Published in final edited form as:
J Consult Clin Psychol. 2013 June ; 81(3): 566–572. doi:10.1037/a0032003.

A Randomized Controlled Trial of COMPASS Web-Based and Face-to-Face Teacher Coaching in Autism

Lisa A. Ruble,
Department of Educational, School, and Counseling Psychology, University of Kentucky

John H. McGrew,
Department of Psychology, Indiana University–Purdue University at Indianapolis

Michael D. Toland,
Department of Educational, School, and Counseling Psychology, University of Kentucky

Nancy J. Dalrymple, and
Bloomington, Indiana

Lee Ann Jung
Early Childhood, Special Education, and Rehabilitation Counseling, University of Kentucky

Abstract

Objective—Most children with autism rely on schools as their primary source of intervention, yet research has suggested that teachers rarely use evidence-based practices. To address the need for improved educational outcomes, a previously tested consultation intervention called the Collaborative Model for Promoting Competence and Success (COMPASS; Ruble, Dalrymple, & McGrew, 2010; Ruble, Dalrymple, & McGrew, 2012) was evaluated in a 2nd randomized controlled trial, with the addition of a web-based group.

Method—Forty-nine teacher–child dyads were randomized into 1 of 3 groups: (1) a placebo control (PBO) group, (2) COMPASS followed by face-to-face (FF) coaching sessions, and (3) COMPASS followed by web-based (WEB) coaching sessions. Three individualized goals (social, communication, and independence skills) were selected for intervention for each child. The primary outcome of independent ratings of child goal attainment and several process measures (e.g., consultant and teacher fidelity) were evaluated.

Results—Using an intent-to-treat approach, findings replicated earlier results with a very large effect size ($d = 1.41$) for the FF group and a large effect size ($d = 1.12$) for the WEB group relative to the PBO group. There were no differences in overall change across goal domains between the FF and WEB groups, suggesting the efficacy of videoconferencing technology.

Conclusions—COMPASS is effective and results in improved educational outcomes for young children with autism. Videoconferencing technology, as a scalable tool, has promise for facilitating access to autism specialists and bridging the research-to-practice gap.

Keywords

autism; IEPs; RCT; COMPASS; goal attainment scaling

Correspondence concerning this article should be addressed to Lisa A. Ruble, Department of Educational, School, and Counseling Psychology, 237 Dickey Hall, University of Kentucky, Lexington, KY 40506. lisa.ruble@uky.edu.

Each day families across the United States send their children with autism spectrum disorders (ASDs) to schools expecting educators to implement current best practice. Yet, research has suggested that fewer than 5% of educators of students with ASD use evidence-based methods (Morrier, Hess, & Heflin, 2011).

Teacher consultation is a research-based approach for supporting teachers to improve student outcomes (Sheridan, Welch, & Orme, 1996). The current study tests a specific model of consultation called the Collaborative Model for Promoting Competence and Success (COMPASS; Ruble, Dalrymple, & McGrew, 2010; Ruble, Dalrymple, & McGrew, 2012). COMPASS is a manualized intervention and decision-making framework designed to target core deficit areas associated with autism based on parent and teacher priorities: social skills, communication, and independence (National Research Council, 2001). The teaching plans generated for these three core areas are tailored to the specific needs of each child. After the team develops goals and teaching plans for each skill, teachers are asked to add the goals to the individual education programs (IEPs). The initial consultation is followed by four follow-up teacher coaching sessions spaced evenly throughout the school year and designed to facilitate teacher implementation of the teaching plans. Preliminary evidence indicates that both the initial consultation and subsequent coaching are critical in producing a large between-groups effect size ($d = 1.5$; Ruble et al., 2010) when compared against services as usual. Also, at least four sessions of coaching were needed, as teacher adherence improved over time and was associated with child outcomes for the last coaching session only ($r = .60$, $p = .01$).

For this study, we examined the impact of web-based videoconferencing technology tools that offer sustainability and efficient use of resources. We were particularly interested in whether our promising initial results would replicate in a new school setting and whether web-based coaching could replace face-to-face coaching, potentially reducing consultant burden. We used a randomized pre–post experimental design to compare child outcomes for three groups: (1) teachers who received an online autism training that served as a placebo control (PBO) group; (2) teachers who received COMPASS and face-to-face (FF) teacher coaching sessions; and (3) teachers who received COMPASS and web-based (WEB) teacher coaching sessions. We expected the PBO group to serve as a control because of the research that documents the limited impact didactic training alone has on changes in teacher behavior (Joyce & Showers, 1983, 2002).

Two hypotheses were tested. Based on our prior study, the primary hypothesis was that child goal attainment will be higher for the FF and WEB groups compared to PBO group. Second, it was expected that the WEB group would show lower overall child goal attainment scores than the FF group.

Method

Teachers

Forty-nine special education teachers were recruited and randomized. One child with autism was randomly selected per teacher. One teacher was male, and all were certified. Forty-five percent had a bachelor of arts, 47% a master of arts, and 8% did not indicate the degree earned.

Children With Autism

Children met the definition of autistic disorder according to the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; American Psychiatric Association, 2000) as confirmed by the Autism Diagnostic Observation Schedule (Modules 1 or 2; Lord et al., 2000), a standard diagnostic instrument for identifying individuals with autism that has good

criterion validity, sensitivity and specificity, as well as good reliability. Children also received special services designated by the Individuals with Disabilities Education Act (2004). Children's ages ranged between 3 and 9 years, with a mean of 6 years (*SD* = 1.6). Eighty-six percent of the children were male, and 80% were White, 6% Black, 2% Asian, 6% other, and 6% unidentified. For families, 20% had incomes less than $25,000, 25% were between $25,000 and $49,999, 33% were above $49,999, and 22% did not respond. The distribution of family income was similar to family income reported in our previous randomized controlled trial (RCT; Ruble et al., 2010).

Sampling

Teachers were recruited in a multistep fashion from two mid-southern states. After permission was granted at the district level, the researchers contacted teachers directly. Between August 2009 and August 2010, a total of 180 teacher-child pairs were assessed for eligibility (see Figure 1). The sample included all possible teacher-child pairs identified as potentially meeting the eligibility criteria. For those teachers who met inclusion criteria, 44.5% participated, 12.7% did not respond, and 42.7% refused. This rate of participation was similar to our prior RCT (Ruble et al., 2010). A total of 15 PBO, 14 FF, and 15 WEB-based teacher–child dyads completed the study, but data from all 49 dyads who were initially recruited were used in the intent-to-treat (ITT) analyses. The sample size per group was sufficient to detect differences between the FF and control groups at a power of .80, assuming that our obtained effect size equaled that from our prior study (1.5). A priori, we expected an effect size half that size between the two treatment approaches, requiring a sample size of 23 per group to detect differences at a power of .80. Due to challenges in recruitment, we were able to recruit only about three fourths of the required sample size. Following the Time 1 assessment, teacher–child dyads were randomized into groups. The PBO group of teachers received online training on three evidence-based practices in autism from the Ohio Center for Autism and Low Incidence (2007).

Child Measures to Establish Sample Equivalency

Three reliable child measures (language, cognition, adaptive behavior) were administered at Time 1 to show group equivalency. Language was assessed with the Oral and Written Language Scales (OWLS; Carrow-Woolfolk, 1995). Cognitive level was evaluated using the General Conceptual Ability subscore of the Differential Abilities Scale (DAS; Elliott, 1990). Adaptive behavior was measured with the classroom edition of the Vineland Adaptive Behavior Scales (Sparrow, Cicchetti, & Balla, 2005).

Assessment of Treatment Adherence Procedures

Three adherence measures were applied: (a) consultant adherence to the initial COMPASS consultation protocol and (b) to the teacher coaching sessions and (c) teacher adherence to the implementation of the teaching plans. Copies of the instruments are available in Ruble, Dalrymple, and McGrew (2012).

Consultant adherence to the COMPASS consultation protocol—Consultant adherence was assessed with a 25-item close-ended (yes/no) checklist (e.g., the consultation included goals suggested from home and family, planning was based on input from all participants, and teaching goals included contributions from each member) completed by parents (Kuder–Richardson Formula 20 [*KR20*] = .99 in current sample) and teachers (*KR20* = .99) after consultation.

Consultant fidelity to the coaching protocol—Teachers completed a 16-item scale to measure consultant adherence to the coaching protocol. Consultant adherence to the protocol

was rated using a 4-point Likert-type scale ranging from 1 (*not at all*) to 4 (*very much*; α = .70). Example items included "we reviewed the most current teaching plan," "we evaluated the goal attainment of the child's most current level of progress," and "we discussed the environmental factors that might be helping or hindering student progress."

Teacher fidelity to teaching plans—Immediately following each coaching session, two independent consultants rated the degree to which the teacher followed the teaching plan recommendations using a 5-point Likert-type scale item ranging from 1 (*not very*) to 5 (*very much*). Estimated percentage agreement, weighted Kappa, and interclass correlations (ICC) in the current sample were .79, .80, and .90, respectively.

Psychometrically Equivalence Tested Goal Attainment Scaling (PET-GAS)

Because each child had different goals, different baseline skill levels associated with the goals, and different teaching plans, an idiographic assessment system utilizing PET-GAS was used to measure the amount of progress each student made on the three IEP goals. Goals related to the three core learning domains of communication (e.g., will independently initiate three requests during lunch), social skills (e.g., will take two turns with an object and a peer during free play), and independence (e.g., will independently complete a three-step work activity using visual supports) were identified and prioritized, and each was translated into an IEP objective. Several procedures were implemented to ensure high quality, comparability, and objective goal attainment assessment (see Ruble, McGrew, & Toland, 2012). To enhance between-groups comparability in GAS descriptions, we applied a protocol for goal writing to ensure equivalence in goal difficulty (e.g., goals were selected that were expected to be attainable by most children but not easy), measurability (e.g., use of clear behavioral descriptions including specific wording concerning duration, frequency, and needed supports), and size of rating scale interval (e.g., an equivalence chart was created for percentage accuracy, frequency, number of prompts, and level of support needed in performing behaviors; Ruble, McGrew, Dalrymple, & Jung, 2010). Detailed descriptions are provided in Ruble, Dalrymple, and McGrew (2010) and in Ruble et al. (2012). Each goal attainment scale used the following 5-point rating scale: −2 = *child's present levels of performance*, −1 = *progress*, 0 = *expected level of outcome*, +1 = *somewhat more than expected*, +2 = *much more than expected*. Half-scores were allowed when raters observed skill level between two benchmarks. A score of zero represented improvement consistent with the actual description of the written IEP objective. PET-GAS posttreatment ratings were based on direct observations from an observer unaware of group assignment. During the observation, teachers demonstrated for the independent observer each of the three targeted teaching objectives during an instructional situation. As recommended, only raw scores were used (MacKay, 1996; Schlosser, 2004). Two coders independently coded 39% of the goals. Interrater agreement as measured using the sample ICC for single measures was .82 for the social skills, .86 for communication skills, and .91 for learning skills goals.

In a second validation step, we formally measured the between-groups comparability of the goals created, applying a 3-point ordinal scale to code three features of each goal: (a) goal measurability (i.e., the degree to which the descriptions include prompt level, criterion for success, and an observable skill), (b) equidistant between-goals benchmarks (i.e., the degree to which benchmark descriptions are equilibrated and scaled appropriately), and (c) level of goal difficulty (i.e., the degree to which the present levels of performance indicate that the child is unable to perform the skill with anyone, anywhere, or with any prompts compared to what is written in the objective; see Ruble, Dalrymple, & McGrew, 2012). Two raters independently coded 20% of the GAS forms for the three features of measurability, equidistance, and difficulty. The sample ICC for single measures was 1.0 for measurability, .

92 for equidistance, and .93 for difficulty. Detailed instructions for creating GAS templates are provided in Ruble, Dalrymple, and McGrew (2012).

Intervention

The intervention consisted of a 3-hr parent–teacher consultation and four 1.5-hr coaching sessions. Parents were invited, but not required, to attend coaching sessions.

COMPASS Consultation

The initial consultations were provided by Lisa A. Ruble or Lee Ann Jung. All consultations were conducted in person at the school and occurred within the first 2 months of the start of the school year. Prior to consultation, parents and teachers completed a COMPASS assessment questionnaire, which was collected and summarized into a joint form used for discussion about the child's personal and environmental challenges and supports associated with social, communication, and independent/adaptive skills at school and home. The COMPASS consultation intervention is manualized (Ruble, Dalrymple, & McGrew, 2012) and includes details and case study examples of actual consultations. A unique feature of the consultation is its shared decision-making approach (Ruble, Birdwhistell, Toland, & McGrew, 2011) in the selection of the treatment goals and intervention plans. The child's IEP team met within 2 weeks to update the IEP so that the goals identified in the consultation and targeted for follow-up coaching were reflected in the child's program. Prior to the first coaching session, the consultant crafted the PET-GAS for each skill using the protocol described earlier. PET-GAS was used for progress monitoring at the four teacher-coaching sessions and at the final outcome assessment.

Web-Based Group

Teachers assigned to the WEB group received a 30-min technology training session prior to their first coaching session. Teachers were given a laptop computer, a webcam, headphones, and a video camera. Teachers were shown how to operate each piece of hardware and how to connect to the Adobe Connect Pro videoconference website. This software was chosen because of its ability to (a) support video and audio from the webcams, (b) share and view documents of all users at the same time, (c) view recorded video simultaneously, and (d) maintain security (users sign in using a unique username and password). Together, these features allowed consultants to conduct the same confidential, structured, and interactive interview that was used with the FF group. The real-time interaction allowed teachers and consultants to view videos of the teacher–student instructional situation jointly, pause videos to have live conversations, and review teaching plans and PET-GAS forms.

Teacher Coaching

Coaching sessions took place every 5 weeks. Two sessions occurred during the fall semester, and two during the spring semester. Similar to the initial consultation, a written protocol was developed and followed for each coaching session that included (a) observing a teacher-made videotape of instructing the child on the three targeted objectives and soliciting teacher feedback on what was observed, (b) scoring the child's progress using the GAS form, and (c) discussing the teaching plans and making any adjustments to the plans based on discussion and review of the video. The same protocol was implemented for both FF and WEB groups. An adherence checklist was used to ensure that the consultant implemented all procedures similarly for both WEB and FF groups. The protocol followed for each coaching session is described in detail in the manual (Ruble, Dalrymple, & McGrew, 2012).

Results

Analyses were conducted from an intent-to-treat perspective, with multiple imputation used to handle missing data and pooling of results via SAS Version 9.3. All available demographic variables were used as auxiliary variables in the imputation phase. Based on research by Graham (2009), several imputed data files (m = 20) were used to increase statistical power. Missing data analyses indicated that older children were less likely to have completed the study after baseline, $t(7.32) = 5.40$, $p = .001$.

Between-Groups Comparisons at Baseline

No mean differences between groups were observed on child, teacher, and family variables (see Table 1), except for DAS scores. Accordingly, DAS scores were used as a covariate in tests of between-groups differences. Although no difference in mean number of years teaching students with autism was noted between groups, the mean years was relatively low and ranged between 0.9 and 2.3 years.

Consultant Fidelity to COMPASS Consultation and Coaching Protocols

No mean differences were observed between WEB and FF groups on teachers' and parents' ratings of consultant adherence to COMPASS and teachers' ratings of coaching adherence (see Table 2). Out of a possible score of 25, overall mean adherence for the initial COMPASS consultation was 23.1 ($SD = 2.0$) as rated by teachers (i.e., 92% of the components were implemented), and 20.1 ($SD = 6.0$) as rated by parents (i.e., 80% of the components were implemented).

Teacher Fidelity to Teaching Plans

Mann-Whitney U tests indicated no differences between WEB and FF groups in teacher adherence for implementing teaching plans across Coaching Sessions 1–4 ($z = -0.3$, $p = .8$; $z = -0.5$, $p = .7$; $z = -0.3$, $p = .8$; $z = -0.3$, $p = .8$, respectively).

Results for Hypotheses

The assumption of independence was tested and considered tenable because individuals were randomly assigned to conditions, and a run chart of residuals showed no clear systematic pattern. Examination of boxplots, skewness, and kurtosis statistics, as well as statistical tests of normality, indicated that scores were approximately normally distributed within each group. Also, the homogeneity of variances assumption was tenable using Levene's test of equality of variances. Analyses of variance indicated no statistically significant between-groups differences on mean scores for the three comparison features of PET-GAS goals—measurability, equidistance, and level of difficulty.

Tables 3 and 4 highlight the results of the t test planned comparisons between groups on PET-GAS change scores unadjusted and adjusted for initial DAS scores. In general, both the unadjusted and adjusted results showed that the mean PET-GAS change score for the PBO group was significantly lower than for the FF and WEB groups. However, when analyses were adjusted for differences in initial DAS scores, the effect size differences (Cohen's d) decreased between the WEB and FF groups (0.56 to 0.27) and increased between the WEB and control groups (0.83 to 1.12).

Discussion

Together with our prior study, this replication adds further evidence that COMPASS consultation improves individualized outcomes in autism. There are few randomized controlled intervention studies, using either indirect (consultation) or direct (behavioral)

approaches, that have been shown to impact general outcomes for children with autism (Warren et al., 2011). Moreover, results from this second demonstration add to our confidence that COMPASS is efficacious. In addition, both studies detected large effects, replicating the size as well as the significance of the effect. Many interventions have a narrow focus on specific skills (Odom, Boyd, Hall, & Hume, 2010); importantly, these findings offer some evidence that this model has promise for improving outcomes across core areas of social, communication, and learning skills.

The results also provide initial evidence for the effectiveness of the WEB group compared to the control group. Further, we were unable to detect differences in effectiveness between the WEB and FF conditions, although the comparisons may have been somewhat underpowered. Moreover, and importantly, despite lack of direct contact with teachers, the WEB teachers were able to demonstrate fidelity similar to that of teachers in the FF group. These results provide hope for solutions for the need of empirically supported teacher coaching models that can be implemented in community-based settings (Odom, 2009) and have the potential to produce significant cost savings for schools as well as families. The efficacy of the use of web-based videoconferencing technology is important because it suggests a research-supported practice for facilitating the educational programs of students with autism regardless of geographic location. It also suggests ways that large urban school systems can use web technologies so that consultants can spend less time traveling between schools and more time having efficient and direct access to classroom teachers.

One issue of note was the unexpected finding that, despite randomization, IQ was dissimilar between groups. One explanation may be the relatively low sample size per group. This finding does demonstrate the need for researchers to conduct thorough baseline assessments that will allow verification of group equality so that appropriate adjustments can be made.

Future research is needed to adapt the model for different age groups and to test its implementation by different consultants. Both RCTs were conducted with young children and by consultants from the research team. In addition, continued study of the critical features necessary for optimal web-based implementation of the model is vital if children whose teachers receive web-based support are to have equal outcomes.

Conclusions

Although the limitations of traditional methods for training classroom teachers have been known for at least 30 years, research has lagged behind in offering alternative and innovative solutions that impact teacher behavior and, more important, can be shown to favorably impact student outcomes. Given the relative lack of scientific evidence for other training approaches, it is unsurprising that conferences and workshops, with their relative ease of delivery, continue to be a primary mode of training for classroom teachers of students with autism. Without supported models for improving teacher instruction, the gaps in knowledge about autism and limited use of research-supported interventions in classrooms will continue. No Child Left Behind (2001) mandates the provision of research-based educational interventions. Should not the same rigor of evidence extend to teacher training, teacher support, and professional development approaches?

Acknowledgments

This work was supported by Grants R34MH073071 and 1RC1MH089760 from the National Institute of Mental Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

We are grateful to the teachers, families, and children who generously donated their time and effort. We wish to thank special education directors and principals for allowing their teachers to participate. We also want to acknowledge our research team members Ryan Johnson, Rachel Aiello, Jessica Birdwhistell, Jennifer Hoffman, and Lauren Feltner for their efforts.

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed.. Washington, DC: Author; 2000. text rev.

Carrow-Woolfolk. Oral and Written Language Scales (OWLS). Circle Pines, MN: AGS; 1995.

Elliott, CD. Differential Ability Scales (DAS). New York, NY: Psychological Corporation; 1990.

Graham JW. Missing data analysis: Making it work in the real world. Annual Review of Psychology. 2009; 60:549–576.

Individuals With Disabilities Education Act. [Retrieved April 29, 2008] Building the legacy: IDEA 2004. 2004. from http://idea.ed.gov/

Joyce, BR.; Showers, B. Power in staff development through research on training. Alexandria, Va: Association for Supervision and Curriculum Development; 1983.

Joyce, BR.; Showers, B. Student achievement through staff development. 3rd ed.. Alexandria, VA: Association for Supervision and Curriculum Development; 2002.

Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, Rutter M. The Autism Diagnostic Observation Schedule–Generic: A standard measure of social and communication deficits associated with the spectrum of autism. Journal of Autism and Developmental Disorders. 2000; 30:205–223. [PubMed: 11055457]

MacKay DJC. Maximum likelihood and covariant algorithms for independent component analysis. 1996 from http://www.inference.phy.cam.ac.uk/mackay/ica.pdf.

Morrier M, Hess K, Heflin J. Teacher training for implementation of teaching strategies for students with autism spectrum disorders. Teacher Education and Special Education. 2011; 34:119–132.

National Research Council. Educating children with autism. Washington, DC: National Academy Press; 2001.

No Child Left Behind. Public Law 107–110, 115 Stat. 1425. 2001

Odom SL. The tie that binds: Evidence-based practice, implementation science, and outcomes for children. Topics in Early Childhood Special Education. 2009; 29:53–61.

Odom SL, Boyd BA, Hall LJ, Hume K. Evaluation of comprehensive treatment models for individuals with autism spectrum disorders. Journal of Autism and Developmental Disorders. 2010; 40:425–436. [PubMed: 19633939]

Ohio Center for Autism and Low Incidence. Autism Internet modules. 2007 from http://www.autisminternetmodules.org/user_about.php.

Ruble L, Birdwhistell J, Toland MD, McGrew J. Analysis of parent, teacher, and consultant speech exchanges and educational outcomes of students with autism during COMPASS consultation. Journal of Educational and Psychological Consultation. 2011; 21:259–283. [PubMed: 22639523]

Ruble LA, Dalrymple NJ, McGrew JH. The effects of consultation on Individualized Education Program outcomes for young children with autism: The Collaborative Model for Promoting Competence and Success. Journal of Early Intervention. 2010; 32:286–301. [PubMed: 21691449]

Ruble, L.; Dalrymple, N.; McGrew, J. The collaborative model for promoting competence and success for students with ASD. New York, NY: Springer; 2012.

Ruble LA, McGrew J, Dalrymple N, Jung LA. Examining the quality of IEPs for young children with autism. Journal of Autism and Developmental Disorders. 2010; 40:1459–1470. [PubMed: 20373007]

Ruble L, McGrew J, Toland MD. Goal attainment scaling as outcome measurement for randomized controlled trials. Journal of Autism and Developmental Disorders. Advance online publication. 2012

Schlosser RW. Goal attainment scaling as a clinical measurement technique in communication disorders. Journal of Communication Disorders. 2004; 37:217–239. [PubMed: 15063144]

Sheridan SM, Welch M, Orme SF. Is consultation effective? A review of outcome research. Remedial and Special Education. 1996; 17:341–354.

Sparrow, SS.; Cicchetti, DV.; Balla, DA. Vineland Adaptive Behavior Scales, Second edition (Vineland-II). San Antonio, TX: Psychological Corporation; 2005.

Warren Z, McPheeters ML, Sathe N, Foss-Feig JH, Glasser A, Veenstra-VanderWeele J. A systematic review of early intensive intervention for autism spectrum disorders. Pediatrics. 2011; 127:e1303–e1311. [PubMed: 21464190]

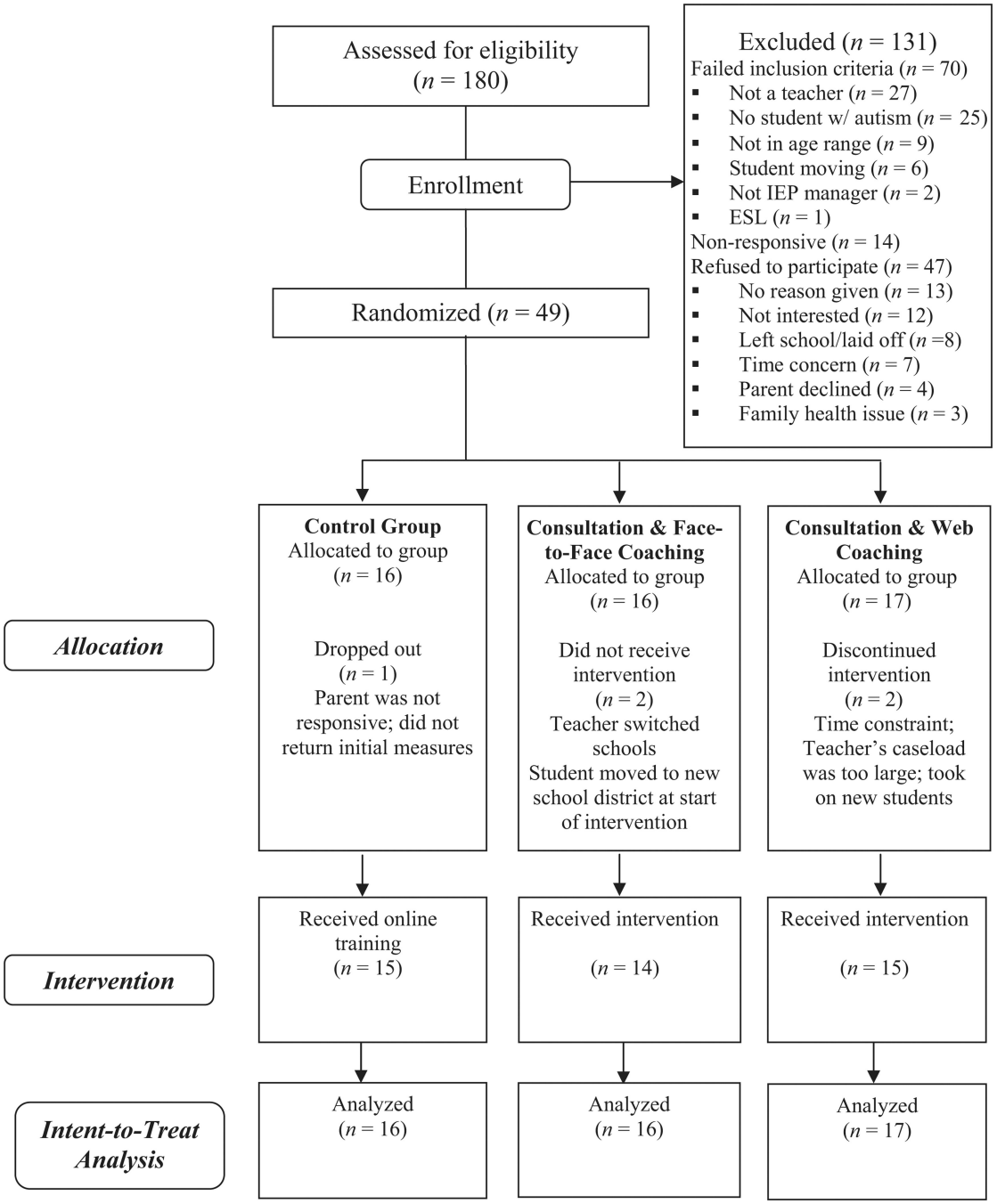

Figure 1.
Consort flowchart. IEP = individual education program; ESL = English as a second language.

Table 1

Comparison of Groups on Child, Teacher, and Family Characteristics at Baseline

	Enhanced services as usual (*n* = 15)		Face-to-face (*n* = 16)		Web-based (*n* = 18)			
Variable	***M***	***SD***	***M***	***SD***	***M***	***SD***	***F*(2, 46)**	***p***
DAS[a]	61.3	24.6	60.9	17.0	44.6	20.6	3.5	.03
OWLS[a]	53.8	13.7	57.3	14.7	49.6	10.7	1.5	.23
Vineland (TR)[a]	58.6	12.8	62.0	13.5	58.3	13.8	0.4	.67
Child age (years)	5.6	1.5	6.4	1.6	5.9	1.7	1.0	.61
Years teaching[b]	1.2	2.2	0.9	3.0	2.3	3.6	1.9	.15
Students taught	3.6	4.5	9.0	7.3	7.0	6.9	2.8	.06
Services received[c]	1.4	1.4	1.0	1.1	1.7	1.4	1.1	.32
Hours of services[c]	12.3	20.8	5.9	7.0	6.8	5.6	1.1	.34
Family income[d]	26.5		21.4		26.9		1.6	.51

Note. DAS = Differential Abilities Scale; OWLS = Oral and Written Language Scales; Vineland (TR) = teacher report on the Vineland Adaptive Behavior Scales; Years teaching = teacher report of number of years teaching students with autism; Students taught = number of students with autism taught across career; Services received = number of services students received outside of school during study duration; Hours of services = number of hours of services students received outside of school during study duration.

[a] Standard score.

[b] Welch test.

[c] Based on the end of the school year report.

[d] Kruskal Wallis test with $\chi^2(2)$ and *M* = mean rank.

Table 2

Comparison of Group Ratings on Consultant Fidelity to COMPASS Consultation and Coaching Protocols

	Face-to-face (*n* = 16)		Web-based (*n* = 17)				
Variable	***M***	***SD***	***M***	***SD***	***t*(*df*)**	***p***	**Cohen's *d***
Teacher ratings of consultant fidelity	23.21	1.95	22.97	2.13	0.34 (968)	.74	0.12
Parent ratings of consultant fidelity	20.65	4.55	19.62	4.96	0.49 (1,643)	.62	0.22
Teacher ratings of coaching adherence	3.74	0.27	3.78	0.27	0.42 (227)	.68	0.15

Note. COMPASS = Collaborative Model for Promoting Competence and Success. Cohen's $d = \frac{M_1 - M_2}{\sqrt{\frac{SD_1^2 + SD_2^2}{2}}}$.

Table 3

Comparisons of PET-GAS Change Scores Unadjusted and Adjusted for DAS Scores by Group

	Unadjusted			Adjusted for DAS scores	
Group	***M***	***SD***	***SE***	***M***	***SE***
Placebo	4.80	2.56	0.66	4.51	0.62
Face-to-face	8.43	2.60	0.65	8.16	0.61
Web-based	6.96	2.67	0.63	7.45	0.61

Note. PET-GAS = Psychometrically Equivalence Tested Goal Attainment Scaling; DAS = Differential Abilities Scale; Placebo = placebo control. DAS $t(42.16) = 2.76$, $p < .001$, with DAS $M = 22.93$.

Table 4

Planned Comparisons of PET-GAS Change Scores

	Unadjusted						Adjusted for DAS scores					
Comparison	M_{diff}	***SE***	***t*(*df*)**	***p***	**95% CI**	**Cohen's *d***	M_{diff}	***SE***	***t*(*df*)**	***p***	**95% CI**	**Cohen's *d***
FF vs. Placebo	3.63	0.92	3.94 (43.17)	<.001	[2.09, ∞]	1.41	3.65	0.92	4.25 (43.06)	<.001	[2.10, ∞]	1.41
WEB vs. Placebo	2.16	0.91	2.46 (41.83)	.01	[0.63, ∞]	0.83	2.94	0.91	3.30 (42.01)	.001	[1.41, ∞]	1.12
FF vs. WEB	1.47	0.90	1.63 (40.69	.06	[−0.05, ∞]	0.56	0.71	0.90	0.80 (40.53)	.22	[−0.81, ∞]	0.27

Note. PET-GAS = Psychometrically Equivalence Tested Goal Attainment Scaling; DAS = Differential Abilities Scale; M_{diff} = mean difference; *p* = *p* value (one-sided); CI = confidence interval (one-sided); FF = face-to-face; Placebo = placebo control; WEB = web-based. Cohen's *d* for PET-GAS adjusted for DAS scores were calculated using unadjusted variances with adjusted means.