International Journal of Education and Social Science Research

ISSN 2581-5148

Vol. 1, No. 02; 2018

PSYCHOMETRIC PREDICTION OF THE DIAGNOSIS OF AUTISM SPECTRUM DISORDER

Scott Spreat
Woods Services Incorporated PO Box 36 Langhorne,
Pennsylvania, USA, 19047 Email Address: sspreat@woods.org

ABSTRACT
Children with and without diagnoses of autism spectrum disorder were compared on measures of adaptive behavior, challenging behavior, and emotional well being. Sixteen specific items were selected from the Behavior Development Survey, the Challenging Behavior Frequency Scale, and the Achenbach. The sixteen items were selected because of their congruence with the diagnostic guidelines set forth in the Diagnostic and Statistical Manual – IV. It was possible to reliably distinguish the two groups based on this block of 16 items. Follow up analysis revealed that as few as three items were necessary to distinguish the groups. Retained predictors (Speaks in Full Sentences, Repetitive Behaviors, and Orders Meals in Public) were able to explain approximately 40% of the variance in diagnostic assignment

KEYWORDS: autism, diagnosis

The tendency for behavioral definitions to change over time has been noted repeatedly in the behavior analysis literature. Using the term, Observer Drift, Kent, O'Leary, Diament, and Dietz (1974) described a phenomenon in which trained observers will "drift" in their application of an observational scheme. This "drift" can be of sufficient magnitude as to render observations collected early in a study incomparable with observations collected later in a study (Kazdin, 1977). Repeated training in the application of the definition has been cited as one possible remedy (Kazdin, 1977), as has the scoring of videotaped behavior samples in an order unrelated to their actual temporal occurrence.

Although observer drift has primarily been raised as a concern in research studies employing trained observers, one cannot help but note the similarities between the clinical diagnostic process and research data collection. Behavior is observed, and a judgment is made by a trained individual based on those observations. Observer drift must be considered a threat to the clinical diagnostic process as well as the research process. Diagnostic skills, developed in graduate school and refined through internships, would seem to be just as susceptible to drift as are the observations collected during the course of a research study. Arguably, the subjectivity that is explicit in most clinical diagnoses made on the basis of the Diagnostic and Statistical Manual (American Psychiatric Association, 1994) would seem to make the threat even greater for clinical diagnosis.

Let us consider the phenomenon of observer drift and autism. The recent increase in the number of

children diagnosed with autism has attracted widespread public interest. While most older textbooks place the prevalence of autism in the neighborhood of about 4.5 per 10,000 children, more recent researchers have reported dramatically higher rates of the disorder. A 1999 California study (California Department of Developmental Services, 1999) presented rates as high as 20-40 per 10,000 children. A number of cross-sectional studies done over time have suggested similarly higher estimates of the prevalence of autism (Gurney, Fritz, Ness, Sievers, Newschaffer, & Shapiro, 2003; Croen, Grether, Hoogstrate, &Selvin, 2002; Shattuck, 2006). In a series of literature review articles, Fombonne (2000; 2006) reported that the prevalence of autism had increased from around 5.2 per 10,000 births in 1966 to about 60 per 10,000 births in 2001. A study conducted by the Center for Disease Control (Rice, Baio, Van Naarden Braun, Doernberg, Pettygrove, Conniff, Meaney, Newschaffer, Zaharadny, Desposito, Charles, & Becker-Cottrill, 2007) has reported even high prevalence rates. Using a chart review methodology, these researchers suggested that the prevalence of autism might be as high as 1 in 150 (or about 67 per 10,000). A follow up to the initial CDC investigation reported a prevalence rate of 1 in 68 (Christensen, Baio, Braun, Bilder, Charles, Constantino, Daniels, Durkin, Fitzgerald, Kurzius-Spencer, Lee, Petty grove, Robinson, Schulz, Wells, Wingate, Zahorodny, &Yeargin-Allsopp, 2016).

.

Although there may be some disagreement regarding specific prevalence estimates, there appears to be no disagreement that the autism diagnosis is being used more frequently today than in the relatively recent past. Autism has come to be recognized as a spectrum disorder (Romanczyk, Arnstein, Soorya, & Gillis, 2003), in which symptoms may vary widely on the three primary diagnostic dimensions. The notion of a spectrum disorder is in contrast to the more narrow traditional definition of what is now sometimes called “classical autism” (Scott, Baron-Cohen, Bolton, &Brayne, 2002), in which one observes the characteristic presentation of a withdrawn, uncommunicative child engaging in stereotypy for lengthy periods of time. Many professionals now consider the apparent increase in the prevalence of autism to be an artifact of a broadening of the definition of autism, combined with improved case finding and greater public awareness. The evolution of the diagnosis from classical autism to a spectrum disorder is consistent with the notion of observer drift. The rules for making such a diagnosis have changed over time.

The most current version of the Diagnostic and Statistical Manual for Mental Disorders (American Psychiatric Association, 2013) provides guidelines for making a diagnosis of autism. These guidelines, however, rely heavily on the subjective assessment of symptoms, and they allow multiple routes to a diagnosis of autism. One might argue that these subjectively interpreted guidelines are precisely the condition that enabled the occurrence of diagnostic drift with respect to autism.

In similarly vague diagnostic situations, such as the assignment of levels of intellectual disability, some researchers have attempted to construct what they call a paramorphic representation of the clinical decision making process. Using a multivariate analysis, the association of a linear combination of clinical descriptors is assessed with respect to a given diagnosis. Roszkowski, Spreat, and Isett (1983), for example, used this sort of approach to describe how clinicians assign levels of what is now called intellectual disability. They found that contemporaneous with their study, clinicians relied primarily on IQ scores to establish level of intellectual disability, with only a limited contribution from measures of adaptive behavior. They noted that they were able to create a mathematical model that described, or offered a paramorphic representation, of the diagnostic

process. This representation did not actually identify the factors employed by the clinician, but rather described the outcome.

To the extent that the diagnosis of autism has 'drifted' from classical autism (Scott, Baron-Cohen, Bolton, &Brayne, 2006) it may be of interest to attempt to create a mathematical model of the diagnostic decision making process with respect to autism spectrum disorder. The resultant product should afford a perspective on those factors that effectively describe the Autism Spectrum Disorder diagnostic process as it is currently being used. The purpose of this study is to attempt to develop a mathematical representation of the contemporary diagnostic process with respect to autism spectrum disorder. Selected items from the Behavior Development Survey, the Challenging Behavior Frequency Scale, and the Achenbach will be used in an effort to predict whether a child has an autism diagnosis.

Methods

Participants

Participants in this study were 156 children who lived in a residential Approved Private School for individuals with significant developmental disabilities concomitant with the emission of dangerous assaultive or self-injurious behaviors. There were 123 males and 33 females. The average age was 16.6 (sd = 2.52); all were under 21 years of age. All of these diagnoses fell within the realm of autism spectrum disorder. A breakdown by level of intellectual disability revealed the following: 11 without intellectual disability; 52 with mild intellectual disability; 35 with moderate intellectual disability; 29 with severe intellectual disability; and 5 with profound intellectual disability. There were 24 individuals for whom level of intellectual disability was not classified. It should be noted that these 156 children came from seven different states, ensuring that we are studying the diagnostic processes of more than a handful of clinicians.

Instrumentation

Because the diagnosis of autism is based on impairments in social interaction, impairments in communication, and presence of restricted, repetitive and stereotyped patterns of behavior, items were selected from three routinely administered scales to assess these three dimensions.

The Behavior Development Survey is a shortened version of the American Association on Mental Retardation's Adaptive Behavior Scale (Nihira, Foster, Shellhaas, and Leland, 1974). It consists of 32 items, and it yields raw scores that range from 0 to 129, with lower scores indicating increasingly diminished levels of adaptive behavior. It is administered by interviewing a third party who is familiar with the person whose adaptive behavior is being assessed. Pawlarczyk and Schumacher (1983) reported that the Behavior Development Survey is a concurrently valid method of assessing adaptive behavior, based on a review of correlation matrices from simultaneous administrations of the Behavior Development Survey, the AAMD Adaptive Behavior Scale, and the Vineland Social Maturity Scale. An unpublished study (Devlin, 1989) revealed an inter-rater reliability of .91 for the Behavior Development Survey. From the Behavior Development Survey, the following items were selected for analysis:

Orders Meals in Public; Demonstrates Pre-verbal Expressions; Speaks in Full Sentences; Shows Awareness of Others; Interacts with Others; and Participates in Group Activities

The Challenging Behavior Frequency Scale (Conroy, 1997) is a 19 item scale that assesses the frequency of various challenging behaviors via report of a knowledgeable third party respondent. Each item represents a discrete behavior, and each item is rated within the context of the previous month as none, mild, moderate, major, or extreme. Higher scores are indicative of greater degrees of challenging behaviors. Reliability of this instrument is reported to be in the neighborhood of .65 (Fullerton, Douglas, & Dodder, 1999), a figure that is generally considered adequate for research purposes, although insufficient for use with individuals. From this instrument, the following items were selected for inclusion in the study: Self Injurious Behavior; Social Withdrawal; and Repetitive Behaviors.

The Achenbach System of Empirically Based Assessment (Achenbach and Rescorla, 2001) for school aged children was administered to all study participants, using psychologists and case managers as informants. The Achenbach consists of 113 discrete items that assess competencies, adaptive functioning, and problems. The Achenbach manual (Achenbach and Rescorla, 2001) reports scale reliability indices in excess of .90 (From this instrument, the following items were selected for inclusion in the study: Refuses to talk; Speech Problems; Rather Be Alone; Harms Self; Doesn't Eat Well; Repeats Acts Over and Over; and Strange Behaviors.

Results

The 16 selected items from the three scales were submitted to discriminant function analysis in an effort to predict membership in either the autism group or the non-autism group. Because the selection of items was based on similarity to the diagnostic schema, the items were entered simultaneously.

The initial step in a discriminant function analysis involves determining whether there is an overall between groups difference on the linear combination of predictor variables. A Wilks lambda of .597, approximately equivalent to a Chi-square value of 72.245 (df = 16, p = .000), indicated that the group centroids were indeed significantly different with respect to the selected items. Wilks Lambda is an inverse measure of variance explained, and in this case, it indicated that the linear combination of the 16 variables was able to explain just over 40% of the variance in diagnostic group membership. This would be approximately equivalent to a multiple R of .63. The analysis was able to correctly classify 79.3% of the individuals in the study. Examination of the resultant structure matrix identified the key contributors to the discriminant function as Use of Sentences, Ordering Meals in Public, Awareness of Others, interaction with Others, Repeats Over and Over, Group Participation, and Speech Problems.

These findings were generally supported by a supplemental logistic regression analysis, in which the same 16 predictor variables achieved statistical significance (Chi-Square [16] = 75.824, p = .000), and a Nagelkerke r-square value of .541 was determined. The Nagelkerke value is not directly comparable with a Wilks value or a traditional R square value. Of particular note is the finding that the Hosmer/Lemeshowtest suggested that our model was an adequate fit for the data, able to correctly classify 80.00% of our sample.

Because the analysis of the discriminant function structure matrix had identified stronger and weaker

contributors, we sought to determine whether a lesser number of items might be used to adequately predict membership in the autism and non-autism diagnostic groups. To this end, the above described discriminant function analysis was repeated using stepwise entry procedures. A Wilks lambda of .639 was determined, and the linear combination of predictor variables Speaks in Full Sentences, Repeats Over and Over, and Orders Meals in Public achieved statistical significance (Chi-square = 65.597, df = 3, p = .0000). The three items were able to correctly classify 81.33% of the participants). As expected, the same three variables were identified in the same order by a stepwise entry logistic regression analysis. The linear equation achieved statistical significance (Chi-square = 64.981, df = 3, p = .000), with a Nagelkerke R square value of .479. The equation was not revealed to be a poor fit by the non-significant Hosmer / Lemeshow Goodness of fit test, and it was able to correctly classify 81.33% of the sample correctly. It should be noted that the diagnosis of Autism Spectrum Disorder was associated with higher scores on the Repeats Over and over item and lower scores on the Speaks in Full Sentences and Orders Meals in Public items.

The jackknifing technique described by Eyman, Meyers, and Bendel (1973) was used to cross validate the results obtained from the 16 item analysis. The total sample was divided into four smaller cohorts of 39 each, and each cohort was held out in turn while the discriminant function analysis was performed on the group comprised of the remaining three cohorts. Pseudo values were calculated for each analysis using the procedure described by Kier (1997). The resultant four pseudo-values for Wilks lambda were .624, .675, .615, and .702. These four values were averaged using Fishers Z-score transformation, and a jackknifed Wilks' lambda estimate of .654 was obtained. This less biased estimated compared favorably with the Wilks' lambda value of .597 from the original analysis, and it suggests that sample specific variance did not unduly inflate the results of the original analysis.

Discussion

Absent biological or genetic markers, the process of diagnosing conditions such as Autism Spectrum Disorder falls to a set of guidelines. General guidelines, such as those used for diagnosis of autism, are subject to variable application, and many clinicians believe that we have observed an evolution, or drift, in the application of autism diagnosis guidelines. It is not the place of this article to argue whether the current higher estimates of autism prevalence are more or less valid that the lower reported values of 25 years ago. The study does, however, demonstrate that the contemporary process of diagnosing Autism Spectrum Disorder can be fairly well described by a limited sample of relatively reliable behavioral items. It is conceivable that a psychometric approach to diagnosing autism could be developed, and it could serve to eliminate some of the controversy regarding the diagnosis of Autism Spectrum Disorder by ensuring the ongoing integrity of the diagnostic process. Adoption of an actuarial approach should stop further diagnostic drift. It was noted years ago and probably still true today (Wiggins, 1981) that actuarially based procedures were demonstrably superior to clinically determined ones in most situations.

Although the statistical analyses achieved significant predictability from both the pre-selected 16 items and the statistically reduced listing of 3 items, in both cases the percentage of explained variation in the application of the diagnosis was in the neighborhood of 40%. The selected variables were able to explain less than half the variance in diagnostic group membership. Clearly other variables and factors are at work, and we would be remiss to suggest that the current process could

be adequately replaced with just three (or 16) items. Additional research is necessary to delineate other factors. Further, one must note that the statistical procedures employed in this study are based on a linear combination of predictor variables. Clinicians, however, may view predictor variables from more of a configurational perspective (Anderson, 1972). Nevertheless, it must be recognized that the three key items (Speaks in Full Sentences, Repeats Over and Over, and Orders Meals in Public) would appear to constitute a simple screening tool for possible use by pediatricians. Note also that it is likely that other combinations of as little as three variables might achieve near equal levels of predictability.

It is perhaps interesting that predictability was only slightly diminished by dropping 13 of the 16 predictors, leaving only Speaks in Full Sentences, Orders Meals in Public, and Repetitive Behaviors. Items representing only two of the three diagnostic indices were necessary to reliably predict the diagnosis of autism. None of the five items that assess socialization were retained, and their absence did little to harm predictability. In an effort to offer a retrospective hypothesis regarding this finding, it is noted that the three socialization items from the Behavior Development Survey (Awareness of Others, Interaction with Others, and Participates in Groups) yielded a mean correlation of .51 with Orders Meals in Public and .57 with Speaks in Full Sentences. In contrast, neither the Challenging Behavior "Social Withdrawal" nor the Achenbach "Rather be Alone" item achieved correlations over .12 with either of the measures of speech use. Multicollinearity may be a partial explanation for the non-inclusion of the BDS variables, but not the Challenging Behavior or Achenbach items. One might be tempted to argue that the retained predictor items are actually measures of both communication and socialization. The use of communication skills, particularly within the context of the social situation of dining out, demands a certain degree of socialization.
As a limitation to the study, it should be noted that reliability of individual items is understood to be lower than the reliability of the scales from which the items were selected. This reduced reliability may result in an increased level of Type 2 error, in that it will be more difficult to find and identify statistically significant findings.

REFERENCES

Achenbach, T. and Rescorla, L. (2001).Manual for the ASEBA School AgeForms and Profiles. Burlington, Vermont: ASEBA.
American Psychiatric Association.Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Association; 2013.
Anderson, J. (1972). Looking for configuarility in clinical judgement.
Psychological Bulletin, 78-93-102.
California Department of Developmental Services. (1999). A report to the
legislature: Changes in the population of persons with autism and pervasive developmental disorders in California's Developmental Services system: 198 through 1998. Sacramento, CA: California Health and Human Services Agency.
Conroy, J. (1997). Adjustment and Challenges Scale: During the past 4 weeks.
Ardmore, PA: Center for Outcome Analysis.
Croen, L., Grether, J., Hoogstrate, J., &Selvin, S. (2002). The changing prevalence of autism in California. Journal of Autism and Developmental Disorders, 32(3), 207-215.
Devlin, S. (1989). Reliability assessment of the instruments used to monitor the
Pennhurst class members. Philadelphia: Temple University Developmental Disabilities Center.

Eyman, R., Meyers, C., &Bendel, R. (1973). New methods for test selection and reliability assessment using stepwise multiple regression and jackknifing. Educational and Psychological Measurement, 33, 883-894.
Fombonne, E. (2006). The "autism epidemic" and know causes of autism. Presentation at 10th International Child Neurology Conference, Montreal, June 15, 2006.
Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: an update. Journal of Autism and Developmental Disorders. 33(4), 365-382.
Fullerton, A., Douglass, M., & Dodder, R. (1999). A reliability study of measures assessing the impact of deinstitutionalization. Research in Developmental Disabilities, 20(6), 387-400.
Gurney, J., Fritz, M., Ness, K., Sievers, P., Newschaffer, C., & Shapiro, E.(2003). Analysis of prevalence trends of autism spectrum disorder in Minnesota. Archives of pediatrics and adolescent medicine, 157(7), 622-627.
Kazdin, A. (1977). Artifact, bias, and complexity of assessment: The ABCs of reliability. Journal of Applied Behavior Analysis, 10(1), 141-150.
Kier, F. (1997). Ways to explore the replicability of multivariate results (since statistical significance testing does not). Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, Texas, January, 1997.
Kent, R., O'Leary, K., Diament, C., & Dietz, A. (1974). Expectation biases in observational evaluation of therapeutic change. Journal of Consulting and Clinical Psychology, 42, 774-780.
Nihira, K., Foster, R., Shellhaas, M, & Leland, H. (1974). The AAMD Adaptive Behavior Scale. Washington, D.C.: American Association on Mental Deficiency.
Pawlarczyk, D. & Schumacher, K. (1983). Concurrent validity of the Behavior Development Survey. American Journal on Mental Deficiency, 87(6), 619-626.
Rice, C., Baio, J., Van Naarden Braun, K., Doernberg, N., Pettygrove, S., Conniff, C., Meaney, F., Newschaffer, C., Zaharadny, W., Desposito, F., Charles, J., & Becker-Cottrill, B. (2007). Prevalence of autism spectrum disorders: Autism and developmental disabilities monitoring network, six sites, United States, 2000. Morbidity and Mortality Weekly Report, 56, 1-10.Christensen D., Baio J., Braun K., Bilder, D., Charles, J., Constantino, J., Daniels, J., Durkin, M., Firzgerald, R., Kurzius-Spenser, M., Lee, L., Pettygrove, S., Robinson, C., Schulz, E., Wells, Cl;, Wingate, M., &Zahorodny, W., and Yeargin-Allsopp, M. (2016). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. MMWR SurveillSumm 2016;65(No. SS-3)(No. SS-3):1–23. DOI: http://dx.doi.org/10.15585/mmwr.ss6503a1.
Romanczyk, R., Arnstein, L, Soorya, L., & Gillis, J. (2003). The myriad of controversial treatments for autism. Chapter in S. Lilienfeld, S. Lynn, & J. Lohr (Eds.).Science and pseudoscience in clinical psychology. NYC: Guilford Press, pp. 363-395.
Roszkowski, M., Spreat, S., &Isett, R. (1983). A paramorphic representation of psychologists' clinical impression of degree of mental retardation. Journal of Psychoeducational Assessment, 1(3), 243-251.
Scott, F., Baron-Cohen, S, Bolton, P., &Brayne, C. (2002). Brief report prevalence of autism spectrum conditions in children aged 5-11 years in Cambridgeshire, UK. Autism, 6(3), 231-237.

Shattuck, P. (2006). The contribution of diagnostic substitution to the growing administrative prevalence of autism in US Special Education. Pediatrics, 117(4), 1028-1037.
Wiggins, J. (1981). Clinical and statistical prediction: Where are we and where do we go from here? Clinical Psychology Review, 1, 3-18.