attentional Control

Table 3

Differences Between Children Who Have, and Do Not Have, a Substantial Discrepancy Between Intelligence and a Specific Ability

	Average IQ	Average LA	Average MC	Average EA	Average AC
IQ>Lang (n=4)	113	72	83	78	102
IQ=Lang (n=156)	100	99	99	101	100
	t=2.01, p<.05	t=4.02, p<.001	t=1.51, ns	t=3.07, p<.01	T=0.19, ns
IQ>Motor (n=28)	113	105	68	103	100
IQ=Motor (n=102)	99	98	99	98	100
	t=5.05, p<.01	t=2.29, p<.10	t=9.56, p<.01	t=1.28, ns	t=0.11, ns
IQ>Empathy (n=10)	115	90	93	78	93
IQ=Empathy (n=129)	99	99	101	100	100
	t=3.44, p<.01	t=1.64, ns	t=1.35, ns	t=4.79, p<.01	t=0.97, ns
IQ>Attention (n=18)	115	112	97	102	76
IQ=Attention (n=65)	101	102	107	104	101
	t=4.39, p<.01	t=2.99, p<.01	t=2.19, p<.05	t=0.48, ns	t=8.17, p<.01

Abbreviations: LA = Language Ability, MC = Motor Coordination, EA = Empathic Ability, AC = Attentional Control

Table 4

Mean CBCL Scores of Children with Specific Ability Quotients < 85 on Intelligence, Language, Motor Coordination, Empathic Ability, or Attentional Control Measures

	Anx/Dep	Aggression	Attention	Delinq	Other	Social	Somatic	Thought	Withdrawn
SAQ-IQ	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
< 85 (n=65)	3.31 (4.36)	6.37 (6.39)	2.93 (3.03)	1.39 (1.88)	5.84 (6.48)	1.73 (2.21)	1.76 (2.93)	0.92 (1.72)	1.53 (2.23)
> 84 (n=318)	2.77 (3.27)	6.00 (5.51)	2.59 (2.97)	1.32 (1.72)	4.33 (3.88)	1.46 (2.07)	1.15 (1.57)	0.50 (1.00)	1.27 (1.71)
	F=1.25, ns	F<1, ns	F< 1, ns	F<1, ns	F=6.23, p<.05	F<1, ns	F=5.86, p<.05	F=6.77, p<.01	F=1.04, ns
SAQ-Lang	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
< 85 (n=58)	3.18 (3.95)	6.29 (5.95)	3.12 (3.23)	1.46 (1.86)	5.96 (5.86)	2.08 (2.59)	1.74 (2.71)	0.91 (1.66)	1.50 (1.94)
> 84 (n=324)	2.81 (3.39)	6.02 (5.62)	2.57 (2.93)	1.31 (1.72)	4.33 (4.11)	1.40 (1.98)	1.16 (1.68)	0.51 (1.04)	1.28 (1.79)
	F<1, ns	F<1, ns	F= 1.67, ns	F<1, ns	F=6.67, p<.01	F=4.24, p<.05	F=4.56, p<.05	F=5.73, p<.05	F<1, ns
SAQ-Motor	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
< 85 (n=65)	2.77 (3.79)	6.79 (6.46)	3.09 (3.25)	1.29 (1.83)	5.40 (5.62)	1.90 (2.15)	1.69 (2.43)	0.73 (1.40)	1.42 (2.07)
> 84 (n=313)	2.90 (3.44)	5.93 (5.44)	2.56 (2.91)	1.34 (1.73)	4.43 (4.16)	1.42 (2.08)	1.15 (1.74)	0.54 (1.11)	1.30 (1.76)
	F<1, ns	F=1.23, ns	F= 1.67, ns	F<1, ns	F=2.54, ns	F=2.76, p<.10	F=4.32, p<.05	F=1.42, ns	F<1, ns
SAQ-Empathy	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
< 85 (n=59)	3.65 (4.69)	7.86 (5.73)	2.89 (2.91)	1.52 (1.87)	5.91 (6.04)	2.03 (2.45)	1.94 (3.01)	1.00 (1.65)	1.55 (2.17)
> 84 (n=324)	2.72 (3.20)	5.73 (5.43)	2.60 (2.99)	1.30 (1.72)	4.34 (4.06)	1.41 (2.01)	1.12 (1.57)	0.50 (1.03)	1.27 (1.74)
	F=3.50, p<.10	F=7.14, p<.01	F< 1, ns	F<1, ns	F=6.31, p<.05	F=4.41, p<.05	F=9.62, p<.01	F=9.33, p<.01	F=1.22, ns
SAQ-Attention	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
< 85 (n=39)	3.78 (3.78)	6.94 (5.30)	4.71 (3.73)	1.52 (2.00)	5.68 (5.45)	2.73 (2.75)	2.12 (2.85)	0.92 (1.71)	1.73 (2.21)
> 84 (n=199)	3.00 (3.59)	5.83 (5.67)	2.56 (2.89)	1.46 (1.78)	4.05 (3.83)	1.51 (2.04)	1.22 (1.57)	0.52 (1.01)	1.44 (1.92)
	F=1.51, ns	F=1.24, ns	F=15.84, p<.01	F<1, ns	F=4.99, p<.05	F=10.15, p<.01	F=7.87, p<.01	F=3.81, p<.10	F<1, ns

Significant differences are in **boldface**

Table 5

Average Specific Ability Quotients of Children With a Specific Ability Quotient Less Than 85 on an Index Ability

Index Ability	SAQ-IQ	SAQ-LA	SAQ-MC	SAC-EA	SAQ-AC
SAQ-Intelligence					
< 85 (n = 66)	77	85	93	87	95
> 84 (n = 324)	104	102	101	102	101
	F=323.71, p<.01	F=95.00, p<.01	F=11.98, p<.01	F=65.77, p<.01	F=5.94, p<.05
SAQ-Language					
< 85 (n = 59)	84	75	92	83	97
> 84 (n = 331)	102	104	101	103	100
	F=92.85, p<.01	F=325.99, p<.01	F=12.14, p<.01	F=114.61, p<.01	F<1, ns
Motor < 85					
< 85 (n = 68)	94	94	71	91	95
> 84 (n = 317)	101	101	106	101	100
	F=11.42, p<.01	F=10.61, p<.01	F=419.25, p<.01	F=25.43, p<.01	F=3.14, p<.10
Empathy < 85					
< 85 (n = 59)	88	84	93	75	96
> 84 (n = 323)	102	102	101	104	100
	F=47.82, p<.01	F=85.93, p<.01	F=9.35, p<.01	F=370.18, p<.01	F=2.16, ns
Attention < 85					
< 85 (n = 39)	95	97	94	95	74
> 84 (n = 206)	100	100	105	100	104
	F=3.73, p<.10	F=1.21, ns	F=15.75, p<.01	F=3.81, p<.10	F=300.39, p<.01

Significant differences are in **boldface**

Table 6

Average Specific Ability Quotients of Children with a Clinically Significant Behavioral Disturbance

CBCL Scale	Intelligence	Language	Motor	Empathy	Attention
Aggressive					
> 2sd (n=5/10)	90	95	90	93	97
< 2sd (n=245/380)	100	100	100	100	100
	F=3.82, p<.10	F<1, ns	F=3.17, p<.10	F=1.84, ns	F<1, ns
Anx/Depressed					
> 2sd (n=10/12)	96	97	90	90	97
< 2sd (n=235/378)	99	100	100	100	100
	F<1, ns	F<1, ns	F=3.48, p<.10	F=4.89, p<.05	F<1
Attention					
> 2sd (n=8/11)	93	87	92	98	92
< 2sd (n=237/379)	100	100	100	100	100
	F=2.19, ns	F=8.24, p<.01	F=2.05, ns	F<1, ns	F=2.20, ns
Delinquent					
> 2sd (n=13/17)	89	91	97	93	95
< 2sd (n=245/373)	100	100	100	100	100
	F=7.84, p<.01	F=5.50, p<.05	F<1, ns	F=3.86, p<.05	F=1.03, ns
Other					
> 2sd (n=8/16)	90	92	89	92	89
< 2sd (n=237/374)	100	100	100	100	100
	F=6.14, p<.05	F=4.31, p<.05	F=5.91, p<.05	F=4.48, p<.05	F=3.88, p<.05
Social					
> 2sd (n=16/19)	93	91	95	93	89
< 2sd (n=229/371)	100	100	100	100	100
	F=3.57, p<.10	F=6.15, p<.05	F=1.59, ns	F=4.35, p<.05	F=8.88, p<.01
Somatic					
> 2sd (n=12/20)	93	94	97	91	92
< 2sd (n=233/370)	100	100	100	100	100
	F=3.27, p<.10	F=2.54, ns	F<1, ns	F=7.22, p<.01	F=3.38, p<.10
Thought					
> 2sd (n=5/8)	93	90	91	90	99
< 2sd (n=240/382)	100	100	100	100	100
	F=1.33, ns	F=3.13, p<.10	F=1.76, ns	F=3.00, p<.10	F<1, ns
Withdrawn					
> 2sd (n=4/6)	93	95	87	94	94
< 2sd (n=241/384)	100	99	100	100	100
	F = 1.00, ns	F < 1, ns	F= 2.05, p < .10	F < 1, ns	F < 1, ns

The first n in each set refers to the number of participants who completed the attentional control measures, and the second n refers to the number of participants who completed the full set of measures.

Significant differences are in **boldface**