Title
Measuring Developmental Delays: Comparison of Parent Report and Direct Testing.

Authors
Ozonoff, Sally
Gangi, Devon
Corona, Laura
et al.

Publication Date
2024-02-01

DOI
10.1007/s10803-024-06292-8

Peer reviewed

Brief Report

Measuring Developmental Delays: Comparison of Parent Report and Direct Testing

Sally Ozonoff,[1] Devon Gangi,[1] Laura Corona,[2] Tori Foster,[2] Monique Moore Hill,[1] Makayla Honaker,[2] Shyeena Maqbool,[1] Rachel Ni,[1] Amy Nicholson,[2] Chandni Parikh,[1] Caitlin Stone,[2] Anna Kathleen Spitler,[2] Amy Swanson,[2] Alison Vehorn,[2] Liliana Wagner,[2] Amy Weitlauf,[2] & Zachary Warren[2]

[1] MIND Institute, Department of Psychiatry and Behavioral Sciences, University of California-Davis, Sacramento CA, USA

[2] Treatment and Research Institute for Autism Spectrum Disorders (TRIAD), Department of Pediatrics, Vanderbilt University Medical Center, Nashville TN, USA

Corresponding author: Sally Ozonoff PhD, MIND Institute, 2825 50th Street, Sacramento CA 95817, USA. sozonoff@ucdavis.edu, ORCID 0000-0002-1230-0794

Running Head: Comparison of Mullen and Developmental Profile scores

Declarations: SO reports funding from National Institutes of Health and Autism Science Foundation, travel reimbursements and honoraria from Autism Speaks and Autism Science Foundation and book royalties from Guilford Press. ZW reports funding from Centers for Disease Control and Prevention, Department of Defense, Health Resources and Service Administration, National Institutes of Health, and National Science Foundation. Other authors declare that they have no relevant financial or non-financial interests to disclose.

Acknowledgments: We are grateful to the families for generous participation in these studies. We thank our partners at Far Northern Regional Center and Valley Mountain Regional Center (California) for referring participants. This project was supported by National Institute of Mental Health (R01 MH127228), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (U54 HD08321; P50 HD103256). Support was also provided by the Vanderbilt Institute for Clinical and Translational Research (VICTR) which is funded by the National Center for Advancing Translational Sciences (NCATS) Clinical Translational Science Award (CTSA) Program, Award Number 5UL1TR002243-03.

Abstract

Purpose: Developmental assessment is part of a comprehensive autism evaluation. During in-person evaluations, developmental assessment is completed via direct testing by an examiner. In telehealth evaluations, developmental assessment relies on caregiver-report instruments. This study examined correspondence between caregiver report and direct testing of developmental skills.

Methods: Participants were 93 children, aged 18-42 months, undergoing evaluation for possible autism spectrum disorder (ASD). Caregivers were interviewed with the Developmental Profile, 4th edition (DP-4) via telehealth platform and children were tested in person 2-4 weeks later using the Mullen Scales of Early Learning (MSEL).

Results: Correlations between the DP-4 and MSEL were high (ranging from .50 to .82) across standard scores, age equivalents, and functional categories, as well as across individual subtests and overall composite scores.

Conclusion: The high convergent validity found in this study suggests that the DP-4 provides a suitable proxy for direct developmental testing using the MSEL in the context of telehealth evaluations for ASD in young children, delivering a good estimate of both developmental functioning and presence of delays.

Keywords: autism; developmental delay; measurement; convergent validity

Data were obtained from registered clinical trial NCT05047224, date of registration 2021-09-07

Introduction

Developmental assessment is an essential component of a comprehensive autism evaluation. During traditional in-person evaluations, developmental assessment is completed via direct testing by an examiner. In telehealth evaluations, developmental assessment relies on caregiver-report instruments. Caregiver-mediated telehealth evaluation of autism spectrum disorder (ASD) increased during the COVID-19 pandemic (Jang et al., 2022; Wagner et al., 2022). It continues to be used, given its ability to provide access to diagnostic services for families who live in rural communities (Marcin et al., 2016), alleviate travel burdens, and reduce waitlists for evaluations (Zwaigenbaum & Warren, 2021; Zwaigenbaum et al., 2021).

Telehealth instruments that measure autism symptoms have been developed, such as the TELE-ASD-PEDS (TAP; Corona et al., 2021; Wagner et al., 2021) and recent studies have demonstrated strong diagnostic correspondence with directly administered tests such as the Autism Diagnostic Observation Schedule (Corona et al., 2023; Warren et al., 2023), particularly for children 3 years or younger. In contrast, telehealth options for broader developmental and cognitive assessment "have been particularly lacking" (Holtman, Winans & Hoch, 2022, p. 5115) and most studies have relied upon caregiver-report instruments (e.g., Ages & Stages Questionnaire, Developmental Profile). Only one study, to our knowledge, has compared directly administered and caregiver-report instruments that measure broader developmental skills and delays in children being evaluated for ASD. Girard et al. (~~2021~~2022) compared the Mullen Scales of Early Learning to the Vineland Adaptive Behavior Scales in 64 autistic and 73 neurotypical children, aged 28-69 months. They found statistically significant correlations between the global standard scores provided by each measure (*r*=.55 in the autism group and *r*=.32 in the neurotypical group). Understanding how well other caregiver report instruments correspond with traditional in-person testing is crucial to validating their future use and sustainability for telehealth evaluations. To address this gap in the literature, the current study examined correspondence between caregiver report of young children's development during telehealth evaluation, using the Developmental Profile, and direct in-person testing of general developmental skills using the Mullen Scales of Early Learning within a sample of children being evaluated for ASD.

Method

As part of a larger study evaluating teleassessment methods, 93 participants (18-42 months of age, *M*=30.94 months, SD=6.53) completed first a telehealth and, 2-4 weeks later (*M*=25 days, SD=15.84), an in-person evaluation, conducted by independent examiners unaware of results of the other assessment. Participants were referred to the study due to concerns about possible ASD raised by their primary care provider, early intervention team, or positive screen on the M-CHAT-R. Parents provided informed consent prior to assessment and all methods were approved by university Institutional Review Boards in accordance with ethical standards. After both evaluations had been completed, 71 participants (76%) were diagnosed with ASD, while n=22 (24%) had non-ASD outcomes, including speech-language delay, global developmental delay, subthreshold social-communication difficulties, regulatory/behavior problems, and typical development. See Table 1 for descriptive characteristics of the sample.

Insert Table 1 about here

Measures

Developmental Profile, 4th Edition (DP-4; Alpern, 2020). This parent-report instrument measures developmental functioning in five domains: Physical, Cognitive, Communication, Social-Emotional, and Adaptive. The Caregiver Interview, which takes 20-30 minutes to complete, was administered via telehealth platform. Items were scored solely based on parent responses to the interview questions; parents were not asked to probe for behaviors or demonstrate child skills while clinicians observed. The Caregiver Interview is normed for ages one month to 21 years and provides standard scores, age equivalents (provided as a range, midpoint used for analyses), and descriptive functional categories (Delayed, Below Average, Average, Above Average, Well Above Average). The DP-4 also yields an overall General Development Score based on the 5 subscales. Reported reliability is high, with test-retest correlations ranging from .65 to .84 over a two-week interval and inter-rater correlations ranging from .60 to .92 (Alpern, 2020). The DP-4 has been used in previous studies of ASD teleassessment (Holtman et al., 2022; Jang et al., 2022).

Mullen Scales of Early Learning (MSEL; Mullen 1995). This is a standardized developmental test for children birth to 68 months. Four subscales were administered at the in-person visit: Fine Motor, Visual Reception, Expressive Language, and Receptive Language. The Gross Motor subscale was not administered since norms are provided only for children younger than 33 months and it is not used in calculation of the composite score. Raw scores were transformed, using published norms, to T scores (M=50, SD=10), age equivalents, and functional categories (Very Low, Below Average, Average, Above Average, Very High). For ease of interpretation and comparison to the DP-4, subscale T-scores were converted to standard scores

(M=100, SD=15) for analysis. The MSEL also yields an overall Early Learning Composite based on the four administered subscales. The MSEL subscales have excellent internal consistency (median 0.91) and test-retest reliability (median 0.84).

Analytic Plan

We identified four subtests that corresponded conceptually across the two measures: 1) MSEL Fine Motor and DP-4 Physical scales, 2) MSEL Visual Reception and DP-4 Cognitive scales, 3) MSEL Receptive Language and DP-4 Communication scales, and 4) MSEL Expressive Language and DP-4 Communication scales. Relationships among these subtests were examined in three ways: standard scores, age equivalents, and functional categories. Pearson correlation coefficients were computed for continuous variables and Spearman's Rho for ordinal variables (i.e., functional categories). To examine whether relationships remained when controlling for the effect of ASD, we used regression models that included a dichotomous ASD outcome variable. Finally, we compared performance on 14 items measuring similar developmental skills across instruments (e.g., identifying colors, cutting with scissors, following directions; see Table 2). We created a *similar item composite* by summing pass/fail scores on these 14 items for each measure. We calculated the correlation of the composites, as well as percent agreement for each item.

Insert Table 2 about here

Results

First, we examined correlations between standard scores of corresponding DP-4 and MSEL subtests (Table 3 and Figure 1). All correlations were statistically significant and of large effect size (Cohen, 1988), *r*s>.5, *p*s<.001. Findings were similar accounting for ASD diagnosis, *β*s>.4, *p*s<.001.

Insert Table 3 and Figure 1 about here

We then examined whether participants obtained similar age equivalent scores between the MSEL and DP-4 (Table 3 and Figure 2). Age equivalents of DP-4 subtests were positively correlated with corresponding MSEL subtests, *r*s>.6, *p*s<.001. Findings were similar accounting for ASD diagnosis, all *β*s>.5, *p*s<.001.

Insert Figure 2 about here

Next, we examined categorical measures of functioning (Delayed/Very Low, Below Average, Average, Above Average, Well Above Average/Very High) across measures (Table 3 and Figure 3) using Spearman's rho for rank order data. Functional categories of DP-4 subtests were positively correlated with corresponding MSEL subtests, *r*s ≥ .5, *p*s<.001. Of particular interest was the frequency of misclassifications between measures, such that a child was classified in the average or above average range by parent report on the DP-4 but in the below average range or lower by direct testing with the MSEL. As can be seen in Figure 3, the percentages of these misclassifications were low: 17% for the DP-4 Physical/MSEL Fine Motor, 12% for the DP-4 Cognitive/MSEL Visual Reception, 10% for the DP-4 Communication/MSEL Receptive Language, and 11% for the DP-4 Communication/MSEL Expressive Language scales.

Comparison of Mullen and Developmental Profile scores

Insert Figure 3 about here

We also examined overall scores across measures. The DP-4 General Development Score was positively correlated with the MSEL Early Learning Composite, r=.71, p<.001. Findings were similar accounting for ASD diagnosis, β=.55, p<.001. The *similar items* composites from each measure were also highly positively correlated, r=.82, p<.001. Findings were similar accounting for ASD diagnosis, β=.75, p<.001. Percent agreement was relatively high for individual items, ranging from 58% to 97% (mean percent agreement = 78%), despite some variability across measures in definitions of passes for certain items (see Table 2). Percent agreement was inflated for a few items which had high fail rates on both measures, as noted in Table 2.

Discussion

Telehealth platforms were widely used for ASD evaluations during the COVID-19 pandemic (Jang et al., 2022; Wagner et al., 2022). General use (beyond COVID-19) for rural families and medically underserved regions has been encouraged to address social and medical disparities (Zwaigenbaum et al., 2021). The Society of Developmental and Behavioral Pediatrics published a policy statement (Keder et al., 2022) advocating for continued offering of telehealth services, because they not only provide more equitable access to care but also offer unique benefits such as naturalistic observation of the child in their home environment, relief from transportation burden, and reduced costs. A potential barrier to the scalability and future sustainability of ASD telehealth evaluations, however, has been the lack of a suitable counterpart to in-person testing of cognitive and developmental levels to identify developmental delays (Holtman et al., 2022).

In the current study, we found that caregiver report of a child's development using the DP-4 was highly positively correlated with direct testing using the MSEL, with all correlations of standard scores and age equivalents of large effect size (i.e., .5 or above; Cohen, 1988), statistically significant, and in the moderate-to-high range of convergent validity as defined by Abma, Rovers, and van der Wees (2016). We also examined agreement across measures in the presence of developmental delays, which is important from a service eligibility perspective. There was strong agreement in functional classifications made by the two instruments. In the communication domain, for example, only 10% of children were classified as below average by direct testing but average based on caregiver report, providing reassurance that most children who would be eligible for speech-language services based on a direct assessment would also be eligible based on caregiver report alone.

Correspondence across instruments was also high when examining composite scores. Both instruments provide an overall score (i.e., MSEL Early Learning Composite and DP-4 General Development Score), and these were highly correlated (r=.71) despite differences in the constructs measured across instruments. The DP-4's General Development Score, for example, measures not only communication, cognitive, and fine motor skills (like the MSEL's Early Learning Composite) but also gross motor, social-emotional, and adaptive skills, which are not included in the MSEL Early Learning Composite. Analysis of a composite of items we constructed that capture similar skills across the two measures also showed strong correspondence (r=.82) between the MSEL and DP-4.

There are several limitations of this study. The sample was composed largely of young children with ASD. Although we did attempt to account for this through regression analyses controlling for diagnostic outcome, more research is needed to determine if the present findings are generalizable to children for whom concerns about ASD have not been raised. The high rate of ASD in the current sample may also have led to floor effects, which could have increased cross-measure correspondence, particularly in analyses of standard scores. It is reassuring that age equivalents, which are less susceptible to floor effects due to a more complete range of low values, also demonstrated strong correspondence across instruments. Finally, the sample was 42 months and younger, so it is not yet clear how strong the correspondence of the DP-4 and MSEL is in older preschool children.

In conclusion, the strong convergent validity found in this study suggests that the DP-4 provides a suitable proxy for direct developmental testing (via the MSEL) in the context of telehealth evaluations for ASD in young children, delivering a good estimate of both dimensionally-measured developmental functioning and categorically-determined presence of delays. This is crucial for the continuation of telehealth evaluations for ASD, which are important to ensuring access to diagnostic services in underserved communities. These findings are relevant not only to telehealth assessments but also to the use of parent report measures of early development in general (e.g., administered during an in-person interview). In the future, it would be helpful for studies to examine correspondence among other directly administered developmental tests (e.g., Bayley Scales of Infant and Toddler Development) and additional caregiver report measures of development (e.g., Vineland Adaptive Behavior Scales, Ages & Stages Questionnaire). Demonstration that caregiver report of development provides a reasonable analog of directly administered developmental tests could permit clinicians using telehealth platforms to identify developmental delays beyond ASD and provide intervention recommendations in specific areas of development (Girard et al., ~~2021~~2022). In the future, it may be useful to develop instruments measuring developmental function that can be administered by caregivers, with examiner guidance, via telehealth platform (see Jang et al., 2022, and Talbott et al., 2022, for such approaches). Until such measures are available, the present results suggest that the DP-4 provides a reasonable proxy for direct testing of developmental function and delays.

Declarations: All procedures involving human participants were performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Institutional Review Boards of both universities.

References

Abma, I.L., Rovers, M. & van der Wees, P.J. (2016). Appraising convergent validity of patient-reported outcome measures in systematic reviews. *BMC Research Notes, 9,* 226. https://doi: 10.1186/s13104-016-2034-2

Alpern, G.D. (2020) *Developmental Profile, 4th edition.* Los Angeles: Western Psychological Services.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Routledge Academic: New York, NY.

Corona, L.L., Wagner, L., Wade, J., Weitlauf, A.S., Hine, J., et al. (2021). Toward novel tools for autism identification: Fusing computational and clinical expertise. *Journal of Autism and Developmental Disorders, 51,* 4003-4012. https://doi: 10.1007/s10803-020-04857-x

Corona, L.L., Wagner, L., Hooper, M., Weitlauf, A., Foster, T.E., et al. (2023). A randomized trial of the accuracy of novel telehealth instruments for the assessment of autism in toddlers. *Journal of Autism and Developmental Disorders, online ahead of print.* https://doi: 10.1007/ s10803-023-05908-9

Girard, D., Courchesne, V., Degre-Pelletier, J., Letendre, C. & Soulieres, I. (2022). Assessing global developmental delay across instruments in minimally verbal preschool autistic children: The importance of a multi-method and multi-informant approach. *Autism Research, 15,* 103-116. https://doi: 10.1002/aur.2630

Holtman, S.J., Winans, K.S., & Hoch, J.D. (2022). Utility of diagnostic classification for children 0-5 to assess features of autism: Comparing in-person and COVID-19 telehealth evaluations. *Journal of Autism and Developmental Disorders, 52,* 5114-5125. https://doi: 10.1007/s10803-022-05606-y

Jang, J., White, S.P., Esler, A.N., Kim, S.H., Klaiman, C., et al. (2022). Diagnostic evaluations of autism spectrum disorder during the COVID-19 pandemic. *Journal of Autism and Developmental Disorders, 52,* 962-973. https://doi: 10.1007/s10803-021-04960-7

Keder, R.D., Mittal, S., Stringer, K., Wallis, K.E., Wallace, J.E. & Soares, N.S. (2022). Society for Developmental and Behavioral Pediatrics position statement on telehealth. *Journal of Developmental and Behavioral Pediatrics, 43,* 55-59. https://doi:10.1097/DBP.0000000000001046

Marcin, J.P., Shaikh, U., & Steinhorn, R.H. (2016). Addressing health disparities in rural communities using telehealth. *Pediatric Research, 79,* 168-176. https://doi: 10.1038/pr.2015.192

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. Circle Pines MN, AGS.

Talbott, M.R., Dufek, S., Young, G.S., & Rogers, S.J. (2022). Leveraging telehealth to evaluate infants with prodromal autism spectrum disorder characteristics using the Telehealth Evaluation of Development for Infants (TEDI). *Autism, 26,* 1242-1254. https://doi:10.1177/13623613211045596

Wagner, L., Corona, L.L., Weitlauf, A.S., Marsh, K.L., Berman, A.F., et al. (2021). Use of the TELE-ASD-PEDS for autism evaluations in response to COVID-19: Preliminary outcomes and clinician acceptability. *Journal of Autism and Developmental Disorders, 51,* 3063-3072. https://doi: 10.1007/s10803-020-04767-y

Wagner, L., Weitlauf, A.S., Hine, J., Corona, L.L., Berman, A.F., et al. (2022). Transitioning to telemedicine during COVID-19: Impact on perceptions and use of telemedicine procedures for the diagnosis of autism in toddlers. *Journal of Autism and Developmental Disorders, 52,* 2247-2257. https://doi: 10.1007/s10803-021-05112-7

Warren, Z., Corona, L., Gangi, D., & Ozonoff, S. (2023). Diagnostic accuracy and acceptability of the TELE-ASD-PEDS (TAP): Preliminary findings from a multi-site study of in-home use. Paper presented at the Society of Developmental and Behavioral Pediatrics, Minneapolis, September.

Zwaigenbaum, L., Bishop, S., Stone, W.L., Ibanez, L., Halladay, A., et al. (2021). Rethinking autism spectrum disorder assessment for children during COVID-19 and beyond. *Autism Research, 14,* 103-116. https://doi: 10.1002/aur.2615

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,* 143-145. https://doi: 10.1111/jcpp.13271

Table 1. Sample characteristics by diagnostic outcome group.

	ASD	No ASD
Sex	*n (%)*	*n (%)*
Male	43 (61%)	11 (50%)
Female	28 (39%)	11 (50%)
Ethnicity		
Hispanic/Latino	15 (21%)	5 (23%)
Non-Hispanic/Latino	55 (78%)	17 (77%)
Prefer not to report	1 (1%)	0 (0%)
Race		
American Indian or Alaska Native	3 (4%)	0 (0%)
Asian	4 (6%)	4 (18%)
Black or African-American	8 (11%)	0 (0%)
Native Hawaiian or Other Pacific Islander	2 (3%)	0 (0%)
White	43 (61%)	16 (72%)
More than one race	7 (10%)	1 (5%)
Prefer not to report	4 (5%)	1 (5%)
Maternal Education		
High school or less	21 (30%)	5 (23%)
Some college	19 (27%)	8 (36%)
Bachelor's degree	18 (25%)	6 (27%)
Graduate degree	13 (18%)	3 (14%)
Annual Household Income		
< $40,000	28 (40%)	5 (23%)
$40,001-100,000	20 (28%)	9 (41%)
$100,001-200,000	13 (18%)	3 (14%)
>$200,000	3 (4%)	3 (14%)
Prefer not to report	7 (10%)	2 (8%)
Visit Characteristics	*M (SD)*	*M (SD)*
Age at Visit 1 (months)	31.10 (6.42)	30.41 (6.98)
Days between visits	23.59 (14.32)	28.32 (19.96)
DP-4 Standard Scores		
Physical	77.28 (13.92)	88.23 (17.90)
Cognitive	74.93 (16.37)	86.14 (13.63)
Communication	66.92 (18.94)	90.36 (18.81)
Adaptive Behavior	74.51 (13.96)	89.23 (16.87)
Social-Emotional	66.51 (11.47)	85.77 (14.70)
General Development Score	68.14 (12.42)	84.71 (15.02)
MSEL T-Scores		
Visual Reception	32.23 (12.30)	44.18 (14.98)
Fine Motor	39.30 (11.05)	44.27 (15.93)

Receptive Language	25.90 (9.09)	44.14 (11.92)
Expressive Language	27.58 (9.16)	41.00 (12.78)
Early Learning Composite	62.41 (14.02)	89.52 (21.32)

Table 2. Corresponding items from Developmental Profile and Mullen Scales of Early Learning used in *similar item* composites.

DP-4 (Subscale and Item Number)	MSEL (Subscale and Item Number)	Percent Agreement
Makes a circular hand movement to copy a circle (P 18)	Copies circle (FM 22)	67%
Copies an up and down straight line (not scribble) (P 21)	Imitates crayon line, vertical or horizontal (FM 15)	58%
Copies two intersecting lines to make a cross or "X" (P 23)	Copies shapes and letters (FM 30)	97% [a]
Uses child-safe scissors with one hand to cut (P 25)	Cuts with scissors (FM 24)	95% [b]
Uses pencil/crayon to mark any writing surface (Cog 10)	Imitates crayon lines, any direction (FM 15)	90%
Points/looks at object in book when named (Cog 11)	Identifies pictures (RL 21)	78%
Identifies at least two colors (Cog 14)	Identifies colors (RL 27)	67%
Counts up to 6 objects (Cog 16)	Has concept of 6 (RL 32)	76% [c]
Sorts things by color, shape, or size (Cog 17)	Matches by shape or by size/color (VR 20 *or* 22)	77%
Clearly understands the meaning of "no" (Comm 9)	Understands inhibitory words (RL 11)	87%
Follows instructions of three or more words (Comm 11)	Follows directions (RL 17)	69%
Uses words or signs to express a want or need (Comm 12)	Says first words (EL 11)	82%
Follows two-step verbal instructions (Comm 13)	Follows related or unrelated commands (RL 20 *or* 25)	73%

Note. P = DP-4 Physical scale, Cog = DP-4 Cognitive scale, Comm = DP-4 Communication scale, FM = MSEL Fine Motor subtest, RL = MSEL Receptive Language subtest, VR = MSEL Visual Reception subtest, EL = MSEL Expressive Language subtest.
The following items were rarely passed (<10%), potentially inflating percent agreement: [a] only 4 children (4%) passed DP-4 item and 1 child (1%) passed MSEL item; [b] only 6 children (7%) passed DP-4 item and 8 children (9%) passed MSEL item; [c] only 1 child (1%) passed MSEL item.

Table 3. Correlations between DP-4 and MSEL subtests.

	MSEL Fine Motor	MSEL Visual Reception	MSEL Receptive Language	MSEL Expressive Language
Standard Scores				
DP-4 Physical	.52*	-	-	-
DP-4 Cognitive	-	.57*	-	-
DP-4 Communication	-	-	.64*	.76*
Age Equivalents				
DP-4 Physical	.63*	-	-	-
DP-4 Cognitive	-	.68*	-	-
DP-4 Communication	-	-	.75*	.80*
Functional Categories+				
DP-4 Physical	.50*	-	-	-
DP-4 Cognitive	-	.53*	-	-
DP-4 Communication	-	-	.53*	.67*

Note. *p<.001; +Spearman's rho

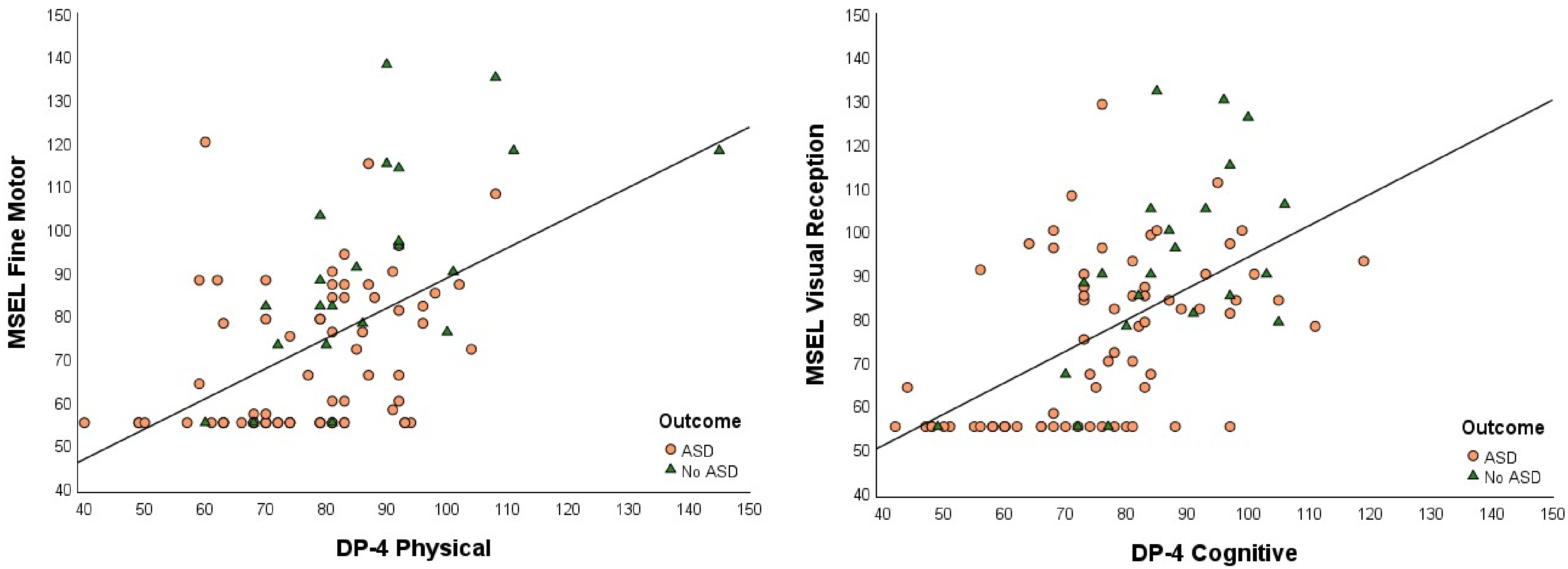
Figure Captions

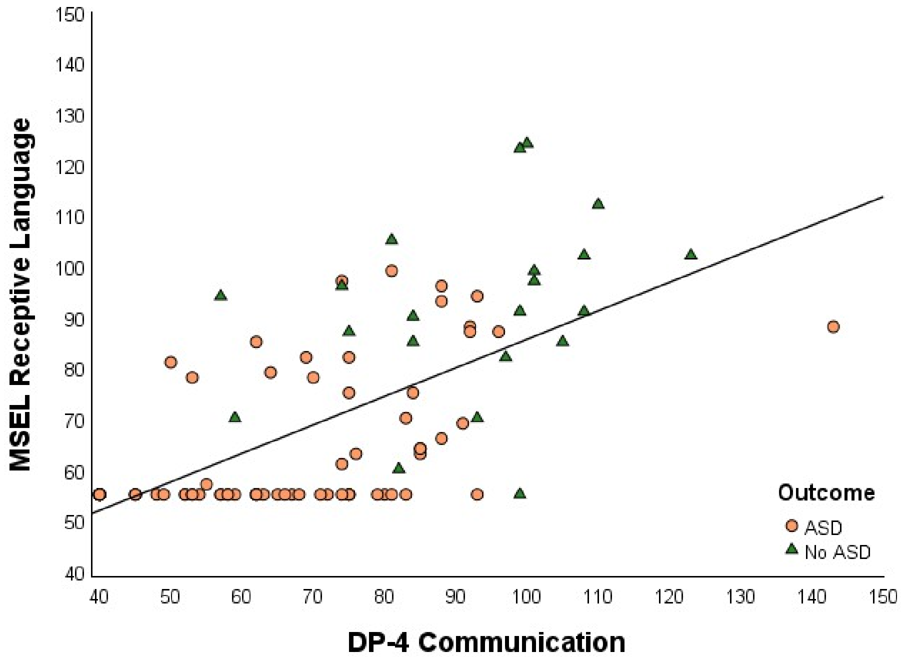
Fig 1 Scatterplots of DP-4 and MSEL standard scores

Fig 2 Scatterplots of DP-4 and MSEL age equivalents (in months)

Fig 3 Scatterplots of DP-4 and MSEL functional categories

Fig 1 Scatterplots of DP-4 and MSEL standard scores




MSEL Receptive Language
DP-4 Communication
Outcome
ASD
No ASD

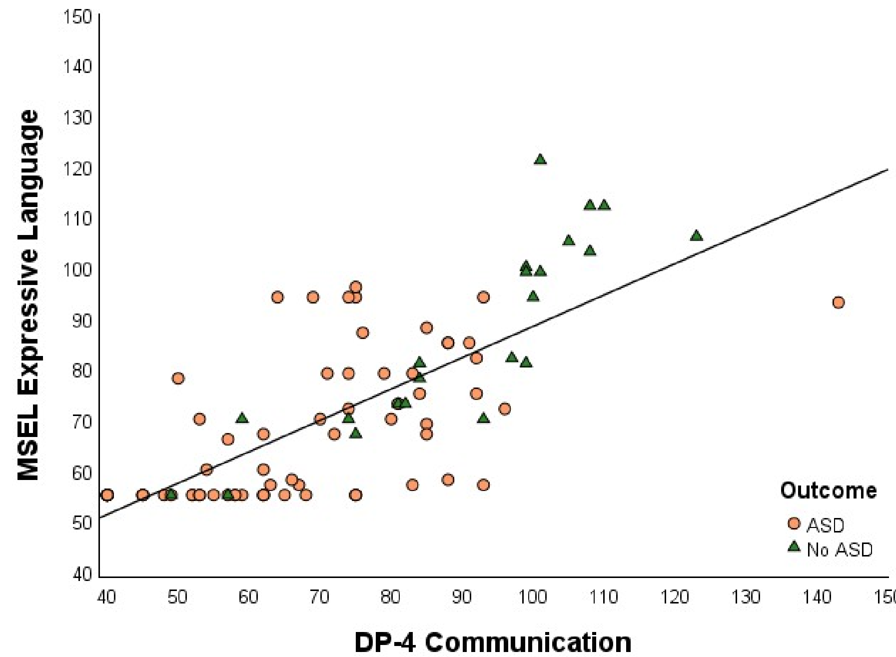
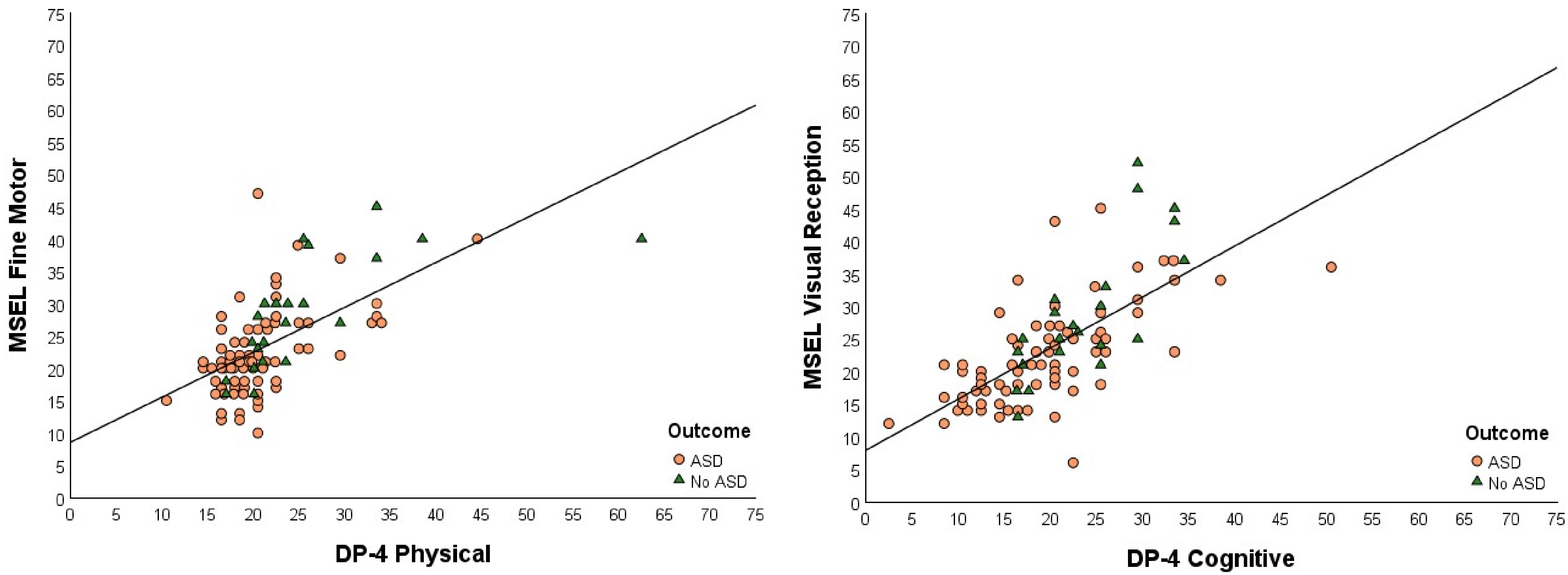

MSEL Expressive Language
DP-4 Communication
Outcome
ASD
No ASD

Fig 2 Scatterplots of DP-4 and MSEL age equivalents (in months)



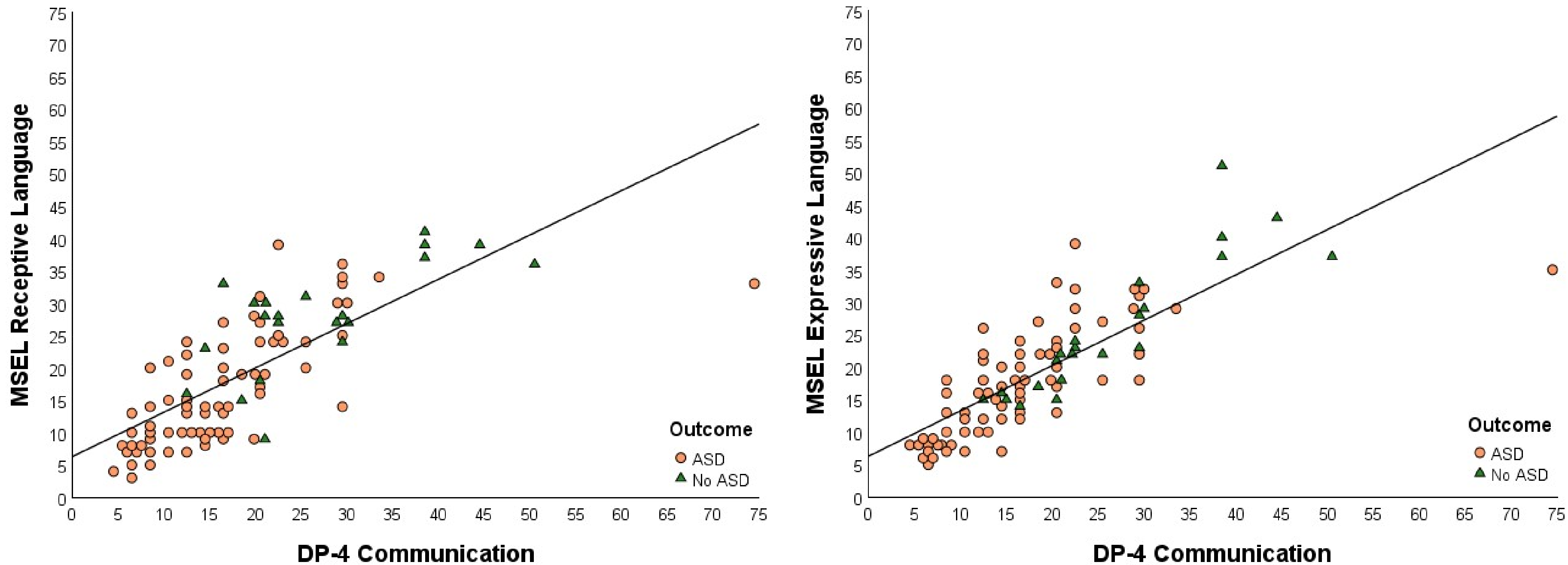
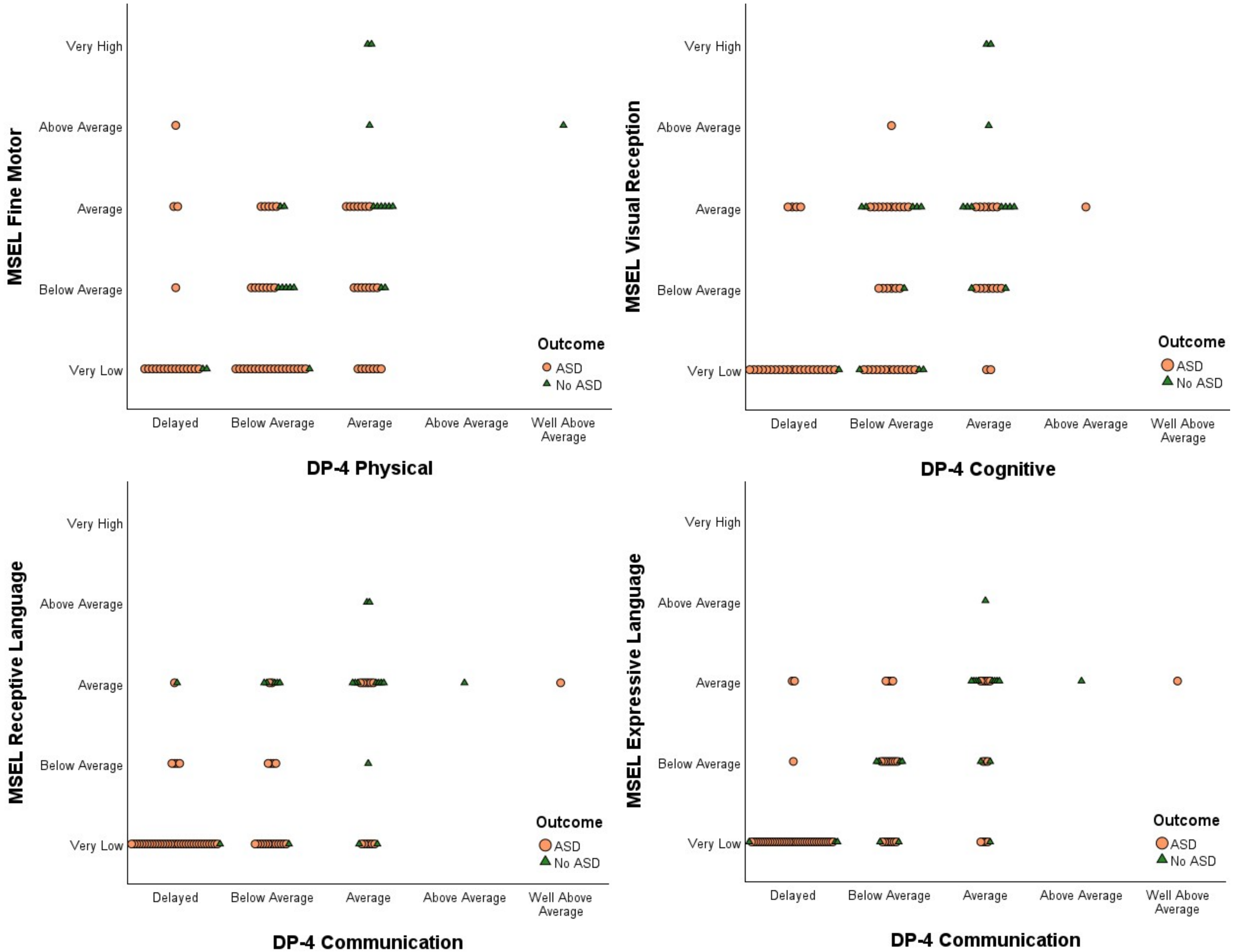


Fig 3 Scatterplots of DP-4 and MSEL functional categories


MSEL Fine Motor
Very High
Above Average
Average
Below Average
Very Low
Delayed
Below Average
Average
Above Average
Well Above Average
DP-4 Physical
Outcome
ASD
No ASD
MSEL Visual Reception
DP-4 Cognitive
MSEL Receptive Language
DP-4 Communication
MSEL Expressive Language
DP-4 Communication