UCSF
UC San Francisco Previously Published Works

Title

Initial diagnostic impressions of trainees during autism evaluations: High specificity but low sensitivity

Permalink

https://escholarship.org/uc/item/6pr0c0dt

Journal

Autism Research, 16(6)

ISSN

1939-3792

Authors

Marchena, Ashley
Wieckowski, Andrea Trubanova
Algur, Yasemin
et al.

Publication Date

2023-06-01

DOI

10.1002/aur.2933

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Initial diagnostic impressions of trainees during autism evaluations: High specificity but low sensitivity

Ashley de Marchena*[1], Andrea Trubanova Wieckowski*[1], Yasemin Algur[2], Lashae N. Williams[1], Sherira Fernandes[1], Rebecca P. Thomas[3], Leslie A. McClure[2], Sarah Dufek[4], Deborah Fein[3], Aubyn C. Stahmer[4], Diana L. Robins[1]

Affiliations: [1] A.J. Drexel Autism Institute, Drexel University, Philadelphia, PA, [2] Department of Epidemiology & Biostatistics, Drexel University, Philadelphia, PA, [3] Department of Psychological Sciences, University of Connecticut, Storrs, CT, [4] Department of Psychiatry and Behavioral Sciences, University of California, Davis MIND Institute, Sacramento, CA

* Ashley de Marchena and Andrea Trubanova Wieckowski should be considered joint first author.

Corresponding Author:
Ashley de Marchena
Abd64@drexel.edu

Acknowledgments:
We acknowledge the contributions and memory of Lauren B. Adamson, who passed away in 2022. We thank all of the children and families for participating in the study. In addition, we thank members of the research team who participated in data collection, specifically Sarah Nanovic and Katherine Sand for their management of the data. Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD039961, and by the National Institute Of Mental Health of the National Institutes of Health under Award Number R01MH115715. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability Statement
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Compliance with Ethical Standards:
Drs. Robins and Fein are co-owners of M-CHAT LLC, which licenses use of the M-CHAT in electronic products. No royalties were received for any of the data presented in the current study. Dr. Robins serves on the advisory board of Quadrant Biosciences Inc. The other authors have indicated they have no potential conflicts of interest to disclose. The research involved human participants. The study protocol was approved by the Institutional Review Board at each data collection site for R01HD039961 grant, and at Drexel University for the multi-site study for R01MH115715 grant.

Lay Summary

After only a five-minute observation, trainee clinicians noted whether or not they believed that a child attending an autism evaluation would be diagnosed with autism. When they indicated a child was autistic, they were correct in 86% of cases, but when they indicated a child was not autistic, they *missed* 29% of children ultimately diagnosed. These findings suggest that clinical judgments of suspected autism should be taken seriously, but *lack of practitioner concern* should not be used to rule out autism.

Abstract

Reducing the age of first autism diagnosis facilitates access to critical early intervention services. A current "waitlist crisis" for autism diagnostic evaluation thus demands that we consider novel use of available clinical resources. Previous work has found that expert autism clinicians can identify autism in young children with high specificity after only a brief observation; rapid identification by non-experts remains untested. In the current study, 252 children ages 12-53 months presented for a comprehensive autism diagnostic evaluation. We found that junior clinicians in training to become autism specialists (n=29) accurately determined whether or not a young child would be diagnosed with autism in the first five minutes of the clinic visit in 75% of cases. Specificity of brief observations was high (.92), suggesting that brief observations may be an effective tool for triaging young children toward autism-specific interventions. In contrast, the lower negative predictive value (.71) of brief observations, suggest that they should not be used to rule *out* autism. When trainees expressed more confidence in their initial impression, their impression was more likely to match the final diagnosis. These findings add to a body of literature showing that clinical observations of suspected autism should be taken seriously, but *lack of clinician concern* should not be used to rule out autism or overrule other indicators of likely autism, such as parent concern or a positive screening result.

Keywords: Autism; Early Detection; Diagnosis; Toddlers; Initial Impression; Diagnostic Confidence

Initial diagnostic impressions of trainees during autism evaluations: High specificity but low sensitivity

Introduction

Reducing the age of autism diagnosis is a pressing public health need. Parents of young children diagnosed with autism often report being concerned about their child's development in infancy (Richards et al., 2016), and autism can be diagnosed reliably by the second birthday (Pierce et al., 2021; Wieckowski, Hamner, et al., 2021), demonstrating that detection and diagnosis of autism can happen during the toddler years in many cases. However, the average age of autism diagnosis in the US is 4 years (Maenner et al., 2021) – a 2-year gap between when an autism diagnosis *could* be made for many children and when it typically *is* made. Furthermore, evidence suggests that children who receive high-quality early intervention services before their fourth birthday make greater developmental gains (Vivanti & Dissanayake, 2016), emphasizing the critical importance of lowering the age of diagnosis during this period of development.

One contributing factor to the long latency to diagnosis is that only a minority (i.e., 31%) of children who screen positive for autism are referred for diagnostic evaluation (Monteiro et al., 2019), as recommended by the American Academy of Pediatrics (Lipkin et al., 2020). Among pediatric practitioners, the most common reason for non-referral was a clinical impression of "no concerns", even with a positive screener (Pierce et al., 2021). This suggests that in some cases, clinicians are declining to refer children for evaluations because their clinical impressions are not concordant with screening results, clearly demonstrating the relevance of clinical impressions to screening and diagnosis. A second factor is the current shortage of expert clinicians with training

in autism evaluation (Kanne & Bishop, 2021), leading to calls for novel approaches both to expand the number of professionals who conduct autism assessments, and to develop more efficient approaches to maximize expert clinician time (Zwaigenbaum & Warren, 2021).

In our previous work, we looked at the match between expert clinicians' initial diagnostic impressions (i.e., after spending five minutes observing or interacting with a young child who had been referred for an autism evaluation) and final diagnosis made after a comprehensive evaluation (Wieckowski, de Marchena, et al., 2021). After this brief observation, when expert clinicians initially believed a child had autism, they were ultimately diagnosed 92% of the time, suggesting that a brief observation by an expert can detect autism in the vast majority of cases. Critically, experts also *missed* 24% of children who were later diagnosed. That is, even experts did not accurately rule out autism based on a brief encounter. Experts in this study also appeared to have good insight into the validity of their initial impressions: self-reported confidence in the initial impression was a strong predictor of match between initial impression and final diagnosis.

The current study seeks to extend these findings to training clinicians (hereafter referred to as "trainees") with less experience evaluating young children for autism. A growing body of literature suggests that even untrained raters can distinguish between autistic and non-autistic samples on the basis of brief observations (DeBrabander et al., 2019; Grossman, 2015; Sasson et al., 2017). Thus, junior or minimally trained clinicians could potentially be valuable resources for expediting the diagnostic process for young children; however, the accuracy of their initial impressions is unknown. In the current study, after spending five minutes with a child, each trainee documented their clinical first impression (i.e., yes-autism or no-autism) and their confidence in that impression. Children then completed a comprehensive diagnostic evaluation

with an expert, allowing us to compare the initial impressions and final diagnosis, and to examine clinician confidence, and agreement with initial impressions made by experts.

Methods

Participants

A total of 252 children (169 males), aged 12–53 months, were included (see Table 1). Participants completed a no-cost diagnostic evaluation at a university-based autism specialty clinic as part of two multi-site research studies examining early autism screening and referral methods. Inclusion criteria included competing an initial autism screener at 12-, 15-, or 18-month check-ups at participating pediatric practices, with potential repeat screening at 24 and/or 36 months and attending a diagnostic evaluation. Participants were referred for evaluation due to a positive result on an autism screener or a primary care practitioner's concern for autism during a pediatric well-child visit. Forty-six percent (n = 117 children) met criteria for autism after completing the evaluation.

Clinicians

During each child's diagnostic evaluation, one senior clinician and one trainee worked together to simultaneously collect child and parent measures. Senior clinicians were largely psychologists with some physicians, behavior analysts, occupational therapists, or master's level social workers with specific expertise in autism diagnosis. Trainees (n = 29; 28 female) were students and professionals at different levels. All trainees, regardless of professional/educational level, were training in autism diagnosis; all were interested in careers related to autism, and they varied in their previous exposure to young children on the autism spectrum. Critically, although trainees were relatively new to autism assessment, they were all in the autism field, and thus, not

naïve to the behavioral presentation of autism. The majority of the trainees contributing ratings were psychology graduate students (n = 17); other trainees included: undergraduate research assistant (n = 1), post baccalaureate research assistant (n = 4), social work graduate student (n = 1), MA-level clinicians in psychology (n = 3) or marriage and family therapy (n = 1), psychology post-doctoral fellow (n = 1), and one physician (licensed abroad in Family Medicine) re-specializing in autism research. Trainees' first impression data are included in the current study; for data on senior clinicians' first impressions, see Wieckowski and colleagues (2021).

Procedure

All parents gave informed consent; the study was approved by university Institutional Review Boards. More details about the procedure can be found in Wieckowski and colleagues (2021).

Evaluation

Evaluations included measures of cognitive/developmental skills (Mullen Scales of Early Learning; MSEL; Mullen, 1995), adaptive functioning (Vineland Adaptive Behavior Scales – second edition; VABS-2; or third edition; VABS-3; Sparrow et al., 2005, 2016), autism features (Autism Diagnostic Observation Schedule, Second Edition; ADOS-2; Lord et al., 2012), and developmental and medical history.

Evaluations were conducted by two clinicians: one senior clinician and one trainee, both of whom completed an initial rating form within the first five minutes of the evaluation (see below). Clinicians completing the diagnostic evaluations were naïve to referral source and screening results. Final diagnoses were based on International Classification of Diseases, tenth edition (ICD-10) criteria, integrating information obtained across all available measures.

Initial Impression Ratings

Five minutes after first observing the child, both the senior clinician and trainee independently documented their initial impression of whether they thought the child had autism, and how confident they were regarding this impression on a scale of 1 to 5, with 1 being *not very confident*, 3 being *confident*, and 5 being *extremely confident*. The context during which these observations occurred varied between participants.

Results

Clinicians' Initial Impressions of Diagnosis

Children initially classified as autistic were ultimately diagnosed with autism in 86% of cases; children initially believed to be non-autistic were ultimately diagnosed with autism in 29% of cases (Figure 1). Overall accuracy of initial impression was .75. Sensitivity of detecting autism within the first 5 minutes was low (.56) and specificity was high (.92); this is highly consistent with senior clinicians' impressions (Wieckowski et al., 2021). The positive predictive value of the initial impression was .86, and the negative predictive value was .71.

Among the 229 cases rated by both senior and trainee clinicians, there was substantial agreement between initial impressions ($\kappa = .651$, $p < 0.001$); percent agreement was 85.6%. Senior and trainee initial impressions differed in $n = 33$ cases; unsurprisingly, confidence ratings for both the senior ($F(1, 221)$ 19.001, $p < .001$, $\eta_p^2 = 0.079$) and trainee ($F(1, 218)$ 46.69, $p < .001$, $\eta_p^2 = 0.18$) clinicians were lower in these cases compared to the rest of the sample. The vast majority of cases for whom senior and trainee clinicians' initial impressions differed ($n = 25$) were ultimately diagnosed with autism ($\chi 2 = 14.320$, $p < .001$). In 64% of these cases, the

senior clinician gave an initial impression of autism, and in 36% of cases, the junior clinician gave an initial impression of autism.

Relationship Between Initial Impression, Final Diagnosis, and Confidence in Clinicians' Initial Impressions

Across diagnostic groups, trainees on average indicated feeling 'confident' in their initial impression (M = 3.21, SD = 1.14). A two-way ANOVA examined the relationship between initial impression and final diagnosis on clinicians' confidence. Neither initial impression (F(1, 236) < .001, p = .99, η_p^2 = 0.000) nor final diagnosis (F(1, 236) = 1.73, p = .19, η_p^2 = 0.07) had a main effect on trainees' confidence in their impression. However, the interaction between initial impression and final diagnosis – indicative of match – had a significant effect on confidence (F(1, 236) = 33.30, p < .001, η_p^2 = 0.124), with highest confidence when initial impression was non-autism and final diagnosis was non-autism (match; M = 3.60, SD = .98), and lowest confidence when initial impression was non-autism and final diagnosis was autism (mismatch; M = 2.24, SD = 0.89).

As shown in Figure 2, as trainees' confidence in their initial impressions rose, the likelihood that initial impression would match final diagnosis increased. In cases for which the trainee was "not very confident" (n = 22), the chances of match were 36%; when the trainee was "extremely confident (n = 27), however, the match rate was 96%.

Discussion

We found that even clinician trainees – who varied in autism experience and were not experts in autism diagnosis –identified autism after brief observations (five minutes) with high

specificity. In a sample of young children referred for autism evaluation, the greatest match between initial impression and final diagnosis was for children initially believed to have autism who were indeed ultimately diagnosed; 86% of cases initially judged by trainees to have autism were diagnosed following evaluation. Moreover, match between initial impression and final diagnosis increased with self-reported confidence: when trainees were extremely confident in their initial impression, overall match between initial impression and final diagnosis rose to 96%. Taken together with previous research on experts' initial impressions of autism (Wieckowski et al., 2021), these findings suggest that in *children whose autism is very apparent*, it is often readily observable to both experts and non-experts alike. Clinically, these findings indicate that when examiners suspect autism even after just a brief observation, this information could be utilized to refer the child to begin autism-specific early intervention, even before diagnostic evaluation can be completed.

In contrast to its potential utility as a clinical tool for ruling *in* autism, our findings strongly suggest that brief clinical observations should not be used to rule *out* autism. Indeed, 29% of children ultimately diagnosed with autism were initially believed not to be autistic by training clinicians (i.e., they were false negatives), a number only slightly higher than the 24% miss rate by expert clinicians (Wieckowski, de Marchena, et al., 2021). Therefore, brief observations should not be used to rule out autism, highlighting importance of not relying only on initial impressions but incorporating other strategies – including screening, parent concerns, and use of standardized observational tools – to inform referrals for diagnostic evaluation.

Reliability between initial impressions by senior clinicians was fairly high, and the general pattern of findings across both samples was extremely similar: specificity was high, sensitivity was low, and match between initial impression and final diagnosed increased with

confidence. This pattern of findings is consistent with other work finding higher specificity than sensitivity when autism diagnostic decisions are tested after a brief observation (Gabrielsen et al., 2015), and is also consistent with general patterns of identifying broad developmental-behavioral problems in primary care, where sensitivity is as low as 14-54% and specificity is as high as 69-100% (Sheldrick et al., 2011).

Given the current barriers to autism diagnoses in the United States (e.g., long wait lists, cost) it is important to evaluate the feasibility of tiered diagnostic approaches in which practitioners can rule autism in for children for whom diagnosis is clear, and refer less clear cases for specialist evaluations (e.g., Wieckowski et al., 2022). Results of this study suggest that even clinicians in training to become autism expert diagnostician can detect a large proportion of children within the first few minutes of an unstructured interaction; these children could be referred to start autism services as soon as possible, while awaiting a formal evaluation. Clinical norms suggest that comprehensive evaluations are essential in all cases (Kanne & Bishop, 2020); however, there is limited evidence to support that claim at a systems level. Indeed, putting every child on a referral list for a detailed evaluation prior to start of services displaces clinical resources that could be used to reduce overall waiting list times, and may ultimately reduce equitable access to services (MacLachlan, 2021). Based on the existing evidence and long waiting lists for diagnosis, we argue that comprehensive evaluation should not be the gatekeeper for initiating autism-specific interventions. Rather, these findings support the use of responsible triage processes in which some children are immediately referred to autism-specific interventions, some children are referred to traditional comprehensive evaluations, and yet other are referred for an even higher level of care. Larger system changes (e.g., payors) are required to accommodate more flexible, individualized approaches to autism assessment.

Limited information was available about the trainees in the current study; therefore, it was not possible to evaluate how examiner demographic factors (e.g., age, sex, race) or professional factors (e.g., discipline, years of training) may influence first impressions and confidence. We note that the sample in the current study consistent only of young children; thus, these findings may not apply to older children and adults with more complex or nuanced presentations. Future directions include extending to study the first impressions of primary care practitioners who conduct screening, and are generally responsible for the decision to refer or not refer children for further services (see Penner et al., 2023, for a recent example of autism diagnostic accuracy in primary care); these practitioners have broad expertise, but are generally not highly experienced detecting autism in toddlers. We hope that these findings can contribute to ongoing efforts to reduce age of diagnosis and entry into early intervention for all children on the autism spectrum.

References

DeBrabander, K. M., Morrison, K. E., Jones, D. R., Faso, D. J., Chmielewski, M., & Sasson, N. J. (2019). Do First Impressions of Autistic Adults Differ Between Autistic and Nonautistic Observers? *Autism in Adulthood*. https://doi.org/10.1089/aut.2019.0018

Gabrielsen, T. P., Farley, M., Speer, L., Villalobos, M., Baker, C. N., & Miller, J. (2015). Identifying Autism in a Brief Observation. *Pediatrics*, *135*(2), e330–e338.

Grossman, R. B. (2015). Judgments of social awkwardness from brief exposure to children with and without high-functioning autism. *Autism*, *19*(5), 580–587.

Kanne, S. M., & Bishop, S. L. (2021). Editorial Perspective: The autism waitlist crisis and remembering what families need. *Journal of Child Psychology and Psychiatry*, *62*(2), 140–142. https://doi.org/10.1111/jcpp.13254

Lipkin, P. H., Macias, M. M., COUNCIL ON CHILDREN WITH DISABILITIES, S. O. D. A. B. P., Norwood, K. W., Jr, Brei, T. J., Davidson, L. F., Davis, B. E., Ellerbeck, K. A., Houtrow, A. J., Hyman, S. L., Kuo, D. Z., Noritz, G. H., Yin, L., Murphy, N. A., Levy, S. E., Weitzman, C. C., Bauer, N. S., Childers Jr, D. O., Levine, J. M., … Voigt, R. G. (2020). Promoting Optimal Development: Identifying Infants and Young Children With Developmental Disorders Through Developmental Surveillance and Screening. *Pediatrics*, *145*(1), e20193449. https://doi.org/10.1542/peds.2019-3449

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule: ADOS-2*. Western Psychological Services Los Angeles, CA.

MacLachlan, M. (2021). Commentary: Challenges and opportunities in autism assessment – a commentary on Kanne and Bishop (2020). *Journal of Child Psychology and Psychiatry*, *62*(2), 146–148. https://doi.org/10.1111/jcpp.13360

Maenner, M. J., Shaw, K. A., Bakian, A. V., Bilder, D. A., Durkin, M. S., Esler, A., ... & Cogswell, M. E. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *MMWR Surveillance Summaries*, *70*(11), 1-16. https://doi.org/10.15585/mmwr.ss7011a1

Monteiro, S. A., Dempsey, J., Berry, L. N., Voigt, R. G., & Goin-Kochel, R. P. (2019). Screening and Referral Practices for Autism Spectrum Disorder in Primary Pediatric Care. *Pediatrics*, *144*(4), e20183326. https://doi.org/10.1542/peds.2018-3326

Mullen, E. M. (1995). *Mullen scales of early learning*. AGS Circle Pines, MN.

Penner, M., Senman, L., Andoni, L., Dupuis, A., Anagnostou, E., Kao, S., Solish, A., Shouldice, M., Ferguson, G., & Brian, J. (2023). Concordance of Diagnosis of Autism Spectrum Disorder Made by Pediatricians vs a Multidisciplinary Specialist Team. *JAMA Network Open, 6*(1), e2252879. https://doi.org/10.1001/jamanetworkopen.2022.52879

Pierce, K., Gazestani, V., Bacon, E., Courchesne, E., Cheng, A., Barnes, C. C., Nalabolu, S., Cha, D., Arias, S., Lopez, L., Pham, C., Gaines, K., Gyurjyan, G., Cook-Clark, T., & Karins, K. (2021). Get SET Early to Identify and Treatment Refer Autism Spectrum Disorder at 1 Year and Discover Factors That Influence Early Diagnosis. *The Journal of Pediatrics*, *236*, 179–188. https://doi.org/10.1016/j.jpeds.2021.04.041

Richards, M., Mossey, J., & Robins, D. L. (2016). Parents' concerns as they relate to their child's development and later diagnosis of autism spectrum disorder. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *37*(7), 532.

Sasson, N. J., Faso, D. J., Nugent, J., Lovell, S., Kennedy, D. P., & Grossman, R. B. (2017). Neurotypical Peers are Less Willing to Interact with Those with Autism based on Thin Slice Judgments. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/srep40700

Sheldrick, R. C., Merchant, S., & Perrin, E. C. (2011). Identification of Developmental-Behavioral Problems in Primary Care: A Systematic Review. *Pediatrics*, *128*(2), 356–363. https://doi.org/10.1542/peds.2010-3261

Sparrow, S. S., Cicchetti, D., & Balla, D. (2005). *Vineland Adaptive Behavior Scales, Second Edition*.

Sparrow, S. S., Saulnier, C. A., Cicchetti, D. V., & Doll, E. A. (2016). *Vineland-3: Vineland adaptive behavior scales.*

Vivanti, G., & Dissanayake, C. (2016). Outcome for children receiving the Early Start Denver Model before and after 48 months. *Journal of Autism and Developmental Disorders*, *46*(7), 2441–2449.

Wieckowski, A. T., de Marchena, A., Algur, Y., Nichols, L., Fernandes, S., Thomas, R. P., McClure, L. A., Dufek, S., Fein, D., & Adamson, L. B. (2021). The first five minutes: Initial impressions during autism spectrum disorder diagnostic evaluations in young children. *Autism Research*, *14*(9), 1923–1934.

Wieckowski, A. T., Hamner, T., Nanovic, S., Porto, K. S., Coulter, K. L., Eldeeb, S. Y., Chen, C.-M. A., Fein, D. A., Barton, M. L., & Adamson, L. B. (2021). Early and repeated screening detects autism spectrum disorder. *The Journal of Pediatrics*, *234*, 227–235.

Zwaigenbaum, L., & Warren, Z. (2021). Commentary: Embracing innovation is necessary to improve assessment and care for individuals with ASD: a reflection on Kanne and Bishop (2020). *Journal of Child Psychology and Psychiatry*, *62*(2), 143–145. https://doi.org/10.1111/jcpp.13271

Table 1

Sample Characterization Variables

		Final Diagnosis	
	Total Sample n = 252	**ASD n = 117**	**Non-ASD n = 135**
	N (%)	**N (%)**	**N (%)**
Sex			
Male	169 (67%)	89 (76%)	80 (59%)
Female	83 (33%)	28 (24%)	55 (41%)
Race			
White/Caucasian	126 (50%)	53 (45%)	73 (54%)
Black/African American	52 (21%)	22 (19%)	30 (22%)
Asian	17 (7%)	11 (9%)	6 (4%)
Bi- or multiracial	22 (9%)	13 (11%)	9 (7%)
American Indian	6 (2%)	3 (3%)	3 (2%)
Other	9 (4%)	7 (6%)	2 (2%)
Unknown	20 (8%)	8 (7%)	12 (9%)
Ethnicity			
Hispanic	73 (29%)	34 (29%)	39 (29%)
Non-Hispanic	155 (62%)	71 (61%)	84 (62%)
Unknown	24 (10%)	12 (10%)	12 (9%)
Maternal Education			
Less than high school or GED	17 (7%)	5 (4%)	12 (9%)
High school/GED	64 (25%)	37 (32%)	27 (20%)
Some college, technical or trade school	52 (21%)	21 (18%)	31 (23%)
College degree	66 (26%)	32 (27%)	34 (25%)
Advanced degree	49 (19%)	22 (19%)	27 (20%)
Unknown	4 (2%)	0 (0%)	4 (3%)
	M (SD)	***M (SD)***	***M (SD)***
Age in months	21.19 (5.96)	22.97 (6.31)	19.65 (5.18)
ADOS-2 CSS Total[1] (n = 252)	4.86 (3.14)	7.77 (1.88)	2.34 (1.29)
VABS (n = 240)[2]	83.92 (14.00)	76.77 (12.55)	90.72 (11.77)
MSEL (n = 248)[2]	74.63 (19.24)	66.22 (16.98)	81.89 (18.13)
Clinician Confidence in Initial Impression[3] (n = 240)	3.21 (1.14)	2.88 (1.14)	3.52 (1.06)

[1] CSS ranges from 1-10 with higher values indicating greater severity.
[2] VABS: Vineland Adaptive Behavior Scales, 2nd or 3rd Edition; MSEL: Mullen Scales of Early Learning. VABS and MSEL scores are standard scores, with a mean of 100 and standard deviation of 15.
[3] Clinicians' confidence in initial impression being ASD or non-ASD ranges from 1 to 5, with higher values indicating greater confidence. Confidence scores for 8 toddlers were not available.

Figures

Figure 1. Flow chart indicating match (blue solid lines) and mismatch (red dashed line) between trainee's initial impressions and child's final diagnoses.

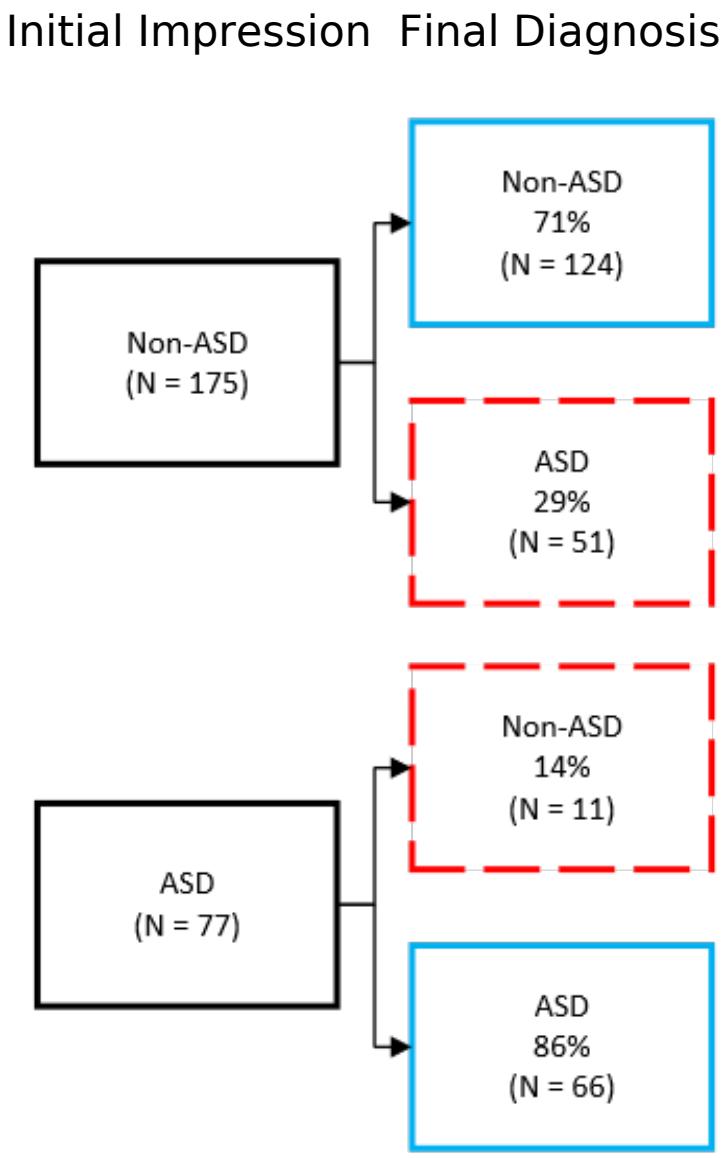

Figure 2. Mean match between initial impression and final diagnosis as a function of trainee's confidence in initial impression. Higher confidence scores indicate greater confidence (1 = 'not very confident;' 3 = 'confident;' 5 = 'extremely confident).