**Supplemental digital content**

**Assessments**

The most common tests administered to the children at baseline were: the Bayley Scales of Infant Development, 2nd Edition (Bayley, 1993; 73% of cognition assessments); the Reynell Developmental Language Scales (Reynell & Gruber, 1990; 77% of receptive and expressive language assessments); and the Vineland Adaptive Behavior Scales: Interview Edition (Sparrow, Balla, & Cicchetti, 1984; 94% of adaptive behavior assessments). Common final assessments were: the Wechsler Intelligence Scale for Children, 3rd and 4th Editions (Wechsler, 1991, 2003; 56% of cognition assessments); the Wechsler Preschool and Primary Scale of Intelligence, Revised and 3rd Editions (Wechsler, 1989, 2002; 35% of cognition assessments); the Receptive One-Word Picture Vocabulary Test and Expressive One-Word Picture Vocabulary Test (Brownell, 2000a, 2000b; 82% of receptive and expressive language assessments); and the Vineland Adaptive Behavior Scales: Interview Edition (Sparrow et al., 1984; all adaptive behavior assessments). Intervals between the baseline and the final assessment for each domain are summarized in Supplemental Table 1.

**Supplemental Table 1. Mean (*SD*) Months between Baseline and Final Assessments**

|  |  |  |
| --- | --- | --- |
| Domain(s) | Eclectic Treatment | ABA Treatment |
| Cognitive | 52.65 (16.69) | 50.74 (23.96) |
| Expressive, receptive | 49.23 (17.53) | 51.20 (22.67) |
| Adaptive composite | 51.96 (16.88) | 52.59 (23.37) |
| Communication, self-help, social | 53.03 (16.40) | 50.09 (24.00) |
| Motor | 26.93 (8.74) | 30.62 (8.25) |

**Analyses of standard scores and developmental quotients**

The analyses in the main article dichotomized baseline and final assessments for each child as normal (standard score [SS] ≥ 85 or developmental quotient [DQ] ≥ 85) or subnormal (SS < 85 or DQ < 85). Statistical significance was determined using Fisher’s exact test, which is designed for dichotomous variables. This approach was adopted because families of children with autism have an implicit understanding of what it means for a child to be “normal,” and so should find analyses of the dichotomized outcomes meaningful. However, several sets of analogous statistical analyses were also conducted using the original standard scores and developmental quotients. Parents and guardians who lack familiarity with these measures will find such analyses opaque; however, they have greater statistical power than analyses designed for dichotomous variables. Thus, the results obtained from analyses of SS and DQ scores are reported below.

One set of analyses used paired t-tests to compare the average (mean) baseline SS or DQ score of each domain with the average final SS or DQ score for that domain, to determine whether treatment typically results in statistically significant gains or declines. Separate t-tests were conducted for each treatment group. The tests revealed that children in the eclectic group exhibited statistically significant gains in their average cognitive, expressive communication, and receptive communication scores following treatment (Supplemental Table 2). However, these gains were accompanied by statistically significant *declines* in the average adaptive composite and social domain scores. Average scores declined as well for the communication, self-help, and motor domains following eclectic treatment, although these declines were not statistically significant (all *p* > .10). A different pattern was evident in the ABA treatment group. For these children, the average scores for all but one domain exhibited statistically significant gains following treatment. The sole exception was the motor domain, which exhibited a small, nonsignificant decline (*p* = .836).

**Supplemental Table 2. Baseline and Final Scores (*M* ± *SD*)**

|  |  |  |
| --- | --- | --- |
| Domain | Eclectic Treatment | ABA Treatment |
| Baseline Score | FinalScore | BaselineScore | FinalScore |
| Cognitive | 57.23 ± 14.62 | 70.35 ± 24.10\*\* | 64.45 ± 17.25 | 90.60 ± 22.28\*\*\* |
| Expressive | 46.86 ± 11.15 | 63.68 ± 27.47\*\* | 51.24 ± 12.78 | 86.36 ± 26.45\*\*\* |
| Receptive | 46.97 ± 14.51 | 61.27 ± 26.35\*\* | 50.35 ± 16.47 | 81.91 ± 26.38\*\*\* |
| Adaptive comp. | 69.86 ± 9.65 | 60.93 ± 18.63\* | 71.75 ± 7.50 | 78.57 ± 15.33\*\* |
|  Self-help | 70.06 ± 10.55 | 64.90 ± 19.16 | 70.72 ± 11.73 | 82.91 ± 17.47\*\*\* |
|  Social | 75.06 ± 11.92 | 67.35 ± 18.44\* | 73.48 ± 11.05 | 81.37 ± 15.48\*\* |
|  Comm. | 64.61 ± 8.91 | 63.32 ± 20.17 | 65.32 ± 10.35 | 81.30 ± 16.90\*\*\* |
|  Motor | 90.37 ± 11.70 | 84.80 ± 15.87 | 93.76 ± 11.95 | 93.24 ± 12.47 |

\* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001 for t-test comparing baseline and final means

Next, analyses of covariance (ANCOVAs) were conducted to compare the relative efficacy of the two treatment types, with each child’s change score–the gain or decline in the SS or DQ score from baseline to the final assessment–serving as the dependent measure. Covariates were the baseline SS or DQ score, the child’s age at diagnosis, and the parents’ mean years of education. The ANCOVAs found that, after controlling for covariates, ABA treatment resulted in significantly larger average improvements than eclectic treatment in all domains other than the motor domain, where ABA treatment resulted in significantly smaller *losses* than eclectic treatment (Supplemental Table 3). Taken together, the t-tests and ANCOVAs suggest that all domains responded positively to ABA treatment, while some domains also responded–but less strongly–to eclectic treatment.

Despite this general pattern across all domains, some domains demonstrated larger gains than others. This was confirmed by a 2 (domain: cognitive or adaptive composite) by 2 (treatment type: ABA or eclectic) by 2 (time: baseline or final) repeated measures analysis of variance (ANOVA). This ANOVA tested three main effects, three two-way interactions, and one three-way interaction. However, only three of these terms were of particular interest: the treatment by time interaction, which indicated if scores changed more following one type of treatment than the other (similar to the ANCOVAs described above); the domain by time interaction, which indicated whether one domain was more responsive to treatment than the other; and the three-way interaction, which indicated whether any such difference in domain responsiveness varied, depending upon the type of treatment.

**Supplemental Table 3. Changes (*M* ± *SD*) from Baseline to Final Assessment**

|  |  |  |
| --- | --- | --- |
| Domain | Eclectic Treatment | ABA Treatment |
| Cognitive | 13.13 ± 22.47 | 26.15 ± 22.74\*\* |
| Expressive | 16.82 ± 25.40 | 34.89 ± 27.58\*\* |
| Receptive | 14.30 ± 21.64 | 31.57 ± 24.97\*\* |
| Adaptive composite |  -8.93 ± 19.33 | 6.82 ± 14.93\*\*\* |
|  Self-help |  -5.16 ± 21.71 | 12.19 ± 17.54\*\*\* |
|  Social |  -7.71 ± 19.09 | 7.89 ± 16.55\*\*\* |
|  Communication |  -1.29 ± 18.91 | 15.98 ± 17.70\*\*\* |
|  Motor |  -5.57 ± 19.35 | -0.52 ± 17.03\* |

\* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001 for ANCOVA comparing eclectic and ABA mean change scores, after controlling for differences in domain baseline scores

The treatment by time interaction in the repeated measures ANOVA confirmed that ABA yielded larger treatment gains than eclectic treatment ( = .15, *p* = .001). The domain by time interaction was also statistically significant ( = .47, *p* < .001), indicating that treatment had a larger impact on cognition than on adaptive skills. This is evident from Figure 2 in the main article, which shows that cognitive scores improved more than adaptive composite scores: At the final assessment, more children in both treatment groups were in the normal range for cognition (to the right of the vertical line in the right panel) than for adaptive skills (above the horizontal line in the right panel). However, the three-way interaction was not significant ( = .003, *p* = .67). Thus, even though gains were greater from ABA than from eclectic treatment, and even though cognitive gains were greater than adaptive composite gains, ABA treatment was no more effective than eclectic treatment in narrowing the difference between cognitive and adaptive skills. In fact, this difference grew in both groups following treatment: Adaptive composite scores were an average of 7-13 points *higher* than cognitive scores at baseline, but 9-13 points *lower* at the final assessment (Supplemental Table 4).

**Supplemental Table 4. Baseline and Final Adaptive Composite and Cognitive Scores**

|  |  |  |
| --- | --- | --- |
| Treatment Group | Mean (*SD*) Baseline Score | Mean (*SD*) Final Score |
| Adaptive Composite | Cognitive | Difference | Adaptive Composite | Cognitive | Difference |
| ABA(*n* = 44) | 71.75(7.50) | 64.68(17.25) | 7.07(16.20) | 78.57(15.33) | 91.09(21.99) | -12.52(15.38) |
| Eclectic(*n* = 28) | 69.86(9.65) | 56.96(15.29) | 12.89(19.13) | 60.93(18.63) | 69.93(25.36) | -9.00(12.64) |
| Overall(*N* = 72) | 71.01(8.39) | 61.68(16.84) | 9.33(17.50) | 71.71(18.69) | 82.86(25.40) | -11.15(14.38) |

*Note*. Excludes children who had unknown cognitive or adaptive composite scores at the baseline or final assessment.

**Supplemental references**

Bayley, N. (1993). *Bayley Scales of Infant Development, second edition.* San Antonio, TX: The Psychological Corporation.

Brownell, R. (Ed.). (2000a). *Expressive One-Word Picture Vocabulary Test—2000* *(EOWPVT-2000)*. Novato, CA: Academic Therapy Publications.

Brownell, R. (Ed.). (2000b). *Receptive One-Word Picture Vocabulary Test—2000* *(ROWPVT-2000)*. Novato, CA: Academic Therapy Publications.

Reynell, J. K. & Gruber, G. P. (1990*). Reynell Developmental Language Scales*. Los Angeles, CA: Western Psychological Services.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales (interview edition).* Circle Pines, MN: American Guidance Service.

Wechsler, D. (1989). *Wechsler Preschool and Primary Scale of Intelligence—Revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1991). *The Wechsler Intelligence Scale for Children—third edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence™—third edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2003). *The Wechsler Intelligence Scale for Children®—fourth edition*. London: Pearson.