SPRINGER BRIEFS IN PSYCHOLOGY
CHILD DEVELOPMENT

Lisa A. Ruble
John H. McGrew

COMPASS and Implementation Science

Improving Educational Outcomes of Children with ASD

Springer

SpringerBriefs in Psychology

Child Development

More information about this series at http://www.springer.com/series/10210

Lisa A. Ruble · John H. McGrew

COMPASS and Implementation Science

Improving Educational Outcomes of Children with ASD

Springer

Lisa A. Ruble
University of Kentucky
Lexington, KY
USA

John H. McGrew
Department of Psychology
Indiana University-Purdue University
Indianapolis
Indianapolis, IN
USA

ISSN 2192-8363 ISSN 2192-8371 (electronic)
SpringerBriefs in Psychology
ISSN 2192-838X ISSN 2192-8398 (electronic)
SpringerBriefs in Child Development
ISBN 978-3-319-18554-5 ISBN 978-3-319-18555-2 (eBook)
DOI 10.1007/978-3-319-18555-2

Library of Congress Control Number: 2015938723

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Acknowledgments

This work was supported by Grant Numbers 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, Lauren Feltner, Alexis Rodgers, Medina Jackson, Teri Belkin, Yue Yu, and Wing Hang Wong for their efforts. We wish to thank our colleagues, Jennifer Grisham-Brown for her work with the Common Elements of Teaching Index and Lee Ann Jung for her contribution to web-based teacher coaching. Lastly, and most importantly, we recognize Nancy Dalrymple, co-developer of COMPASS, for her leadership, vision, teaching, patience, and unwavering dedication to improving the lives of individuals with ASD and their families.

Lisa A. Ruble
John H. McGrew

Contents

List of Figures

List of Tables

About the Authors

Lisa A. Ruble is a Professor in the Department of Educational, School, and Counseling Psychology at the University of Kentucky. She is a licensed psychologist and clinician-researcher, who has provided social skills and behavioral interventions, school consultation and training, and parent training in autism spectrum disorders (ASD) for more than 25 years. She has authored more than 50 journal articles, book chapters, and manuscripts. Dr. Ruble's work in implementation research involves understanding how evidence-based practices can best be provided in community-based settings. In two randomized controlled trials funded by the NIH, Drs. Ruble and McGrew, and co-developer Nancy Dalrymple, tested and replicated the effectiveness of a parent-teacher consultation intervention called the Collaborative Model for Promoting Competence and Success (COMPASS). COMPASS is an effective practice for creating personalized and collaborative treatment plans that are guided by outcomes-based monitoring and coaching. She and Dr. McGrew have also created tools for measuring educational quality and educational outcomes. New studies are underway that focus on adult outcomes and quality of life, teacher burnout, parent stress, and parent empowerment.

John H. McGrew earned his Ph.D. in Clinical Psychology from Indiana University Bloomington. He is the former Assistant Chairman and Director of Clinical Training, and is currently Professor in the Department of Psychology at Indiana University Purdue University Indianapolis. Dr. McGrew has been principal or co-principal investigator of more than fifteen grants or contracts in the area of mental health services. Dr. McGrew has published over 75 articles in books or professional journals and made more than 80 presentations at international, national and regional conferences. His research focuses on interventions to help persons, and families of those, with severe mental illness, including autism, to improve their lives and to support them to live as fully as possible in the community.

Chapter 1
Overview of Evidence Based Practices in Psychology and Implementation Science

More than 1 % of children and youth have autism spectrum disorder (ASD; Developmental Disabilities Monitoring Network 2014), and the need for evidence-based intervention has never been greater. Unfortunately, the translation from research-to-practice takes time, almost 20 years (Green 2008)—too much time for children, teachers, and families who need high quality services today. For all children, public schools play one of the most important roles in the teaching and learning of critical skills. Attention to public schools is critical because they are the one place that all children with ASD, regardless of income, insurance status, and geographic location, receive services. But as the information on the specialized learning needs and evidence based approaches for students with ASD has mounted, the need for more informed teachers has also increased. The challenge is that special educators are expected to be broadly trained in educational practices across a wide variety of disability groups (Barnhill et al. 2011), and typically lack the time and specialized training to be aware of or competent in the latest advances in ASD intervention research. The need for innovative approaches for bridging the research-to-practice gap calls for new frameworks from implementation science for understanding how to facilitate the adoption and uptake of empirically supported interventions.

The purpose of this manuscript is to describe our work on an implementation strategy called COMPASS, the Collaborative Model for Promoting Competence and Success, a consultation based intervention specialized for students with ASD that has been tested in two randomized controlled trials. We begin by providing an overview of both evidence-based practices (EBP) and the evidence-based practices in psychology (EBPP) framework, followed by a review of what is known about consultation and what is unique about the COMPASS framework, and conclude with a description of the tested and hypothesized active ingredients of COMPASS and teacher influences of consultation effectiveness. We also present within our chapters the questions that guided our work, as well as the questions that remain for future research.

The importance of consultation as an organized and empirically validated human service intervention has been recognized for decades. In 1949 Gerald Caplan described the first consultation model. He was responsible for the mental health needs of 16,000 immigrant adolescents and a small clinical staff in Israel (Caplan 1970). Challenged by the more than 1,000 requests for help made by caretakers responsible

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*,
SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2_1

for the children who lived in more than 100 institutions, Caplan sought to create a more effective and efficient mental health delivery approach. His innovation was to focus on consultative activities. His insight was to understand that consultation provides a multiplier effect, such that a single expert consultant can help provide the needed expertise and support for multiple direct care staff. Similarly, ASD consultation uses an indirect or extender approach to client care; it strives to improve the abilities of caregivers rather than to attempt to provide traditional one-on-one therapy directly. Since Caplan's discoveries, the need for evidence-based consultation is as important today as ever especially in areas with expanding needs and limited numbers of expert providers in ASD.

Overview of Evidence Based Practices in Psychology

Evidence-based practice. Evidence-based practice is the current accepted standard for clinical and intervention practice across a variety of fields (e.g. medicine, nursing, dentistry, psychology) and treated conditions (Baker et al. 2008; Chambless and Ollendick 2001; Nathan and Gorman 2007), including autism (Mesibov and Shea 2011; Reichow et al. 2011). Evidence based practice is defined as an intervention for which there is strong research demonstrating effectiveness in improving client outcomes (Chambless and Ollendick 2001). Within autism, for example, reviews have begun to identify interventions that have been tested empirically and that meet at least one of the evidentiary standards for EBP (e.g., at least two RCTs), although the evidence is still relatively weak for many interventions and concentrated in interventions for a limited portion of those with ASD (higher functioning children and adolescents) (National Autism Center 2009; Wong et al. 2013).

However, although there is general acceptance of the need for and importance of EBPs, there also is resistance to the EBP movement (Bohart and Tallman 2010; Chambless and Ollendick 2001; Miles and Loughlin 2011; Tannenbaum 2005). This resistance springs in part from concerns about the primacy of EBPs as the only model for clinical practice. Specific criticisms cover a range of methodological, conceptual and practical grounds (e.g., unrepresentative client samples and settings, narrow definitions of effectiveness, over-reliance on randomized controlled trial (RCT) designs) (see Chambless and Ollendick 2001; Westen et al. 2004 for reviews). Three particularly salient critiques include (1) definitional confusion about what constitutes an EBP, (2) concerns about the overemphasis on clients with pure single diagnoses with the result that many EBPs do not apply to clients typically seen in therapy, e.g., co-morbid clients or those who present with subclinical symptoms, and (3) concerns that EBPs over emphasize differences between treatments and ignore equally strong evidence for factors common across treatments.

With respect to the first issue, one problem is that the criteria for EBPs differs across investigators (Mesibov and Shea 2011; Nathan and Gorman 2007; Roth and Fonagy 2005; Tannenbaum 2005; Thyer and Pignotti 2011; Westen et al. 2004). COMPASS, for example, fits the criteria for a promising practice, e.g., strong

positive evidence from two separate RCTs, but not an EBP according to the criteria from the original APA Division 12 task force on empirically validated treatments, i.e., a minimum of two RCTs from at least two separate research groups (Chambless and Hollon 1998); however, it meets the criteria for an EBP as outlined by Roth and Fonagy (2005) in their review for the British Health Services, i.e., a controlled replicated demonstration of effectiveness or a single high-quality RCT. Similarly, two recent reviews of ASD interventions used very different criteria for EBP. The National Professional Development Center on ASD (Wong et al. 2015) listed three different criteria for an EBP (e.g., at least two high-quality experimental/quasi-experimental studies conducted by at least two research groups, at least five high-quality single-case design (SCD) studies conducted by at least three different research groups), whereas the National Standards Project (National Autism Center 2009), classified treatments as evidence based on reviewer ratings of three or higher on a Scientific Merit Rating Scale encompassing five weighted domains of methodological quality.

With respect to the second point about limited applicability of EBPs, critics note that psychological practice is not diagnosis-focused (the standard for EBPs) but individual-focused, and is over-simplified by an approach that presumes a simple matching from diagnosis to a list of acceptable interventions for each diagnosis (APA Task Force 2006; Miles and Loughlin 2011; Thyer and Pignotti 2011). That is, intervening with an individual client requires an ongoing decision-making process that must take into account the interplay among three equally critical areas: research evidence (EBPs), patient factors, and clinical expertise (APA Task Force 2006). Unfortunately, most literature on EBPs has focused on the first area, research evidence.

The last issue, speaks to the tendency of the EBP approach to emphasize differences rather than similarities between empirically validated treatments. An alternative approach is to identify factors common across treatments that likely account for most of the variance underlying treatment success (Bohart and Tallman 2010; Lambert 2013; Kazdin 2008). That is, when EBPs are compared against viable alternate treatments or each other, rather than against placebo or 'services as usual', typically no difference is found (Wampold 2006). In contrast to these minimal comparative treatment effects, there is a vast literature on the large impact of therapist (e.g., therapist sense of well-being), client (e.g., IQ, level of functioning, self-efficacy), and relationship (e.g., therapeutic alliance) variables on treatment outcome, beyond the specific effects of treatment (see Bohart and Tallman 2010; Lambert 2013).

Based in part on these concerns about EBPs, the American Psychological Association convened a task force whose final report proposed a new term, Evidence Based Practice in Psychology (EBPP: APA Task Force 2006). The purpose of the task force was to craft an approach to practice that recognized and valued the rigorous empirical approach for identifying what works that characterizes EBP, while also attending to the practical realities of everyday clinical practice with clients with multiple morbidities and unique characteristics that may not align with the use of a particular single EBP. The result was EBPP (see Fig. 1.1), which is

Fig. 1.1 Evidence based practice in psychology framework (APA 2006)

defined as the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences.

Similar issues plague practice and research in autism. Much remains unknown about the integration of science and practice and the effective delivery of evidence-based treatments for persons with ASD in community settings (Office of Autism Research/Interagency Autism Coordinating Committee [OAR/IACC] 2012). Moreover, the majority of the available treatment research has been limited to examinations of the efficacy of a focused intervention on a specific outcome. There has been little to no attention on the practical issues facing clinicians when attempting to implement treatments in the real world, such as the influence of client, family, or therapist characteristics on clinical decision-making and treatment outcomes (Mesibov and Shea 2011).

The EBPP approach is now gaining acceptance. However, although the EBPP approach successfully addresses many of the concerns of clinicians while also integrating the lessons of science, its accurate application highlights several new areas of concern. Two critical areas of concern are the paucity of research on setting and client factors that can help inform an evidence based decision, and how best to make a good clinical decision (McGrew et al. 2015). This latter concern is made more challenging given the vast literature on the superiority of actuarial over clinical decision making (e.g., Dawes et al. 1989; Grove et al. 2000). In this context, it is our belief that COMPASS provides a model for clinical decision making within EBPP.

Consultation is ideal for bridging the research-to-practice gap (Ruble et al. 2012; Sheridan and Kratochwill 2007). That is, COMPASS, as a consultation model, explicitly ties EBPs to EBPP. Specifically, COMPASS is a process-based framework that provides an approach for the clinical decision-making needed to integrate the information from all three overlapping domains of the EBPP model

(see Fig. 1.1), while also systematically gathering the information within each domain—the setting/ecological factors, the family/child with ASD factors, and the teacher/clinician factors that need to be taken into account.

Development of an EBP

The evidence ladder		Intervention science activity
1. Reliable intervention	←	Post recognition quality monitoring
2. Disseminable	←	Disseminability studies
3. Effective	←	Multiple & multi-site replication studies
4. Conditionally effective	←	Initial evaluation studies
5. Emerging	←	Pilot studies: manuals, fidelity & outcome measures
6. Program of interest	←	Discovering & describing interesting Programs: basic research, clinical judgment

Ladder of Evidence. Several intervention development/implementation models informed the development of the COMPASS model. The Ladder of Evidence Model (Leff et al. 2003) provides the largest context and is a good overview of our developmental process (see Fig. 1.2). As shown, the development and eventual dissemination of an EBP progresses through a series of six hierarchical steps. At the first step, the developers discover a promising new approach for some clinical disorder or problem. At this stage, case studies, clinical experience and program evaluation all help to provide the developers with the initial set of ingredients and critical elements that comprise the first iteration of the intervention.

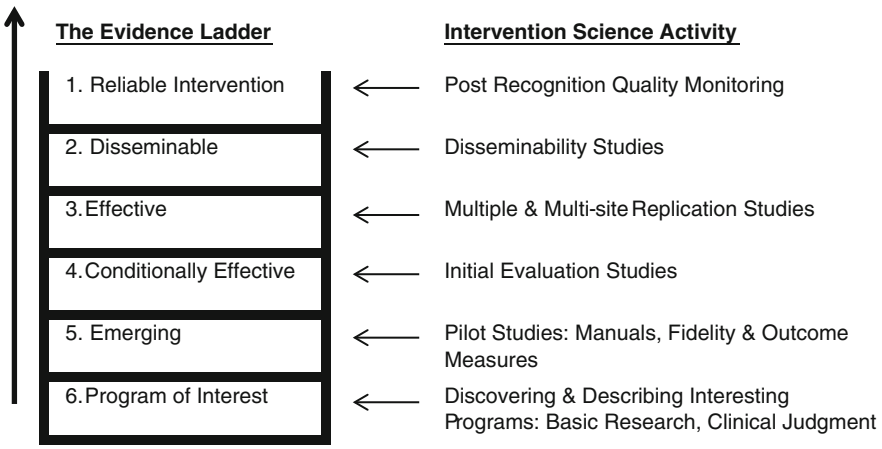


Fig. 1.2 Ladders of evidence

This first step for COMPASS is described in more detail in Chap. 2. The next step comprises the pilot studies, where the initial iteration is first formally tested as a complete package. During this step manuals and fidelity scales begin to be developed. The third step concerns the initial evaluation studies, usually with RCT designs, in which the intervention is first shown to be effective in a rigorous clinical trial. At this point, the intervention is considered to be an emerging or promising practice. The COMPASS studies comprising the second and third steps are described in more detail in Chaps. 4 and 5. The fourth step concerns further effectiveness studies that are larger and multi-site. We are beginning to do these studies for COMPASS. The fifth (dissemination) and sixth stages (reliable intervention) comprise what is often referred to as implementation science. Once an EBP has been identified, there is still a need to insure that it is disseminated and implemented accurately. This requires the development of training protocols, and a suite of fidelity and outcome measures to guide and track faithful implementation. We are also vigorously pursuing these aims, and discuss our progress in this regard in Chaps. 5 and 6. Overall, the Ladder of Evidence Model provides a good overview of our process. However, as discussed in the next section we also supplemented this model with additional frameworks.

Dunst and Trivette Framework. Two further frameworks helped to guide our research program. Both build on the Ladder of Evidence and provide further explication of particular steps. The first by Dunst and Trivette (2012) expands on steps five and six of the Ladder of Evidence. In this framework they make a helpful differentiation between implementation strategies and the intervention strategy. As originally envisioned by Dunst and Trivette, implementation strategies represent those practices used to support the accurate implementation of the intervention (e.g., training, fidelity monitoring, outcomes monitoring, etc.). That is, the implementation does not intervene directly with the intended clients or students, but refers to those strategies that support the intervention implementation, and thus any impact on client or students outcomes is indirect. This is a very helpful framework for understanding a consultation model, such as COMPASS. In this framework, the implementation practice refers to the methods used by consultants, coaches, and trainers to teach the intervention practice or EBP to the teacher, clinician, parent or service provider that will result in improved child or client outcomes. That is, the implementation practice is what the consultant does with the teacher and the intervention practice is what the teacher does with the child. In our work, COMPASS has served as the evidence-based implementation strategy proven to result in better educational outcomes for children with ASD. The link between COMPASS (what the consultant does with the teacher) and child outcomes (what the teacher does with the child) is the intervention practice or EBP. Each of the three areas in Fig. 1.3 represents interdependent activities that are both distinct and also linked to each other. In other words, the quality of the implementation practice (COMPASS consultant fidelity) should be associated with the quality of the intervention practice (teacher fidelity), which subsequently is associated with the effectiveness of the practice outcomes (child educational goal attainment). In later chapters, we will present data that show the relationship between these three areas.

Fig. 1.3 Implementation science framework (Dunst et al. 2013)

Integrated Model. Our integrated model includes both the features of EBPP and the Dunst and Trivette (2012) framework, while also aligning with steps two through six of the Ladder of Evidence (see Fig. 1.4). The EBPP factors are represented by the internal and external factors described under consultant, teacher, and child behavior and in Chaps. 7 and 8. The Dunst and Trivette framework is represented by the hashed lines and includes the quality elements associated with the implementation and intervention practice variables. As shown, there are three primary players (represented by the three central blocks) that impact COMPASS outcomes, the consultant, the teacher and the student with ASD. The outputs of each central block are the specific behaviors of the consultant (e.g., providing feedback/education, providing support), the teacher (e.g., engaging the child directly, providing prompts) and the student with ASD (e.g., engaged with the teacher, compliance with directions, off-task behavior).

Factors that can impact the outputs or behaviors of each actor are modeled as internal and external factors. These factors serve either to support or hinder the individual in performing their specific tasks within COMPASS. Moreover, external and internal factors can refer either to general factors or those specific to COMPASS. For example, for the consultant, external factors include training in

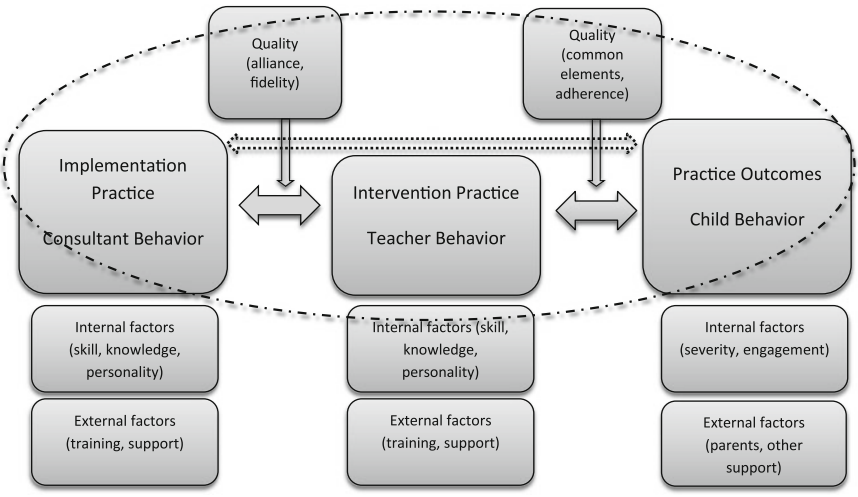


Fig. 1.4 Integrated model

consultation practices generally, training in COMPASS specifically, and support from other consultants or administration. Internal factors could include general skills and knowledge (listening skills, observational or assessment skills, knowledge of autism) and skills specific to COMPASS (ability to create good goals, knowledge of COMPASS model), as well as personal factors (sense of well-being, burnout, personality—outgoing vs. introverted). Similarly for teachers, external factors include training (both general training in special education and specific to COMPASS) and support (other teacher support, general support from family and friends, administrative support, consultant support, workplace supports—time, equipment), and internal factors could include skills/knowledge, again both general and specific to COMPASS (knowledge of autism, skills in data collection, knowledge of COMPASS model) as well as personal factors (burnout, stress, optimism). For students, external factors include supports (teacher, parents, other students or professionals) and training (teacher instruction and feedback) and internal factors include knowledge/skills (good attentional ability, educational attainment, language skills) and personal factors (autism severity, intellectual disability). It should be noted that the initial COMPASS consultation provides a thorough assessment of the internal and external factors impacting the student.

Another critical feature of the model is an assessment of the quality of the interactions between the consultant and the teacher and the teacher and the student. As with the internal and external factors, quality can reflect practices specific to COMPASS or general practices characteristic of good consultant or teacher practice. For example, the quality of the consultant-teacher interaction might reflect elements of good consultation generally (empathy, rapport, reflective statements) or of COMPASS specifically (adherence to COMPASS coaching protocol, feedback of goal attainment). Similarly, the quality of the teacher-student interaction might reflect elements of good teaching generally (prompt feedback, joint attention) or of COMPASS specifically (adherence to COMPASS teaching plan, high quality IEP goals).

As shown in the Fig. 1.4 and explained above, together these quality and internal and external factor elements define the critical factors impacting outcomes. That is, within the Integrated Model, understanding success for the student receiving the COMPASS intervention (teacher behavior) and COMPASS implementation (consultant behavior) requires knowledge of the impact of each of these potential internal, external and interaction quality factors. We will refer to this framework in the subsequent chapters as we describe our questions we tested in our RCTs.

Current Status of EBP and EBPP in Autism

Claims of autism treatment efficacy and purported cures arguably have caused more controversy compared to any other disorder because the large majority of treatment research has not been tested going through the ladders of evidence. Unlike medical disorders that have a recognized biological source, such as diabetes, where there is

an identified underlying causal mechanism that can be objectively measured with medical tests, there is no such understanding of autism. The lack of a biological marker makes autism vulnerable to claims not supported by research (Offit 2008). But even when we do have evidence that an intervention, biological or psychosocial, is helpful, it is necessary to identify why a particular approach works. When we understand the underlying mechanisms of change to explain why something works, then we can further our research to help identify those variables that affect change and more importantly, how we can enhance the effects and make them widely available. Additionally, change mechanisms may have an impact beyond a particular intervention, such as COMPASS, and underlie interventions generally. The identification of such crosscutting principles can have implications for the larger therapy literature (e.g. therapist alliance).

Although autism treatments are abundant, very few have been tested using strong experimental design (Wong et al. 2013) or have been examined for potential mechanisms of action. In fact, according to Wong, there have only been 38 total randomized control trials of interventions targeting autism! To contextualize this low number, it is worth noting that in a review of a single psychosocial intervention for severe mental illness (intensive case management), Dieterich et al. (2010) were able to identify 38 RCTs, matching the total for *all* psychosocial interventions for children and youth with ASD.

Despite the need for more rigorous testing, researchers have reached consensus on key underlying elements important for effective learning common across different treatment models. To obtain this information the National Research Council (Lord and McGee 2001) convened experts in autism interventions to come together and summarize the critical ingredients of effective programs. The committee identified six features that were common across all programs. In addition to these central features, the committee also identified areas of instruction that should be included in a program. These areas are listed in Figs. 1.5 and 1.6 below.

These represent our best current "guesses" about what is critical for autism intervention. As discussed in subsequent chapters, we incorporated these principles within COMPASS. Moreover, through an ongoing series of rigorous study, we

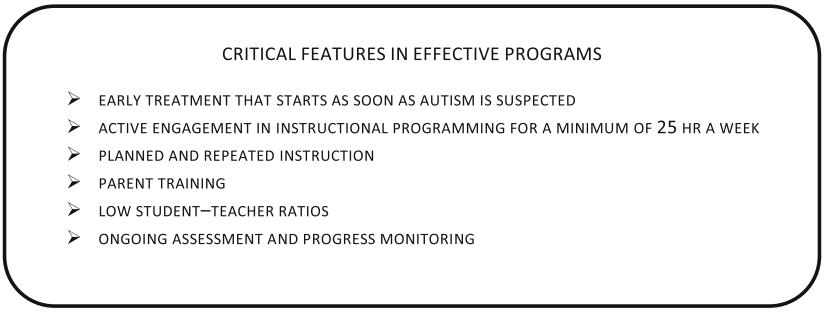


Fig. 1.5 Critical features in effective programs

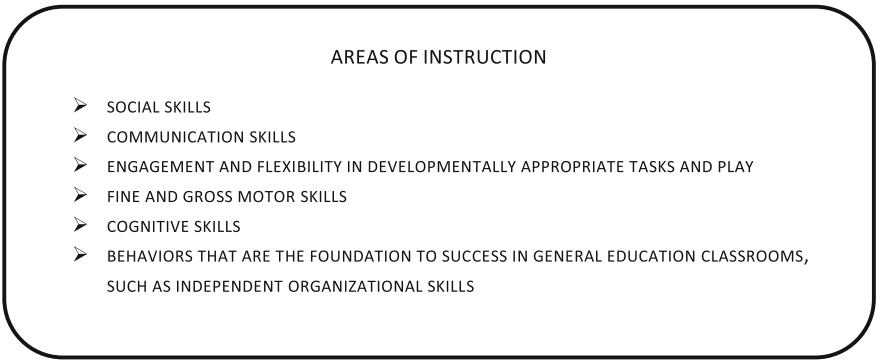


Fig. 1.6 Areas of instruction in effective programs

continue to strive to identify empirically those factors that underlie and explain COMPASS intervention and implementation success. The next chapter details the initial development and history of COMPASS.

References

APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist, 61*, 271–285.

Baker, T. B., McFall, R. M., & Shoham, V. (2008). *Current status and future prospects of clinical psychology: Toward a scientifically principled approach to mental and behavioral health care* (Vol. 9). United Kingdom: Wiley-Blackwell Publishing Ltd.

Barnhill, G. P., Polloway, E. A., & Sumutka, B. M. (2011). A survey of personnel preparation practices in autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 26*(2), 75–86.

Bohart, A. C., & Tallman, K. (2010). Clients: The neglected common factor in psychotherapy. In B. L. Duncan, S. D. Miller, B. E. Wampold & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 83–111, 455, xxix). Washington: American Psychological Association.

Caplan, G. (1970). *The theory and practice of mental health consultation*. New York: Basic Books.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*(1), 7–18.

Chambless, D. L., & Ollendick, T. H. (2001). *Empirically supported psychological interventions: controversies and evidence* (Vol. 52). USA: Annual Reviews.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*(4899), 1668–1674.

Developmental Disabilities Monitoring Network Surveillance. (2014). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *MMWR Surveill Summ, 63*(Suppl 2), 1–21.

Dieterich, M., Irving, C. B., Park, B., & Marshall, M. (2010). Intensive case management for severe mental illness. *The Cochrane Library*.

Dunst, C. J., & Trivette, C. M. (2012). Meta-analysis of implementation practice research. In B. Kelly & D. Perkins (Eds.), *Handbook of implementation science for psychology in education* (pp. 68–91). New York: Cambridge University Press.

Dunst, C. J., Trivette, C. M., & Raab, M. (2013). An implementation science framework for conceptualizing and operationalizing fidelity in early childhood intervention studies. *Journal of Early Intervention, 35*(2), 85–101.

Green, L. W. (2008). Making research relevant: If it is an evidence-based practice, where's the practice-based evidence? *Family Practice, 25*(Suppl 1), 20–24. doi:10.1093/fampra/cmn055.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: a meta-analysis. *Psychological Assessment, 12*(1), 19–30.

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge based, and improve patient care. *American Psychologist, 63*, 146–159.

Lambert, M. J. (2013). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.). New York: Wiley.

Leff, S., Conley, J., Campbell-Orde, T., & Bradley, V. (2003). *Getting to systems that promote self-determination through research and evaluation.* UIC NRTC 2003 Self-Determination and Psychiatric Disability Conference Papers, 158–180.

Lord, C., & McGee, J. P. (2001). *Educating children with autism*. Washington: National Academy Press.

McGrew, J., Ruble, L., & Smith, I. (2015). Current status of evidence based practice in psychology in autism: the field is only partially clothed. Manuscript in preparation.

Mesibov, G. B., & Shea, V. (2011). Evidence-based practices and autism. *Autism, 15*(1), 114–133.

Miles, A., & Loughlin, M. (2011). Models in the balance: Evidence-based medicine versus evidence-informed individualized care. *Journal of Evaluation in Clinical Practice, 17*, 531–536.

Nathan, P. E., & Gorman, J. M. (2007). *A guide to treatments that work* (3rd ed.). New York: Oxford University Press.

National Autism Center. (2009). National standards report. Retrieved March 2, 2015 from http://www.nationalautismcenter.org/reports/.

Office of Autism Research Coordination (OARC). (2012). National institute of mental health and thomson reuters, inc. on behalf of the interagency autism coordinating committee (IACC). *IACC/OARC Autism Spectrum Disorder Research Publications Analysis Report: The Global Landscape of Autism Research*. Retrieved July 2012 from http://iacc.hhs.gov/publications-analysis/july2012/index.shtml.

Offit, P. A. (2008). Vaccines and autism revisited—The Hannah poling case. *New England Journal of Medicine, 358*(20), 2089–2091.

Reichow, B., Doehring, P., Cicchetti, D. V., & Volkmar, F. R. (2011). *Evidence-based practices and treatments for children with autism*. New York: Springer Science & Business Media.

Roth, A., & Fonagy, P. (2005). *What works for whom?: a critical review of psychotherapy research*. New York: Guilford Press.

Ruble, L., Dalrymple, N., & McGrew, J. (2012). *Collaborative model for promoting competence and success for students with ASD*. New York: Springer.

Sheridan, S. M., & Kratochwill, T. R. (2007). *Conjoint behavioral consultation: Promoting family-school connections and interventions*. New York: Springer Science & Business Media.

Tanenbaum, S. J. (2005). Evidence-based practice as mental health policy: Three controversies and a caveat. *Health Affairs, 24*(1), 163–173.

Thyer, B. A., & Pignotti, M. (2011). Evidence-based practices do not exist. *Clinical Social Work Journal, 39*(4), 328–333.

Wampold, B. E. (2006). Not a scintilla of evidence to support empirically supported treatments as more effective than other treatments. In J. C. Norcross, L. E. Beutler, & R. F. Levant (Eds.), *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions* (pp. 299–308). Washington: American Psychological Association.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin, 130*(4), 631–663.

Wong, C., Odom, S. L., Hume, K, Cox, A. W., Fettig, A., Kucharczyk, S., Schultz, T. R. (2013). *Evidence-based practices for children, youth, and young adults with Autism Spectrum Disorder*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute, Autism Evidence-Based Practice Review Group.

Chapter 2
COMPASS Development

Since Caplan's seminal work (1970) of the first demonstration of the potential of consultation, consultation has expanded into an effective implementation strategy for improving educational outcomes of children and youth with academic and behavioral difficulties (Busse et al. 1995; Medway and Updyke 1985; Sheridan et al. 1996), especially for those children with complex needs such as ASD (Ruble et al. 2010). Not surprisingly, then, consultation is a standard intervention provided in schools. Professional development accreditation programs such as the National Association of School Psychology (National Association of School Psychologists 2010) and the American Psychological Association (Fouad et al. 2009) require consultation as a core competency of psychologists. A national survey indicated that about 9 out of 10 school psychologists provided teacher consultations and about two-thirds provide parent consultation regarding ASD. However, only about half worked with parents and teachers of students with ASD together (Aiello and Ruble 2011). This is troublesome because research suggests that the most positive outcomes with the greatest clinical significance from consultation occur when parents and teachers work together (Guli 2005). Parent-teacher collaboration may be even more important for children with ASD given the need for ecological interventions in ASD that bring parents and teachers together to set goals and implement strategies (National Research Council 2001) that address the need for generalization of skills across settings. Another concern identified from the nationwide survey was that consultation models tend to be generic and non-specific to ASD, with little to no empirical evidence for efficacy or effectiveness. More importantly, 4 out of 5 school psychologists surveyed indicated they needed more training in developing family-centered educational plans (Aiello and Ruble 2011; Aiello et al. 2015). This finding is consistent with other research that training for other school professionals that focuses specifically on ASD is insufficient (Singer 2008). Expert consultation can help to fill these gaps in training and practice, specifically, an evidence-based school consultation for ASD that both utilizes an ecological approach and includes the perspectives of the family of children with ASD.

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*,
SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2_2

Consultation and Coaching as an Implementation Strategy for Moving EBT in Classrooms

As mentioned, there is a wealth of evidence that research supported practices for ASD are not utilized in educational settings and if they are, often are implemented with poor fidelity (Hess et al. 2008; Morrier et al. 2011; Stahmer 2007; Stahmer et al. 2005). This research-to-practice gap has led to an entire field of study to identify methods to enhance dissemination and implementation of EBPs—implementation science (Brownson et al. 2012; Kelly and Perkins 2012). Effectively transferring EBPs into school and community settings requires proven implementation strategies (Odom 2009; Odom et al. 2013). Having a tool chest of EBPs and implementation fidelity checklists are not enough. Also necessary are research-supported approaches for supporting the transport of EBPs into classrooms and other community settings. Consultation, as an implementation *strategy*, is ideal for bridging the research-to-practice gap (Ruble et al. 2012; Sheridan et al. 2008). That is, a key role and rationale for COMPASS, as a consultation model, is that it explicitly ties EBPs to EBPP. Specifically, as noted earlier in Chap. 1, COMPASS is a process-based framework that provides a systematic approach to guide the clinical decision-making needed to integrate the information from all three overlapping domains of the EBPP model, Fig. 1.1, while also systematically gathering the information within each domain—the setting/ecological factors, the family/child with ASD factors, and the teacher/clinician factors that need to be taken into account.

Consultation Defined

What exactly does it mean to be a consultant? This is an important question in the area of ASD because depending on one's training background, the definition can vary dramatically and have an impact on consultation delivery and effectiveness. The answer is further complicated because consultation also means different things depending on the context and purpose. Often it is thought of as a brief interaction when an expert shares information with a nonexpert, where the recipient is the eventual intervener with the individual with ASD. Other times it can refer to monitoring and feedback concerning current practices, and the recipient might be administration or supervisory staff. Regardless of the specific understanding and definition adopted, consultation is a distinct activity, different from training, teaching, and supervision (Brown et al. 2011). It is important to clarify our definition because we believe consultation, as we define it, is associated with the measurable and active ingredients of COMPASS. Several formal definitions of consultation exists (e.g., Caplan 1970; Brown et al. 2011), but the definition most closely aligned to our work with COMPASS comes from Erchul and Martens (2010):

> School consultation is a process for providing psychological and educational services in which a specialist (consultant) works cooperatively with a staff member (consultee) to improve the learning and adjustment of a student (client) or group of students. During face-to-face interactions, the consultant helps the consultee through systematic problem solving, social influence and professional support. In turn, the consultee helps the client(s) through selecting and implementing effective school-based interventions. In all cases, school consultation serves a remedial function and has potential to serve a preventative function (p. 12).

This broad definition of school-based consultation describes both features of COMPASS, the initial face-to-face consultation that includes teachers and parents and the subsequent coaching activities that may occur using multiple methods, including face-to-face or via a web-based approach. As consultants, we offer expert skills in ensuring the integrity of COMPASS; but most importantly, teachers and parents offer expert knowledge of the child, their concerns for the child, and their goals for the child. In the next section we describe the key elements of good consultation and the different models and theories of consultation. We also discuss how COMPASS expands on these models and what data we have on factors that influence COMPASS outcomes.

Consultation Models and Their Influence on COMPASS

There are two predominant models of consultation—Behavioral and Mental Health. As the name implies, the Behavioral Consultation Model originally described by Bergan and Tombari (1976) adopts a behavioral approach for understanding and intervening with individuals and takes into account the functional relationships between behaviors and environmental contingencies by emphasizing analysis of antecedents (what occurs before a behavior) and consequences (what occurs after a behavior). More recently, Sheridan and colleagues (Sheridan et al. 2001, 2002, 2006, 2008) have conducted extensive research on an expanded version of the Behavioral Model called Conjoint Behavioral Consultation (CBC). A key difference in the CBC model is the inclusion of both the parent and teacher as part of the consultation team. The obvious advantage of including both key participants is the value of consistency in understanding and approach to problems across the home and school settings. Given that children from birth to age 18 spend more than 90 % of their time outside of the school system, parent/caregiver contribution to developmental outcomes is imperative (Sheridan and Kratochwill 2007). The CBC framework includes four main phases: (a) problem identification, (b) problem analysis, (c) plan implementation, and (d) plan evaluation. Problem identification is concerned with identifying prioritized goals (e.g., decrease aggressive behaviors toward peers). Problem analysis involves review of the observations and goals used to develop the intervention plan. Plan implementation is the implementation and monitoring of the intervention. Plan evaluation is the determination of the student's progress toward goal attainment based on the plan and possible need for

modifications to the plan. COMPASS consultation encompasses these four elements in the initial consultation where parent and teacher priorities for social, learning, and communication skills are identified and intervention plans are developed based on the child's personal and environmental risk and protective factors. The subsequent teacher coaching sessions provide ongoing feedback and support to help ensure the success of the implementation of the intervention plans and identify and craft any needed modifications.

The Mental Health Model (Caplan et al. 1994), in contrast, builds from psychodynamic theories and stresses the importance of interpersonal relationships between the consultant and the consultee, that is, it is more focused on the process of consultation. It emphasizes the significance of consultants being aware of the necessity of understanding the norms, beliefs, habits, and routines of consultees, and that ultimately, the consultee is largely responsible for putting the intervention into effect. In this model consultants are viewed as resources, that is facilitators, rather than as experts. In fact, consultants who assume an expert role are less likely to achieve positive outcomes compared to consultants who present as facilitators or coaches and use key concepts of the Mental Health Model (Caplan et al. 1994). The fundamental factors of this approach include: (a) the relationship between the consultant and the consultee is equitable and nonhierarchical; (b) the consultant does not get involved in the personal problems of the consultee; and (c) the longer-term goal of consultation is to improve the functioning of the consultee to be successful and eventually independent in their work with individuals in the future.

COMPASS is influenced by both of these models. The behavioral model is more focused on the student as primary, with the consultee role as the mediator/implementer of student change, while the mental health model focuses more on the consultee as primary, with the assumption that once the consultee needs have been addressed he/she will be successful with the student. COMPASS incorporates critical aspects of both models. Because we recognize and incorporate into COMPASS the necessity of an antecedent-based approach for understanding the interplay between person-environment interactions as based on the Behavioral Model and also the need for attention to the interactions between the consultant and the consultee, and their own personal situation, as based on the Mental Health Model, we believe that COMPASS is unique. When working with students with autism, the need for a team approach is a necessity for the generalizability of outcomes (as in conjoint behavioral consultation), but the need for supported teachers is also necessary to address and moderate external as well as internal resources and challenges for the teacher, such as burnout or teacher stress, since the focus of consultation is what the teacher does when the consultant is not there. Thus COMPASS addresses both teacher and student needs. Next we describe how COMPASS expands the behavioral and mental health models and Dunst and Trivette framework (Dunst and Trivette 2012).

COMPASS Includes and Expands on Other Models

What makes COMPASS most unique from other consultation frameworks based on the behavioral or mental health models, is its influence by a third model—the Minnesota Competence Enhancement Program (MCEP) developed by August et al. (1992). Unique to the MCEP framework is the focus on the individual's adaptation and resilience as viewed from a community-based prevention and intervention perspective rather than as viewed from a deficit model as emphasized within traditional treatments for psychopathology conducted within clinical contexts. Also innovative is the model's focus on the development of competence as a supportive factor and cushion against challenges and failure. That is, the focus of intervention is not simply reducing deficits but also on identifying and enhancing competence. Probably most critical to the framework was the expansion of the therapeutic scope from narrow antecedent/consequence behavioral strategies to an understanding of the importance of ecological interventions that include the people who have the most frequent interactions with the child in their own environmental contexts and thus provide the necessary opportunity for naturalistic teaching, generalization, and skill maintenance.

As shown in Fig. 2.1, the framework considers four main areas that impact the development of competence—personal and environmental challenges and supports (Ruble and Dalrymple 2002). It is essential that the team understands how each of the four areas affect an individual with ASD (Fig. 2.1).

Challenges. Personal challenges include biological predispositions that increase risks to developing competence. In ASD, neurobiological differences in brain development and function are examples (Volkmar 2005). Such differences lead to impaired ways of processing information from the environment as well as

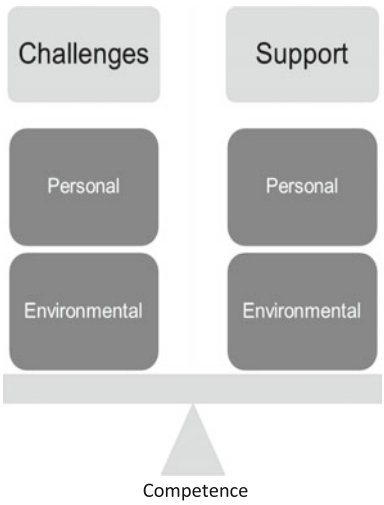


Fig. 2.1 Autism competency enhancement model

difficulties producing competent responses. The information-processing difficulties are manifested in the social communication problems of persons with ASD as well as in their narrow range of interests and unusual sensory or motor behaviors. Importantly these vulnerabilities are likely to occur early in life, impacting typical development and ability to respond competently to the social and communicative demands of the environment.

Adding to the personal challenges are environmental challenges that also interfere with competence development. Some possible environmental challenges include lack of knowledge about ASD, lack of appropriate supports for learning, confusing or loud environments, and punitive behavioral programs. Inadequate supports for direct teaching on communication, social, self-management, independence, leisure, and sensory needs contribute to failure. Stressors on the family system may also lead to further risk of poor competency development in people with ASD. Additional environmental challenges include lack of services, long waiting lists for community-based services, and poorly delivered services.

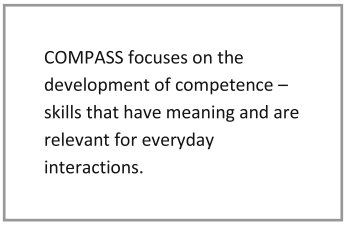
Supports. While it is important to understand the contribution that personal and environmental challenges of persons with ASD have on competence development, the real work comes from understanding how to enhance competence by increasing supports. Supports are the protective factors that serve to balance risk factors in helping to develop competency. During various periods throughout life, the need for protective factors will vary; however, individuals with ASD will always need help to build and keep personal and environmental supports.

Personal supports are the strengths, interests, and preferences that help produce and maintain competence. Assessment of individual strengths, interests, and preferences must be identified and used in treatment planning for the development of functional and meaningful life skills that impact quality of life. The assessment of strengths, interests, and preferences is considered an ongoing activity, not a static activity. These areas will change and expand over time and as the individual ages. Specific foods, riding in a car, rocking, spinning things, routines, sequences, patterns, numbers and letters, and moving—running, pacing, jumping—are examples of preferences that individuals with ASD may demonstrate.

Environmental supports refer to people, teaching methods, reinforcement strategies, and positive behavior supports—anything that assists the person in developing competence. Alone, environmental supports do not eliminate challenges, but rather they provide the balance on which to build competency. Environmental supports must be individualized. They also must be community-based and system-wide to appropriately meet each person's needs and to allow for generalizability to all environments. Within this approach, consultation can serve as an implementation strategy and as an environmental support to ensure consistency and stability through a continuum of services and the numerous teachers, various providers, and family members who all serve as supports. Critically, if we are going to be successful in supporting students and adults with ASD to be competent, we must collaborate across people, agencies, and government. In our book-length manual, we describe in more detail the COMPASS framework for identifying personal and environmental challenges and supports (Ruble et al. 2012).

As noted earlier, one key element of the model is the focus on competence enhancement as opposed to deficit reduction. The concept of competence enhancement as promulgated by Ruble and Dalrymple (1996) was novel because it linked individual learning progress and challenges to the environment. This was innovative because too often program plans were designed to address specific weaknesses, rather than addressing the whole person and how to ensure their strengths and preferences were included in treatment plans. Assessment of the needs of the individual along with stressors, challenges and resources, including strengths and interests is essential when taking into account the entire person. It is vital to focus on increasing protective factors while understanding vulnerabilities and ecological stressors.

The concept of developing competency served as the fundamental measure of quality of life and treatment success or outcome described in our manuscript "An alternative view of outcome" (Ruble and Dalrymple 1996). In this paper, we challenged the traditional approach for measuring adult outcomes and advocated for novel approaches that focused on the development of competence and quality of life as central outcomes that are closely linked to accommodations and social and family support networks. This work helped to reaffirm the evolving model's emphasis on collaboration and building supports rather than emphasizing deficits.



Discovery and Evolution of COMPASS

COMPASS originated from the need for a training framework for community-based service providers, such as teachers, adult residential providers, vocational rehabilitation counselors, and other service personnel, to understand the unique learning challenges, preferences, and strengths of each individual with autism. In 1992, in our first attempt to create a model, we adapted the Minnesota Competence Enhancement Program, which was called the Autism Competency Enhancement framework. In 1996, this model was used as the basis for the Autism Technical Assistance Manual for Kentucky Schools (Ruble and Dalrymple 1996) that was used to train teachers throughout the state of Kentucky. The training was specialized for students with ASD and was adapted and used for a variety of purposes, including educational planning purposes, addressing behavioral problems, and facilitating transitions. Later in 1998 the model served as the consultation framework for TRIAD at Vanderbilt University in the state of Tennessee and was

renamed the Collaborative Model for Promoting Competence and Success of Persons with Autism Spectrum Disorder (COMPASS).

Over the years, the necessity for a comprehensive model has not changed. The model was based on the practical realities of a need for better understanding of autism by those who have the most frequent interactions with individuals as well as a need for enhanced quality of life outcomes measured by participation in work, school, social interactions, in recreational and leisure activities. This is a reality that continues today. In the early 90s, a push for services provided locally and within natural environments led to the demand for knowledgeable community-based service providers of ASD in Indiana, where we were developing and testing COMPASS, as well as throughout the US. Today, we have a lot more knowledge about evidence-based practices, but still require a comprehensive, implementation strategy for improving educational outcomes that takes into account the cultural, psychosocial, developmental and neurobiological needs and resources of the individual considered within an ecological framework in the selection, modification and individualization of EBPs.

A training framework in ASD that explicitly calls for the individualization of teaching and therapeutic strategies is clearly needed, and indeed is mandated, because of the federal requirements for an Individual Education Program for all students with disabilities. This individualization is particularly challenging in ASD given the extreme heterogeneity of the disorder. 8, for example, Fig. 2.2 shows the diverse range in clinical presentation of persons with ASD. About 70 % of individuals have some degree of intellectual impairment, ranging from mild to severe or profound (Fombonne 2005) Social interactions also vary and individuals typically

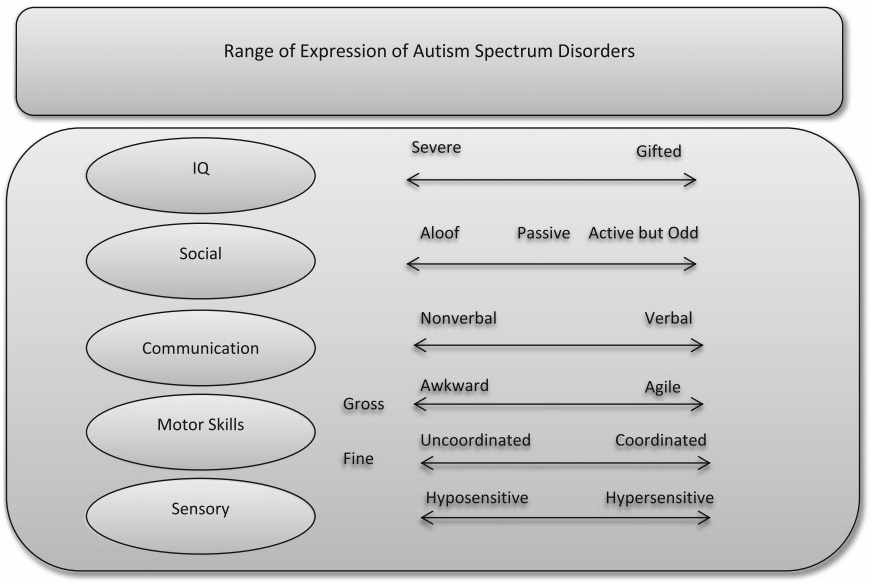


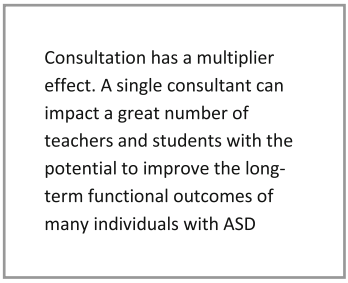
Fig. 2.2 Range of expression in ASD

fall within one of three categories—aloof, passive, and active-but-odd (Wing 2005). Individuals, who appear aloof, may have little interest in interaction with others. Those who are passive demonstrate an interest in interaction, but do not initiate and instead respond. The last group, active-but-odd, characterize individuals who do initiate, but in unusual ways. For verbal communication, about 20 % of individuals never develop spoken speech (Lord et al. 2004), and others may be quite verbal, but have limited reciprocal communication abilities. Gross and fine motor skills also vary from person to person. Some individuals may have well-coordinated fine motor and excellent gross motor skills, while other individuals may struggle with practical tools such as using eating utensils, buttoning shirts, or using a pencil (Rogers et al. 2005). Lastly, sensory processing skills are also quite variable (Behrmann and Minshew 2015). Some individuals may tolerate noises and other environmental sensitivities well and other individuals may become quite upset and unable to function in certain environments.

Following the prior early development work described above in Indiana, the framework was expanded from a training model for community-service providers to a framework for outpatient services for children and youth with ASD in Tennessee and Kentucky. Dismayed by the lack of available information on research supported interventions delivered in outpatient medical settings, the authors applied the COMPASS framework for each of the clinical services provided, which included early childhood, behavior management, social skills, and program planning. It was clear that an informational and process approach was needed because services were often limited by insurance and time. Given the limited number of sessions approved by insurance as well as the 60-min time limit, an approach that enhanced parental involvement and the decision-making of treatment goals and intervention plans was crucial. If we could demonstrate the clinical decision-making that goes into goal selection and intervention planning, then perhaps parents and caregivers would be better informed to make their own decisions and share information with other service providers outside the clinical outpatient setting. Thus, the process approach implied by COMPASS, and adapted from The Mental Health Model (Caplan et al. 1994), was thought to help empower the primary resource of children—their families. Those caregivers that were part of the clinical-decision making were thought to be better informed and equipped to make decisions and evaluate outcomes for facilitating their children's development.

As noted above, COMPASS has been used in a variety of contexts and settings, however, we believe that the underlying tenet of informed clinical-decision making is helpful not only in medical settings but also in educational settings. In fact, COMPASS has primarily been used and tested within the public school setting. The focus on educational settings is a result of the high numbers of students with autism being identified and included in schools and communities, and the corresponding need for professionals and support personnel who are strongly grounded in knowledge and experience of autism. Consultation as an intervention has the potential to facilitate the training and support needed by teachers and staff. Because consultation tends to have a multiplier effect, i.e., a single consultant can impact a great number of teachers and students, the use of consultants who can guide others

in designing and monitoring programs has the potential to improve the long-term functional outcomes of many individuals with autism.



Schools typically invest in professional development and training for improving teacher skills in autism and other areas using unproven methods (Morrier et al. 2011). However, decades of research shows that although large group workshops, in-services, and conferences are helpful for learning new concepts, they are unsuccessful for changing classroom practices (Joyce and Showers 1988, 2002).

Strategies that do work to change classroom practices incorporate three ingredients: (a) activities that allow for reflection and self-assessment of one's own knowledge as a means for identifying future activities of learning, (b) opportunities for mastery that engages the learner in a process of assessing one's experience within the context of a conceptual framework, and (c) real-life, rather than non-contextual learning activities (Dunst and Trivette 2009). Even consultation, as a one-time activity that includes some of these components is not enough. As discussed later, we have important evidence that coaching, that is the follow-up assistance that includes monitoring, feedback and supervised practice, is necessary for ensuring the implementation of teaching plans with high fidelity (Ruble et al. 2010, 2013). Thus, the above evidence-based features of effective training are embedded within COMPASS (Ruble et al. 2012).

Empirical Study and Development of COMPASS

Since 2004, federal funding from the National Institute of Mental Health has enabled the authors to continue to evaluate the effectiveness of COMPASS in three different studies. The first compared its effectiveness to special education services as usual. The second examined effectiveness via web-based technology. The latest study is in progress and will test COMPASS when adapted for older adolescents preparing to transition from school to post-school services using a series of iterative qualitative and quantitative pilot tests.

Additional Critical Factors Informing the Development of COMPASS

Another critical factor included in the model is a focus on measurable goals. As mentioned above, setting goals that are individualized and ecologically valid are a critical part of the COMPASS model, however, equally important is crafting goals that are measurable. Obviously, goals are much easier to evaluate and assess with this model when they are measurable and objective. For example, following goal setting, details about how to teach the goal and objective are generated from a shared understanding of the balance between the student's personal and environmental challenges and supports. The factors that create the balance are the ingredients necessary for achieving competence and are unique for each individual. As a framework, this model also helps train staff to understand and support the person more effectively. Over the years, we have learned that the most important impact we can have in consulting with parents and teachers is empowerment. A team that is empowered is one that has accurate information to make decisions, implement teaching plans and evaluate outcomes long after the consultant leaves.

Another critical factor in COMPASS is the adoption of a lifespan perspective and the creation of a shared understanding that competence looks different across an individual's lifespan. Challenges change over time and are constantly requiring new sets of competence—for the person with autism as well as their families and caregivers. People with ASD must have support from people who understand them, their personal and environmental challenges, and their personal supports all within a developmental framework, in order to know how and what environmental resources will enhance learning. Too often the person with ASD is viewed as the problem because those who are trying to teach and support them do not understand their uniqueness and how the environment contributes to challenges in learning and competence.

In the next section, we summarize how these features are integrated into consulting, as we discuss the two main COMPASS activities: (a) an initial, parent-teacher goal setting and treatment planning session and (b) follow-up teacher coaching and performance-based assessment activities. Rather than repeat what has been detailed in the COMPASS book-length treatment manual (Ruble et al. 2012), we discuss the research behind many of the key elements of COMPASS that help validate underlying assumptions and important mechanisms of change. In the next chapter, we start with a description of the analyses to test the assumption that COMPASS is collaborative. In the following chapters, we describe the approach we took to develop a sensitive outcome measurement tool that was valid. Then we describe our two randomized controlled trials with COMPASS. The first study compared a group of students whose teachers received COMPASS against a group of students who received special education services as usual. The second study included a third group of teachers who received coaching using web-based technology vs teachers who received traditional face-to-face coaching. Following discussion of the RCTs, we will present data on features crucial for positive

COMPASS outcomes. We have evidence for two key features (IEP quality and teacher adherence), but we also have hypothesized elements that will be reviewed. We then describe what we have learned about important teacher and child internal and external factors that impact COMPASS outcomes. We conclude with a discussion of questions answered and future research that is needed.

References

Aiello, R., & Ruble, L. A. (2011). Survey of school psychologists' autism knowledge, training, and experiences. *Poster presented at the annual convention of the national association of school psychologists*. San Francisco, CA.

Aiello, R., Esler, A., & Ruble, L. (2015). National study of school psychologists' use of evidence-based assessment in autism spectrum disorder. *In preparation*.

August, G. A., Anderson, D., & Bloomquist, M. L. (1992). Competence enhancement training for children: An integrated child, parent, and school approach. In S.-L. Christenson & J.-C. Conoley (Eds.), *Home-school collaboration: enhancing children's academic and social competence* (pp. 175–192). Maryland: National Association of School Psychologists.

Behrmann, M., & Minshew, N. (2015). Sensory processing in autism.

Bergan, J. R., & Tombari, M. L. (1976). Consultant skill and efficiency and the implementation and outcomes of consultation. *Journal of School Psychology, 14*(1), 3–14.

Brown, D., Pryzwansky, W. B., & Schulte, A. C. (2011). *Psychological consultation and collaboration: introduction to theory and practice* (7th ed.). New Jersey: Pearson.

Brownson, R. C., Colditz, G. A., & Proctor, E. K. (2012). *Dissemination and implementation research in health: translating science to practice*. Oxford: Oxford University Press.

Busse, R. T., Kratochwill, T. R., & Elliott, S. N. (1995). Meta-analysis for single-case consultation outcomes: Applications to research and practice. *Journal of School Psychology, 33*(4), 269–285.

Caplan, G. (1970). *The theory and practice of mental health consultation*. New York: Basic Books.

Caplan, G., Caplan, R. B., & Erchul, W. P. (1994). Caplanian mental health consultation: historical background and current status. *Consulting Psychology Journal: Practice and Research, 46*(4), 2–12.

Dunst, C. J., & Trivette, C. M. (2009). Let's be PALS: an evidence-based approach to professional development. *Infants and Young Children, 22*(3), 164–176.

Dunst, C. J., & Trivette, C. M. (2012). Meta-analysis of implementation practice research. In B. Kelly & D. Perkins (Eds.), *Handbook of implementation science for psychology in education* (pp. 68–91). New York: Cambridge University Press.

Erchul, W. P., & Martens, B. K. (2010). *School consultation: conceptual and empirical bases of practice* (3rd ed.). New York: Springer.

Fombonne, E. (2005). The changing epidemiology of autism. *Journal of Applied Research in Intellectual Disabilities, 18*(4), 281–294. doi:10.1111/j.1468-3148.2005.00266.x.

Fouad, N. A., Grus, C. L., Hatcher, R. L., Kaslow, N. J., Hutchings, P. S., Madson, M. B., & Crossman, R. E. (2009). *Competency benchmarks: a model for understanding and measuring competence in professional psychology across training levels* (Vol. 3). US: American Psychological Association.

Guli, L. A. (2005). *Evidence-based parent consultation with school-related outcomes* (Vol. 20). US: Guilford Publications.

Hess, K. L., Morrier, M. J., Heflin, L., & Ivey, M. L. (2008). Autism treatment survey: services received by children with autism spectrum disorders in public school classrooms. *Journal of Autism and Developmental Disorders, 38*(5), 961–971.

Joyce, B. R., & Showers, B. (1988). *Student achievement through staff development*. New York: Longman.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria: Association for Supervision and Curriculum Development.

Kelly, B., & Perkins, D. F. (Eds.). (2012). *Handbook of implementation science for psychology in education*. New York: Cambridge University Press.

Lord, C., Risi, S., & Pickles, A. (2004). Trajectory of language development in autistic spectrum disorders.

Medway, F. J., & Updyke, J. F. (1985). Meta-analysis of consultation outcome studies. *American Journal of Community Psychology, 13*(5), 489–505.

Morrier, M. J., Hess, K. L., & Heflin, L. J. (2011). Teacher training for implementation of teaching strategies for students with autism spectrum disorders. *Teacher Education and Special Education, 34*(2), 119–132.

National Association of School Psychologists. (2010). Professional standards and training. 2014, from http://www.nasponline.org/standards/practice-model/domains.aspx.

National Research Council. (2001). *Educating children with autism*. Washington: National Academy Press.

Odom, S. L. (2009). The tie that binds: evidence-based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education, 29*(1), 53–61.

Odom, S., Cox, A., & Brock, M. (2013). Implementation science, professional development, and autism spectrum disorders. *Exceptional Children, 79*(2), 233–251.

Rogers, S. J., Cook, I., & Meryl, A. (2005). Imitation and play in autism. *Handbook of autism and pervasive developmental disorders* (3rd ed., Vol. 1, pp. 382–405).

Ruble, L. A., & Dalrymple, N. J. (1996). An alternative view of outcome in autism. *Focus on Autism and Other Developmental Disabilities, 11*(1), 3–14.

Ruble, L. A., Dalrymple, N. J., & McGrew, J. H. (2010). 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, 32*(4), 286–301.

Ruble, L. A., Dalrymple, N. J., & McGrew, J. H. (2012). *Collaborative model for promoting competence and success for students with ASD*. New York: Springer Science + Business Media.

Ruble, L. A., & Dalrymple, N. J. (2002). COMPASS: a parent-teacher collaborative model for students with autism. *Focus on Autism and Other Developmental Disabilities, 17*(2), 76–83.

Ruble, L. A., Toland, M. D., Birdwhistell, J. L., McGrew, J. H., & Usher, E. L. (2013). *Preliminary study of the autism self-efficacy scale for teachers (ASSET)* (Vol. 7). Netherlands: Elsevier Science.

Sheridan, S. M., Clarke, B. L., Knoche, L. L., & Edwards, C. P. (2006). The effects of conjoint behavioral consultation in early childhood settings. *Early Education and Development, 17*(4), 593–617.

Sheridan, S. M., Eagle, J. W., Cowan, R. J., & Mickelson, W. (2001). The effects of conjoint behavioral consultation results of a 4-year investigation. *Journal of School Psychology, 39*(5), 361–385.

Sheridan, S. M., & Kratochwill, T. R. (2007). *Conjoint behavioral consultation: promoting family-school connections and interventions*. Springer Science & Business Media.

Sheridan, S. M., Kratochwill, T. R., & Burt, J. D. (2008). *Conjoint behavioral consultation: promoting family-school connections and interventions* (2nd ed.). New York: Springer.

Sheridan, S. M., Meegan, S. P., & Eagle, J. W. (2002). Assessing the social context in initial conjoint behavioral consultation interviews: an exploratory analysis investigating processes and outcomes. *School Psychology Quarterly, 17*, 299–324.

Sheridan, S. M., Welch, M., & Orme, S. F. (1996). Is consultation effective? a review of outcome research. *Remedial and Special Education, 17*(6), 341–354.

Singer, K. S. (2008). *The practice and training of school psychologists in the autism syndrome disorder* (Vol. 69). USA: ProQuest Information and Learning.

Stahmer, A. C. (2007). The basic structure of community early intervention programs for children with autism: provider descriptions. *Journal of Autism and Developmental Disorders, 37*(7), 1344–1354.

Stahmer, A. C., Collings, N. M., & Palinkas, L. A. (2005). Early intervention practices for children with autism: descriptions from community providers. *Focus on Autism and Other Developmental Disabilities, 20*(2), 66–79. doi:10.1177/10883576050200020301.

Volkmar, F. R. (2005). *Handbook of autism and pervasive developmental disorders* (3rd ed.). Hoboken: Wiley.

Wing, L. (2005). Problems of categorical classification systems. *Handbook of Autism and Pervasive Developmental Disorders* (3rd ed., Vol. 1, pp. 583–605).

Chapter 3
COMPASS Practice Outcome: Idiographic Assessment and Development of a Measure

The Challenge of Assessing Outcomes for Children in Special Education

As shown in our Integrated Model (Fig. 3.1), the practice outcome, i.e., student goal accomplishment, is the most critical aspect of our intervention. Thus, the first question that needs to be answered for a new intervention is whether or not it makes a significant impact, i.e., is effective. Goal setting and goal attainment are the key inputs and outcomes of COMPASS. As part of the development of COMPASS, we needed an outcome measure that was consistent with the goals of COMPASS, namely to improve educational outcomes. However, as described below, development of an outcome measure that was both sensitive to individualized goals and produced scores that were equivalent for group comparisons presented several challenges.

Because special education requires strict individualization of student goals for the Individual Education Program (IEP), one critical challenge to educators and to researchers is how to assess progress. For educators, the challenge is the appropriateness of using standardized curriculum based assessments to assess progress of children with specialized instructional goals. For researchers, the challenge is to find a tool that can measure progress when each student has different goals, and therefore different milestones for what constitutes success when conducting group comparison studies. The standard procedure, to use nomothetic assessments, which assumes that all members of the group can be measured equivalently on a test and that expected outcomes for all members are similar, is not appropriate for special education or for COMPASS outcomes. That is, a standardized, norm-referenced outcome measure, such as language skills, is not consistent with the demands of special education that each student's curriculum be individualized

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*,
SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2_3

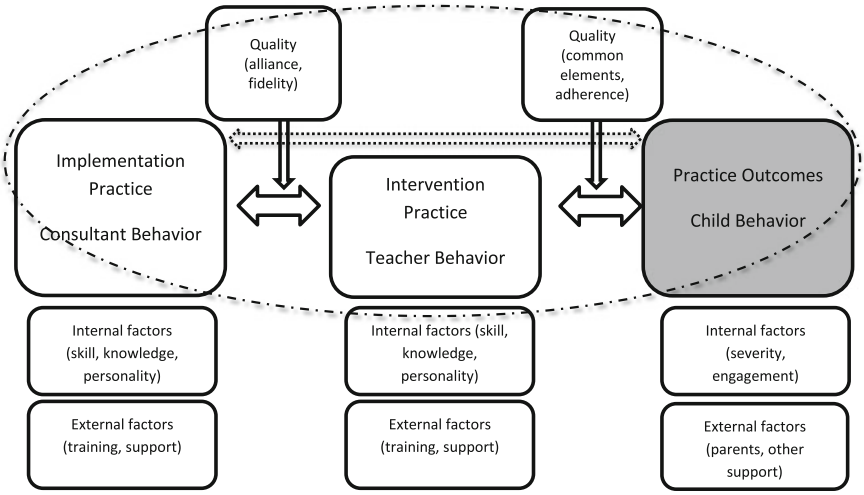
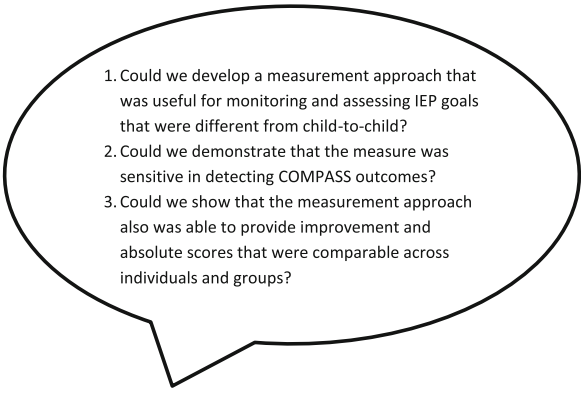


Fig. 3.1 Integrated model with focus on practice outcomes

through the IEP. For example, even children who all have a communication goal, may not have the same communication goal to increase verbal language or sign language or picture symbol usage. Moreover, despite their common use in research for young children with ASD (Reichow 2012), broad constructs like language or IQ are relatively insensitive measures of change since improvement in these areas is typically associated with very young children rather than school-age children. Improvements in standard scores are not anticipated or appropriate IEP goals. IEP goals are personalized and focus on specific changes in behavior that are largely not captured by standard group-based or nomothetic tests. Accordingly, because success in special education, and thus with the COMPASS intervention, is measured by the achievement of student-specific IEP goals, we needed a measure of individual goal attainment as our outcome measure for competence. In creating such a measure, we asked the following questions (see below): (a) Could we develop a measurement approach that was useful for monitoring and assessing IEP goals that were different from child-to-child?; (b) Could we demonstrate that the measure was sensitive in detecting COMPASS outcomes?; and (c) Could we show that the measurement approach also was able to provide improvement and absolute scores that were comparable across individuals and groups, e.g., comparable between control and experimental group participants?



Specifically, as already mentioned, evaluation of the outcomes of special education programs at the student level requires an individualized measure of change, since we are measuring different outcomes for each student. Further, all students start at different baseline levels and require different intervention strategies. The question then is how do we measure the effectiveness of COMPASS when outcomes are different, baseline starting levels are different, and treatment plans are different? As noted above, our answer was to use a sensitive, idiographic approach for measuring individual student outcomes that also ensures psychometric equivalence for group comparisons. Idiographic refers to an individualized assessment approach. Idiographic assessment approaches are not suitable for outcomes that are based on standardized tests, like IQ or other norm-referenced tests. Typically, norm-referenced outcomes, or nomothetic approaches, have been favored in experimental studies because they represent quantitative data collected as mean scores that are readily applied in statistical designs. To answer our questions, our need for an idiographic methodology required that we adapt an individualized approach so that it would be suitable for quantitative data analysis similar to nomothetic approaches. Accordingly, we created a psychometric-equivalence tested approach to goal attainment scaling (PET-GAS) as an outcome measure. Goal attainment scaling is an ideal idiographic method that has been around for decades (Cytrynbaum et al. 1979). It also lends itself well to monitoring IEP goal accomplishment. Figure 3.2 shows a GAS template that is used for establishing goal benchmarks. Although GAS is readily accepted as a technique for measuring individual goal progression, before we could readily adopt and apply the use of GAS for our research, we had to address concerns expressed by researchers who questioned its validity for group design research (Schlosser 2004). Researchers worried that goals set for different individuals may not be equivalent. Some may be harder than others, some may be easier. Similarly, there was concern that progress rating steps may not be equivalently difficult for different individuals. Thus, differences in goal attainment between individuals may not reflect differences in the treatment they received, but differences in how hard or easy it was to achieve the goals.

-2 Present level of performance	-1 Progress	0 Expected level of outcome (GOAL)	+1 Somewhat more than expected	+2 Much more than expected
Baseline performance level	Partial accomplishment of expected goal performance	Goal accomplishment at the end of the school year	Exceeded accomplishment of expected goal performance	Much surpassed accomplishment of expected goal performance

Fig. 3.2 Example of GAS template

Goal attainment scaling helps measure effectiveness in group design research when outcomes are different, when baseline starting levels are different, and when treatment plans are different for each individual.

To respond to these concerns, we created a systematic procedure for developing GAS templates that would result in goals that were measurable and could assess individual change in ways that ensured group comparability (Fig. 3.3). However, we also needed a method to check and ensure that goals created using our GAS templates were equivalent. Thus, we also developed a set of measures to test for baseline equivalence of critical psychometric properties that could inflate or decrease outcomes artifactually if they were not comparable. If we could show that these features were similar between experimental and control group templates at baseline, then we would have evidence of the group comparability needed for statistically valid group analysis. To do this, we evaluated each template using three features: (a) measurability of the goal; (b) level of difficulty of the goal; and (c) equidistance of goal rating steps, that is, that the difficulty in improving performance from one step to the next step is equivalent across goals (see Fig. 3.2).

In our paper *Goal Attainment Scaling as an Outcome Measure in Randomized Controlled Trial x of Psychosocial Intervention x in Autism* (Ruble et al. 2012b), we tested several assumptions of our GAS approach, including the type of scaling level used (i.e., interval or ordinal) and comparison of different measurement sources (live vs. videotaped, teacher vs. researcher) and methods (inter-individual equivalence

GAS Coding Instructions

Instructions: A 3-point Likert Scale is applied for each of the three dimensions described. Each scale has its own anchors.

Level of Difficulty

1	2	3
Not at all Difficult	Somewhat Difficult	Very Difficult

1 – Skill is very close to what the child is already described as able to perform in the present levels of performance

2 – The present levels of performance indicates that the child is able to perform the skill in limited ways compared to what is written in the objective (limited people, prompts, or places); if present level says child has difficulty doing it, score a "2"

3 – The present levels of performance indicates that the child is unable to perform skill with anyone, anywhere, or with any prompts compared to what is written in the objective.

Measurability

1	2	3
Not at all Measurable	Somewhat Measurable	Clearly Measurable

1 – None or only one indicator (prompt level, criterion for success; observable skill) is listed

2 –Two of the three indicators (prompt level, criterion for success; observable skill) are provided

3 – Describes all three indicators (prompt level, criterion for success; observable skill)

Equality

1	2	3
Not at all Equal	Somewhat Equal	Clearly Equal

1 – None or only 1 of the four levels (-2, -1, 1, 2) include reference to the appropriate prompt hierarchy (moving from physical to independent/visual support) and/or the skill frequency increases or redu ces in roughly equal amounts (e.g., moving from 3 to 6 to 9 occurrences for -2, 0, 2)

2–Two out of the four levels (-2, -1, 1, 2) include reference to the appropriate prompt hierarchy (moving from physical to independent/visual support) and/or the skill frequency increases or reduces equal amounts.

3 – At least three of the four levels (-2, -1, 1, 2) include in reference to the appropriate prompt hierarchy (moving from physical to independent/visual support) and the frequency of skill increases or reduces by equal amounts for the majority

Fig. 3.3 GAS equivalence measure

and comparability, and reliability of coding across different behavioral observation methods—video-taped or teacher made video-tape). Group comparability of GAS descriptions generated for individual outcome measurement was tested using data from our randomized trial (i.e., measurability, equidistance, level of difficulty, comparability of behavior samples collected from teachers vs. researchers and live

vs. videotape). We developed detailed guidelines for developing GAS descriptions and were able to verify that psychometrically equivalent GAS descriptions can be created and can be evaluated for equivalency, that teacher collected behavior samples are representative and comparable to researcher collected observations, and that varied sources of behavior samples can be reliably coded. More details on how to create psychometric-equivalent GAS templates and to test the equivalence of the templates between groups is provided in our manual (Ruble et al. 2012a). Essentially, we found that we could develop templates with reliability and that were comparable across individuals and in terms of progress difficulty. That is, we were able to show that goals created using PET-GAS both retained the idiographic advantage of individualization and possessed the "nomothetic" quality of providing equivalent indications of progress when comparing individuals or groups. We also confirmed that teacher-made tapes that were used to score child progress during the coaching sessions were similar to tapes collected by researchers. In other words, the GAS scores from teacher-made tapes were similar to the GAS scores obtained from researcher-collected tapes. Overall, our data suggests that GAS can be a valid and sensitive idiographic approach for use in group design experimental studies (Fig. 3.3).

References

Cytrynbaum, S., Ginath, Y., Birdwell, J., & Brandt, L. (1979). Goal attainment scaling: a critical review. *Evaluation Quarterly, 3*(1), 5–40.

Reichow, B. (2012). *Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders* (Vol. 42). Germany: Springer.

Ruble, L. A., Dalrymple, N. J., & McGrew, J. H. (2012a). *Collaborative model for promoting competence and success for students with ASD*. New York: Springer Science + Business Media.

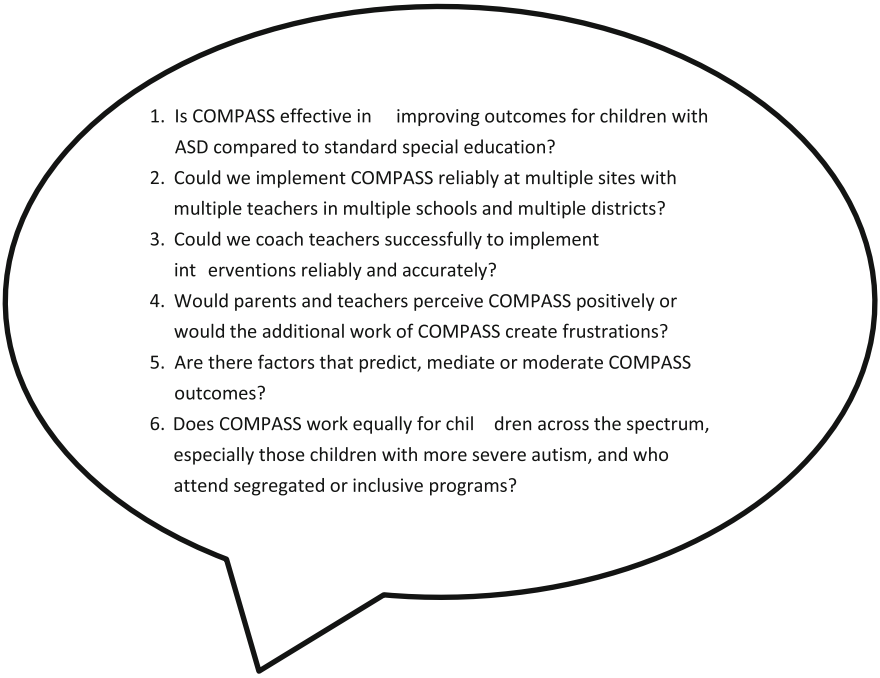
Ruble, L., McGrew, J., & Toland, M. (2012b). Goal attainment scaling as an outcome measure in randomized controlled trials of psychosocial interventions in autism. *Journal of Autism and Developmental Disorders, 42*(9), 1974–1983.

Schlosser, R. W. (2004). Goal attainment scaling as a clinical measurement technique in communication disorders: a critical review. *Journal of Communication Disorders, 37*(3), 217–239. doi:10.1016/j.jcomdis.2003.09.003.

Chapter 4
Efficacy of COMPASS

In this chapter we discuss our two NIH-funded RCTs of COMPASS that are also reported in detail in our published papers (Ruble et al. 2010, 2013) and a book (Ruble et al. 2012). But before we describe the unique aspects of each study below, we want to discuss our rationale for our approach. For our first study, our goal was to conduct a proof of concept study. In other words, does COMPASS help children achieve higher educational outcomes compared to those children who receive special educational services as usual? Our outcomes were based on goal attainment scaling methods described in the previous chapter. Using GAS as our outcome, we began with a small sample of participants and focused on the younger children because we had more information about evidence based practices for this age group that could be used to guide intervention development. In addition to establishing whether or not COMPASS works, we had several other goals and questions for this first study. In the call out figure below, we outline our primary effectiveness question along with our five additional key questions for the first study: (a) Could we implement COMPASS reliably at multiple sites with multiple teachers in multiple schools and multiple districts?; (b) Could we coach teachers successfully to implement interventions reliably and accurately?; (c) Would parents and teachers perceive COMPASS positively or would the additional work of COMPASS create frustrations?; (d) Are there factors that predict, mediate or moderate COMPASS outcomes?; and (e) Does COMPASS work equally for children across the spectrum, especially those children with more severe autism, and who attend segregated or inclusive programs? The first two questions are addressed in this chapter. The other questions about teacher and parent satisfaction and acceptance of COMPASS are addressed in Chap. 5. The fourth question about mediators or moderators is discussed in Chaps. 6 and 7 about intervention quality and teacher factors that impact child outcomes. The last question concerning whether or not COMPASS helps children regardless of level of ASD severity is addressed in Chap. 8.

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*,
SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2_4



To help answer our questions effectively, we applied the same sampling, recruitment, measures, and randomization procedures for both RCTs, which then allowed us to combine samples to answer secondary research questions (such as the impact of teacher burnout on student outcomes). Both studies also targeted the same groups, special education teachers who were responsible for the Individual Education Programs (IEP) of students with autism aged 3–8 in public schools. About half of the teachers were the primary classroom teacher of the child and the other half were the child's resource teacher or general education support teacher. All were the primary person responsible for the implementation of the IEP. Sampling and methods were similar across the studies. Teachers were asked to participate at the start of the school year (Time 1). Those who agreed were then asked to provide the initials of the students with ASD they taught to maintain confidentiality. We selected at random one student and then asked the teacher to forward information about the study to the parents and caregivers of the selected student. After parents agreed to participate and a comprehensive baseline evaluation was completed for each teacher–child dyad, the dyads were randomized to the control or experimental conditions. To determine the impact of COMPASS on child educational goal attainment at the end of the school year, a Time 2 evaluation was

completed using the same measures applied at the start of the school year. To ensure objective and independent assessment, an independent evaluator blind to participant assignment was used to judge child progress on IEP goals.

With one exception, which is described below in the discussion of the individual studies, the COMPASS intervention was implemented identically in both studies. For both studies, teachers in the experimental conditions received the initial COMPASS consultation and four follow-up coaching sessions. The initial consultation included the teacher and parent and lasted approximately 2.5–3 h. Each coaching session lasted between 60 and 90 min and occurred about every 4–6 weeks and in total was less than 10 h across the school year. The Fig. 4.1 describes the activities of the initial consultation and the activities that occur during each coaching session.

The key outcome measure for both studies was Goal Attainment Score (GAS) change from baseline to end of the study. We chose GAS as our primary outcome measure for several reasons as discussed in the prior chapter. Specifically, it allowed us to assess intervention outcomes for group design research when children started at different baseline level of skill, had different goals, and had different intervention plans. This idiographic method has been applied in numerous studies and is described in detail in Chap. 3.

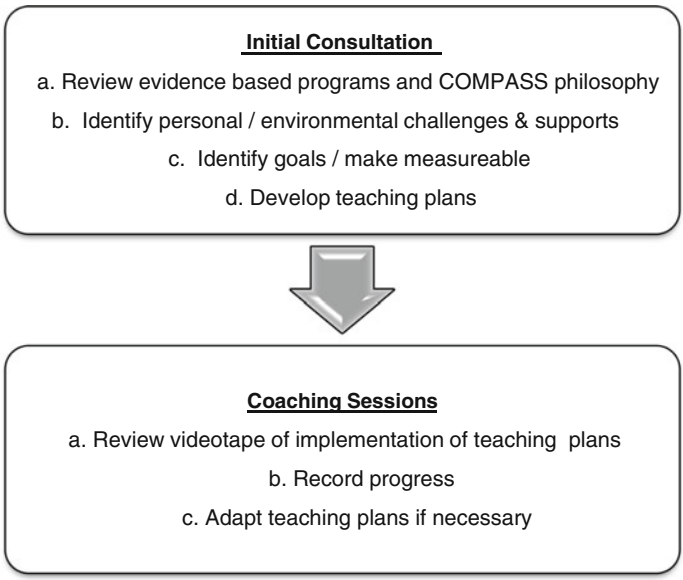


Fig. 4.1 COMPASS initial consultation and follow-up coaching

Description of Two RCTs

Study 1 (Ruble et al. 2010)

As noted earlier, for the first RCT, the primary goal was to establish proof of concept. For this study, the mean age of the children was 6.1 years and they represented children across the autism spectrum who attended special educational classrooms full time and part time or who were educated in general education classrooms. Children were recruited based on an autism diagnosis. No exclusion criteria, other than no sensory impairments (hearing or visual) were made. Teacher-student dyads were randomized into COMPASS or into a comparison group consisting of services as usual. Figure 4.2 shows the sequence of events in carrying out the study. Once the teachers, parents, and children completed baseline measures, teachers were randomized to a control or COMPASS group.

To validate that the randomization procedure worked as intended, child age, autism severity, IQ, language ability, and adaptive behavior were compared between the control and COMPASS group; no significant differences were found. Overall, the COMPASS group teachers received a little less than 10 h of consultation from the researchers. The control group students received their special education program as usual. The students in the COMPASS and control groups had goals that reflected a social skill, a communication skill, and a learning or independent work skill. To determine the amount of progress children made in their IEP goals, at the end of the school year, child goal attainment change, as measured by final goal attainment scores minus beginning of the year goal attainment scores, were collected by an observer who was independent of the research team and unaware of group assignment. *The overall results show that students whose teachers received COMPASS had GAS scores that were significantly higher than those in the control group. Students whose teachers received COMPASS made a*

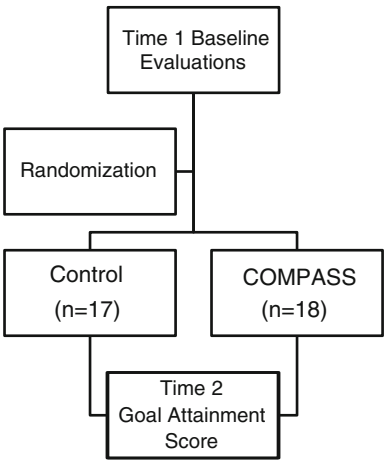


Fig. 4.2 Research design for study 1

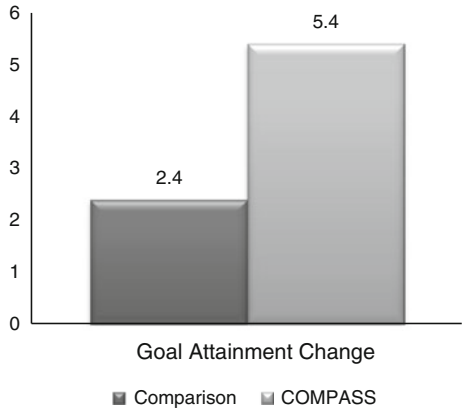
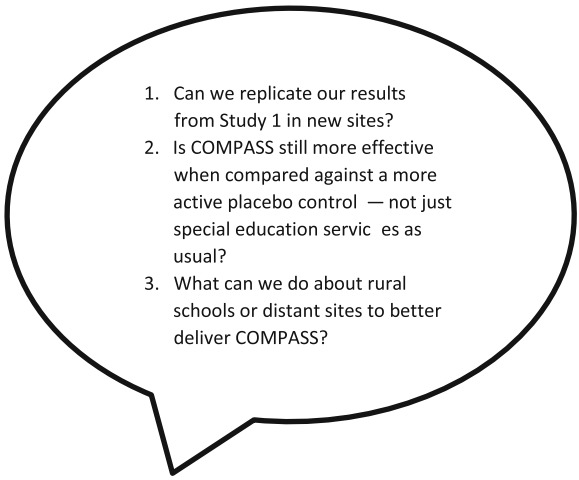


Fig. 4.3 Study 1 GAS outcomes

1.5 standard deviation improvement compared to students who were in the control group (see Fig. 4.3). After controlling for Time 1 GAS scores, there was a statistically significant group difference in change from pretest to posttest scores, $F(1, 29) = 11.08$, $p = 0.002$, indicating greater improvement in scores for children in the experimental group ($M = 5.4$) relative to the control group ($M = 2.4$). Thus, we were able to clearly answer our first question: COMPASS was effective.

Study 2 (Ruble et al. 2013)

Based on the positive results from Study 1, we were able to ask new questions that could be tested in a different study. For study 2, we asked three new primary questions outlined to the right: (a) Can we replicate our results from Study 1 in new sites?; (b) Is COMPASS still more effective when compared against a more active



placebo control—not just special education services as usual?; and (c) What can we do about rural schools or distant sites to better deliver COMPASS? These questions led to a second study. Question one is the replication question, addressing the core issue of whether we can confirm COMPASS effectiveness in a second independent sample and relates to step 3 of the Evidence Ladder (Chap. 1; Fig. 1.2). Question two expands this effectiveness question to add a more active control condition. That is, in study one the comparison was to standard special education, also sometimes referred to as treatment-as-usual comparison. For study two, we wanted to add a more active treatment comparison in addition to standard special education. Question three focused on an important implementation concern, how to provide consulting to distant sites, when travel and face-to-face interventions tend to be difficult. To address question 3 specifically, we added a second experimental condition that tested web-based videoconferencing as a means for coaching classroom teachers. We also asked a set of secondary questions that focused specifically on the WEB condition. We outline these questions after first presenting the answers to our primary questions.

As mentioned, we applied the same baseline and randomization procedures used in the first study to the second RCT. Teacher-student dyads were randomized into one of three groups: (a) COMPASS delivered face-to-face (FF), (b) COMPASS delivered via the web (WEB) (initial consultation was face-to-face but the coaching was provided using a web-based videoconferencing program called Adobe Connect), or (c) a comparison group which received free online professional development training in three evidence-based teaching methods in autism. See Fig. 4.4.

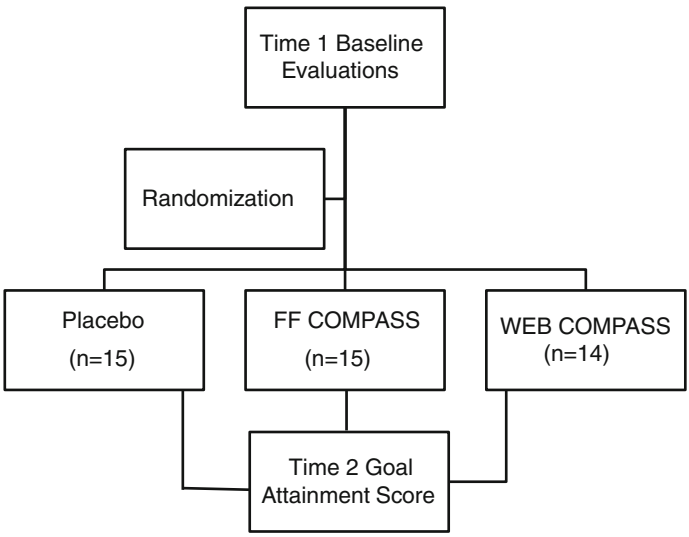


Fig. 4.4 Research design for study 2

In choosing the specific activity for the active placebo that served as our comparison group, we had two clear goals: (1) that the participants would view the online training as plausibly helpful and (2) that the training, although informative, was unlikely to produce impacts on outcomes, i.e., placebo. We asked teachers to watch videos and read materials on three commonly used teaching strategies for children with ASD: (a) structured teaching, (b) peer mediated training, and (c) the picture exchange communication system. Structured teaching is useful for instructing children on independent work skills (Mesibov et al. 2012). Peer-mediated training has been found to be useful in enhancing social interaction skills (Garfinkle and Schwartz 2002). The Picture Exchange System is an evidence based communication system to help non-verbal or minimally verbal children to learn to initiate requests for various wants and needs (Howlin et al. 2007). We considered this group a placebo control because research shows decontextualized training provided without self-reflection and performance-feedback does not result in changes in teacher behavior (Joyce and Showers 2002; Trivette et al. 2009).

As in Study 1, analysis of the similarity between the three groups indicated that the randomization procedure worked as planned with the exception of one child variable, IQ. For all three groups, children had similar levels of autism severity, language, adaptive behavior, and age. Analysis of teacher variables also indicated similarity in experience based on number of years of teaching and numbers of students with ASD taught across all three groups. In addition, the children received similar numbers of services and hours of services outside the home. However, IQ was lowest for the WEB group (M = 44.6) compared to both the placebo (M = 61.3) and FF group (M = 60.9). Because children in the WEB group had more severe intellectual disability, uncorrected group comparisons would likely underestimate effectiveness of the WEB group, thus in our statistical analysis we adjusted outcomes based on this difference of IQ across groups.

Can we replicate our results from Study 1 in new sites and is COMPASS still more effective when compared against a more active placebo control—not just special education services as usual? The results from Study 2 replicated and also extended findings from our first study and demonstrated step 3 of our Evidence Ladder (Chap. 1, Fig. 1.2). Nearly identically to our first RCT, children whose teachers received FF COMPASS made a 1.4 standard deviation improvement compared to the placebo control children ($t = 4.25$, $p < 0.001$; Fig. 4.5). Thus, (1) we were able to reliably implement and deliver COMPASS in both study one and study two, indicating that COMPASS should be disseminable to multiple sites, and (2) we were able to replicate similarly strong effect sizes in both studies, and importantly in study two we did so against a more stringent comparison group, an active placebo control.

What can we do about rural schools or distant sites to better deliver COMPASS? To understand the impact that web-based videoconferencing made on COMPASS outcomes, we analyzed results of the WEB group against the face-

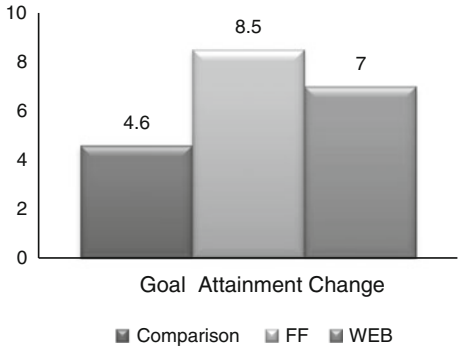
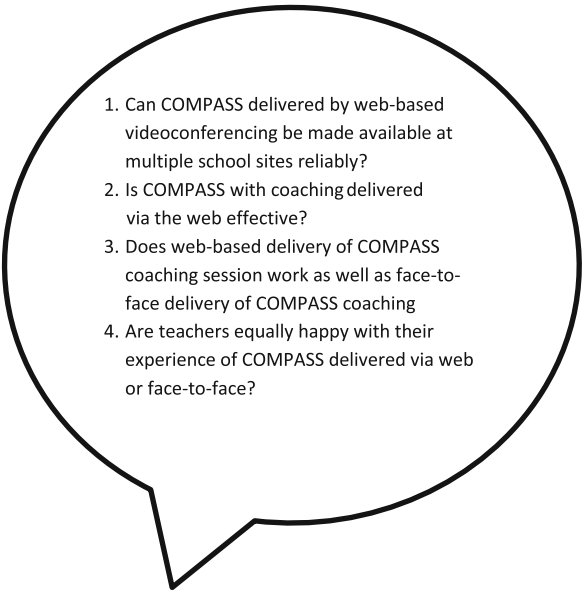


Fig. 4.5 Study 2 GAS outcomes

to-face and placebo groups. We had many unanswered questions about video conferencing as a coaching method. Would web-based coaching be equally effective to face-to-face coaching or would our sample be too small to detect a difference? We were pleasantly surprised to find that the WEB group also made significant and large improvements and that the effect was not statistically different from the face-to-face group ($t = 0.80$, $p = 0.22$). Specifically, children in the WEB group made a 1.1 standard deviation improvement compared to the placebo group ($t = 3.3$, $p = 0.001$) after controlling for differences in IQ. The viability of web-based COMPASS is important, because it demonstrates that similar outcomes can be achieved regardless of mode used for teacher coaching. For schools located in rural areas or in densely populated urban areas where distances can be a barrier, consultant time spent traveling can be reduced and time spent coaching can be enhanced (Fig. 4.5).

As mentioned above we also pursued a set of secondary questions that focused specifically on the WEB condition. Specifically we wanted to know.
(a) if we could provide COMPASS delivered by web-based videoconferencing reliably across multiple schools and classrooms; (b) if the experience teachers had with web-based COMPASS coaching had any negative effects on satisfaction; and (c) if teachers in the WEB condition were able to implement intervention plans with similar quality of fidelity to teachers in the face-to-face condition? Because these questions focus on implementation issues we will discuss our answers in the following chapters that focus more closely on implementation concerns. As a preview, suffice it to say that we were able to provide positive answers to each question.



Finally, now that we know that we can deliver COMPASS using web-based technology with good efficacy, additional questions have emerged. Does efficacy depend on having prior face-to-face contact, which provides the opportunity to establish rapport and alliance needed for the web-based approach to work? Recall that all teachers and parents participated in an initial face-to-face consultation that involved discussion of best practices, goals of COMPASS, the child's COMPASS profile, identification of a social, communication, and learning skills goal, and development of instructional plans. Could a web-based only approach work for the assessment and intervention planning stage of COMPASS? Would other web technologies that would allow group discussions or group supervision be helpful for enhancing teacher's implementation of intervention plans, problem solving, and self-efficacy when teachers can learn from other teachers who are working with different students? These questions will have to wait to be answered by future research.

In the next chapter, we begin to focus not on whether COMPASS works, but on why it works. Up to this point, we established proof of concept that COMPASS is effective. We also demonstrated that COMPASS is not dependent on mode of delivery and works well when given face-to-face or using web-based technology. Our next set of questions concerned why. What is it about COMPASS that makes it effective? This leads to a discussion of active ingredients and the potential impact of consultant, teacher, and child variables on outcomes.

References

Garfinkle, A. N., & Schwartz, I. S. (2002). Peer imitation increasing social interactions in children with autism and other developmental disabilities in inclusive preschool classrooms. *Topics in Early Childhood Special Education, 22*(1), 26–38.

Howlin, P., Gordon, R. K., Pasco, G., Wade, A., & Charman, T. (2007). The effectiveness of picture exchange communication system (pecs) training for teachers of children with autism: a pragmatic, group randomised controlled trial. *Journal of Child Psychology and Psychiatry, 48*(5), 473–481.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Mesibov, G. B., Shea, V., & McCaskill, S. (2012). Structured teaching and the TEACCH program. In D. Zager, M. L. Wehmeyer & R. L. Simson (Eds.), *Educating Students with Autism Spectrum Disorders* (pp. 99–112). New York: Routledge.

Ruble, L. A., Dalrymple, N. J., & McGrew, J. H. (2010). 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, 32*(4), 286–301.

Ruble, L. A., Dalrymple, N. J., & McGrew, J. H. (2012). *Collaborative model for promoting competence and success for students with ASD*. New York, NY: Springer Science + Business Media.

Ruble, L. A., McGrew, J. H., Toland, M. D., Dalrymple, N. J., & Jung, L. A. (2013). *A randomized controlled trial of COMPASS web-based and face-to-face teacher coaching in autism* (Vol. 81). US: American Psychological Association.

Trivette, C. M., Dunst, C. J., Hamby, D. W., & O'Herin, C. E. (2009). *Characteristics and consequences of adult learning methods and strategies* (Vol. 2). Asheville, NC: Winterberry Press.

Chapter 5
COMPASS Implementation Quality

Consultation as a Goal Setting Activity in COMPASS: Analysis of the Communicative Exchanges Between Consultant and Consultees

As described more fully in Chap. 2, the COMPASS intervention is comprised of two distinct, but related implementation practice activities– the initial consultation and the follow-up teacher coaching activities. Both are designed to encourage and empower parents and teachers as the active decision-makers on behalf of a student as well as to assist them in their efforts to directly intervene with the student using performance-based outcome assessment (i.e., timely, interim feedback about student goal progress) and empirically informed decision-making. Now that we have established that COMPASS works (Chap. 4) in two separate rigorous controlled trials we are ready to start asking how and why it works. To answer these questions we will be using our Integrated Model (Fig. 5.1). Recall that this model outlines the factors thought to impact student outcomes. There are three primary actors, the consultant, who is the enactor of the implementation practice (COMPASS), the teacher, who is the enactor of the intervention practice (the EBPs matched to the student's individual goals) and the student. The behavior of each of these actors in turn, is impacted by internal and external factors. The character of the relationship between the actors is modeled in terms of quality elements (e.g., fidelity to COMPASS). The focus of this chapter is on the consultant behaviors, shaded in gray, and what we have learned about quality consultation, its impact on teacher behavior, and whether there is any impact on practice outcomes, via teacher behaviors.

Evidence of implementation quality includes what transpires between consultant and consultee. Fidelity is a critical aspect of implementation/intervention quality. The most common definition of fidelity is the degree to which a given intervention is implemented as intended (Moncher and Pribnz 1991; Yeaton and Sechrest 1981). Recall that fidelity is a critical aspect of intervention quality. In education, fidelity measurement has been typically assessed using two broad dimensions described as

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*,
SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2_5

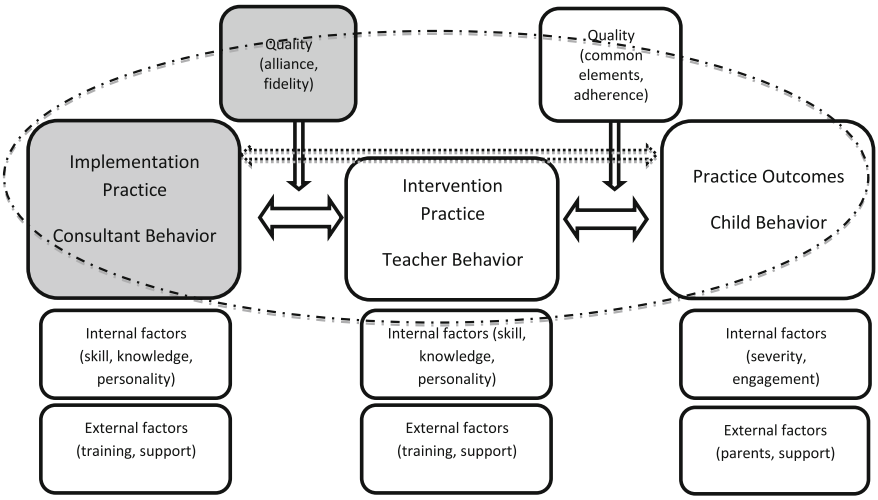


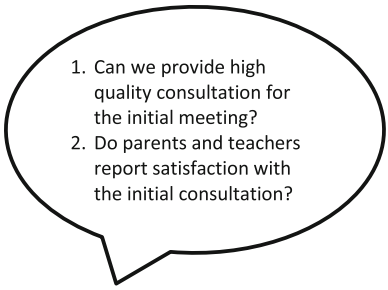
Fig. 5.1 Integrated model with focus on implementation practice quality

(a) structural fidelity and (b) process or procedural fidelity (Gersten et al. 2005; Odom 2009; O'Donnell 2008). Structural fidelity refers to the organizational/structural elements that characterize and describe the intervention and are somewhat dichotomous in nature in terms of measurement, in that they occur or don't (e.g., IEP meeting occurred, goals were set, teacher meets daily with student, consultant met with teacher four times during the year) (Durlak and DuPre 2008; Gersten et al. 2005). On the other hand, process or procedural elements refer to the measurement of the actual behaviors of the teacher or intervener and tend to occur along a continuum (teacher employed joint attention in interaction). Both impact quality and are part of its measurement, and we applied both types of measures in our studies.

Initially we focused on what occurred during the first step of COMPASS, the opening consultation. Also recall in the EBPP framework the importance of clinical decision-making in integrating the overlapping areas of the EBP, the characteristics of the child (preferences, strengths), and the characteristics of the teacher (training, knowledge). Step one of COMPASS, the initial consultation, provides the platform for this clinical decision-making to occur. In the chapters that follow, we will review what we have learned about the intervention practice and practice outcomes. First, as shown in our Integrated Model, we examined how the implementation practice influences the intervention practice directly and also indirectly as a result of continuous performance monitoring and feedback from the practice outcomes.

As discussed in earlier chapters, the initial consultation is a 3-h goal-directed, parent-teacher, activity with three specific aims: (a) to develop a shared understanding of the child's personal and environmental supports and challenges using the COMPASS competency framework (Ruble and Dalrymple 1996, 2002); (b) to use this information for identifying and generating high quality personalized teaching goals; and (c) to develop personalized teaching plans based on the

information obtained in steps a and b for each goal. To guide our analysis of the quality of the initial consultation, we had the following questions: (a) Could we deliver high quality consultation reflected by positive exchanges between parents, teachers, and consultants? and (b) Do parents and teachers report that the initial consultation was high quality and that they were satisfied? The first question examines consultant quality using elements non-specific to COMPASS and thought to underlie good consulting generally, the second question examines quality in terms of adherence to specific desired elements of COMPASS.



Can we deliver high quality COMPASS consultation? Recall that the quality of the consultant-teacher interaction is a function of factors characteristic of general good consulting practice and of factors specific to COMPASS. Research on general elements of good consultation shows that effective processes and strategies used during consultation are based on collaboration, mutual respect, and parity among all team members (Brown et al. 2011). When consultees have input into identifying the problem and goals, they are more likely to be engaged throughout the entire process. This is critical because the real work happens once the consultant leaves. To understand how well the COMPASS consultants engaged teachers and parents as part of the process during the initial consultation, we evaluated a general element of good quality consultation or implementation quality—communication. Specifically, we evaluated the communicative exchanges that occurred during the critical stage of discussing the child's COMPASS profile with the aim of developing a shared understanding of the child's personal and environmental challenges and supports. Although the quality of communicative exchanges is important throughout COMPASS consultation and coaching, we focused on this first stage-setting activity to formally evaluate communication quality using an objective methodology. During the stage-setting activity, the COMPASS profile is reviewed by the parent and teacher and covers nine core areas of learning and development observed at home, in the community and at school: (a) likes, strengths, frustrations and fears, (b) adaptive skills, (b) problem behaviors, (c) social and play skills, (d) communication, (e) sensory preferences and dislikes, (f) learning skills, (g) environmental supports, and (h) environmental challenges. An example of the social and play skills assessment form used to guide discussion of the child's skills at home and at school, with adults and with children is provided in the Fig. 5.2. Our manual on COMPASS provides transcribed examples of the discussion during this activity and also detailed case studies (Ruble et al. 2012).

4. Social and Play Skills

Directions: Please rate the following statements on a scale of 1–4, with 1 meaning "not very well" and 4 meaning "very well." Please answer each question first in terms of the child's interactions with adults, and then with children.

How well does the child/student	With adults				With children			
	Not well			Very well	Not well			Very well
Social awareness								
1. Look toward a person who is talking to him/her	1	2	3	4	1	2	3	4
2. Accept others being close to him/her	1	2	3	4	1	2	3	4
3. Watch people for extended periods of time	1	2	3	4	1	2	3	4
4. Respond to another person's approach by smiling or vocalizing	1	2	3	4	1	2	3	4
5. Initiate interactions for social reasons	1	2	3	4	1	2	3	4
Joint attention skills								
6. Look at something another person points to	1	2	3	4	1	2	3	4
7. Show something to a person and look for person's reaction	1	2	3	4	1	2	3	4
8. Point at an object or event to direct another person's attention to share enjoyment	1	2	3	4	1	2	3	4
9. Share smile by looking back and forth between object and person	1	2	3	4	1	2	3	4
Imitation								
10. Imitate sounds another person makes	1	2	3	4	1	2	3	4
11. Imitate what another person does with an object (such as a person makes toy airplane fly, the child repeats action)	1	2	3	4	1	2	3	4
12. Imitate body movements of others (such as clap when others clap, play Simon Says)	1	2	3	4	1	2	3	4
13. Imitate and expand upon other's actions with toys (such as peer beats drum, child beats drum and also starts to march)	1	2	3	4	1	2	3	4

Fig. 5.2 Example of COMPASS assessment

Using data from our first randomized controlled single-blind trial of COMPASS (Ruble et al. 2010), we analyzed the verbal interactions between the consultant, teacher, and parents during the initial consultation expressed as speech acts and speech act exchanges (Ruble et al. 2011). Speech acts were defined as "a phrase or utterance, bounded by intonation, pauses, or grammar" (Sheridan et al. 2002, p. 311). Speech exchanges, on the other hand, represent the impact of a preceding speech act on the subsequent speech act among the different participants in the consultation. We applied the Psychosocial Processes Coding Scheme (PPCS) developed by Leaper (1991) which allowed us to examine the reciprocal influence and conversational intent and function of participants' speech acts. Three types of speech acts identified were: (a) affiliative or positive speech acts, (b) distancing or negative speech acts, and (c) mixed, a combination of positive and negative speech acts. The speech acts were then transformed into one of three speech exchanges: affiliative, distancing, or mixed. An affiliative speech exchange, involves interactions in which one speaker's collaborative or obliging speech act is followed by another speaker's collaborative or obliging speech act (e.g., a consultant's statement "This sounds like an issue related to lack of understanding of perspective taking" followed by a teacher's statement "I agree"). The second type of speech act exchange, distancing, involves interactions in which one speaker's controlling or withdrawing speech act is followed by another speaker's controlling or withdrawing speech act (e.g., a consultant's statement "You really should do this instead …" followed by a teacher's statement "I don't agree with that…"). The final speech act exchange, mixed, involves exchanges in which one speaker's affiliative statement is followed by another speaker's distancing statement, or vice versa (e.g., a consultant's statement "Let me understand what you are saying" followed by a parent's statement "No, I don't think that you do understand").

Positive verbal interactions between the consultant, parent, and teacher correlated with IEP Quality. IEP Quality correlates with child educational outcomes.

The three categories of speech acts are not equally desirable; affiliative acts are seen as ideal and distancing acts are not. Therefore, we hoped to find more affiliative acts and fewer distancing ones. We had a basic research question: Does COMPASS facilitate good consultant-teacher communication indicative of good quality as shown by high numbers of affiliative acts and low numbers of distancing ones? To answer our question, we analyzed a total of 13,826 speech acts and 9,310 speech exchanges from 18 COMPASS consultations. As we had hoped, and consistent with our research hypothesis that consultant-teacher communication would be high quality, the overwhelming majority of speech act exchanges were affiliative (93.6 %). Very few speech acts were coded as withdrawing and controlling (<2 % each).

Intercorrelation analysis showed that speech exchanges were significantly associated with one another in expected ways. Affiliative speech exchanges were negatively associated with both distancing and mixed exchanges, $r = -0.63$, $p < 0.001$, and $r = -0.99$, $p < 0.001$, respectively, whereas distancing and mixed exchanges were positively associated, $r = 0.57$, $p < 0.001$. Thus, these analyses were able to show that, as hoped, consultants' behavior was judged of high quality, as measured using general indicators of good communication.

We also wanted to know how parents and teachers perceived the initial consultation and if they report that we implemented the initial COMPASS consultation with good quality. Specifically, did we implement COMPASS with high fidelity and were parents and teachers satisfied? Analysis from study 1 showed that fidelity and satisfaction were high (Ruble et al. 2010). Teachers reported 96 % of the elements of COMPASS were implemented. For satisfaction, they reported a mean score of 3.7 from a total of 4, indicating high satisfaction. These results together indicate that the initial consultation was provided with high quality as measured using indicators specific to COMPASS. In study 2, we repeated the measures and replicated the findings. For fidelity, teachers reported that 92 % of the elements were implemented. Further, for satisfaction teachers reported a mean score of 3.6 out of a total of 4, indicating high satisfaction. Thus, we had consistent evidence of high quality implementation.

Next, we wondered if there was an association between our two measures of COMPASS consultation implementation quality, thus, we correlated scores from the COMPASS fidelity checklist with scores from the teacher satisfaction measure and found a significant correlation between the two ($r = 0.41$, $p = 0.03$). The correlation or overlap between the measures of quality provides evidence of convergent validity, which is one of the key indicators of construct validity. That is, our two measures were assessing related but also slightly different aspects of the same construct, consultant quality.

These findings led to a different question that emerged from our Integrated Model (Fig. 5.1). Specifically, we asked whether good quality implementation of the initial COMPASS consultation had any direct impact on intervention practice, i.e., teacher behavior? To answer this question, we correlated affiliative speech exchanges with one of the immediate expected outcomes from the initial consultation—IEP quality. Within the COMPASS model, IEP quality is a measure of intervention quality, in that improved IEP quality indicates that the teacher changed the IEP based on recommendations made following the consultation. Although we provide more description in the following chapter, briefly, IEP quality expected to change as a result of COMPASS was measured by evidence of best practice goals, as indicated by inclusion of goal targets recommended by the literature (i.e., a social goal, a communication goal, and independent work behavior goal) and by evidence that goals were well specified as indicated by their clarity, objectivity, and measurability. These were the targeted IEP elements we expected to change as a result of COMPASS. That is, we purposely identified a goal for each of the recommended domains and worked with the teacher and parent to make sure goals were clear, observable, and easily measured. Our results showed a direct and positive correlation between IEP quality and affiliative exchanges ($r = 0.51$, $p < 0.008$) and negative correlations between IEP quality and both distancing ($r = -0.49$, $p = 0.03$) and mixed speech exchanges ($r = -0.49$, $p = 0.03$). That is, the style and character of the communication between participants during the initial consultation was strongly associated with the quality of the goals generated. An affiliative, positive communication style was more likely to create a consultation environment leading to the production of high quality goals, than a communication style that was distancing or did not use predominantly affiliative exchanges. Good communication is one characteristic of a good working alliance and in this respect, these data are consistent with the general literature on the importance of therapeutic alliance on treatment process and outcomes.

Moreover, as discussed in Chap. 6, we believe IEP quality is one of the active ingredients of COMPASS. In fact, as we will show, IEP quality also correlated with end of the year student goal attainment outcomes in both of our RCTs. Thus, high quality implementation, as measured by process or relationship variables during COMPASS, directly sets the stage for high quality intervention practice or teacher behavior, as indicated by IEP quality. Overall, these results provide convincing evidence that COMPASS, when done well, sets the stage for encouragement and supportive positive and helpful interactions between consultant, teacher, and parents. Specifically, positive speech acts facilitate the quality of the teaching goals (IEP), which in turn are related to positive child outcomes.

We have not yet analyzed the relationships between our COMPASS specific measures of the quality of the consultant's behavior during the initial consultation (e.g., teacher ratings of consultant adherence) and consultant teacher behavior, such as IEP quality. Those analyses if also positive, would provide additional evidence for the critical relationships between implementation quality and intervention quality. We next examined our key questions concerning implementation quality when assessed for the coaching sessions.

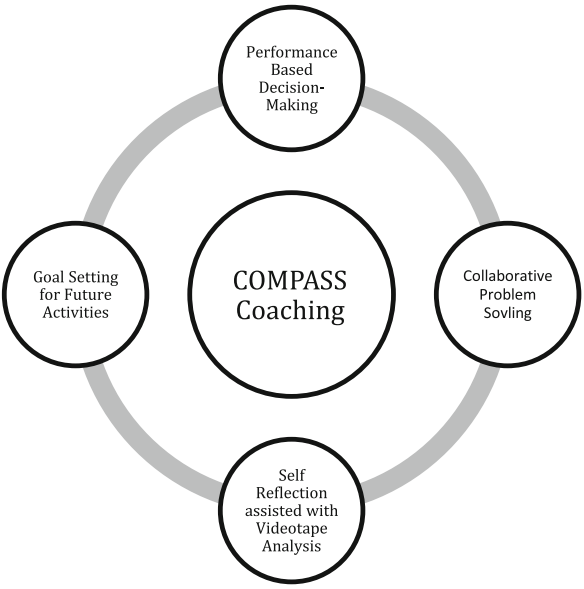


Fig. 5.3 Elements of coaching

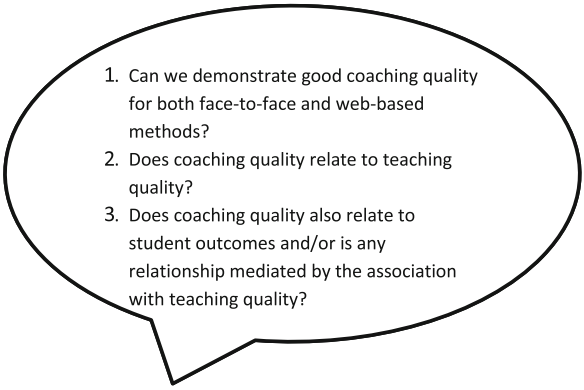
Coaching Quality. To this point we have provided evidence of COMPASS implementation quality of the initial consultation and of the impact of this quality on teacher/intervention quality and satisfaction. We also were interested in consultant coaching quality and its potential impact on teacher behavior. Although the action of coaching technically fits within the definition of consultation, we differentiate coaching as a related, but distinct activity. We define coaching as the subsequent consultant-teacher meetings following the initial consultation meeting. Recall that the initial consultation sets the stage for work with the student by identifying the personalized teaching goals and intervention plans developed based on discussion of the COMPASS profile. The coaching activities involve the follow-up teacher support necessary to ensure high quality and accurate implementation of the intervention. We believe one key feature and critical impact of coaching is to improve teacher competence with, and adherence to the teaching plans using a collaborative, problem-solving and performance-based coaching approach. Coaching within COMPASS includes several distinct activities whose implementation was guided by a semi-structured interview available in the manual (Ruble et al. 2012). Figure 5.3 shows the different elements that underlie COMPASS coaching as defined by the semi-structured interview process. We describe these elements next.

Performance-based monitoring refers to consultant activities in assessing and providing feedback to the teacher about (1) student progress on the goals, as measured by goal attainment scaling, and (2) his/her implementation of the specified teaching plan. Student progress is assessed both against the pre-specified goal

attainment target (i.e., score of 0 on the GAS) and against prior goal attainment scale scores obtained at each coaching session.

The second set of coaching activities involves collaborative problem solving and self-reflection assisted with videotape analysis between the consultant and teacher. This step is actualized by a detailed joint review and observation using videotape analysis. Additional elements include teacher self-reflection about his/her interactions with the student since the last coaching session and identification of potential problems and potential solutions to enhance this interaction during instruction. Based on the results of the collaborative problem solving, the next step is to set goals for future teacher (and student) activities.

Goals for future planning as well as specific strategies to achieve them are then identified. Coaching goals adhere to and advance the larger goals identified at the initial consultation and refer to the priority activities needed to address problem areas. The teacher is expected to work on these interim goals following the coaching session and prior to the next session. These activities form the key elements that take place during coaching. Together we believe these elements of coaching, i.e., a collaborative, problem-solving and performance-based approach, improve teacher competence with, and adherence to the teaching plans.



Assessment of Coaching Quality. We had several questions about coaching quality: (a) Can we demonstrate good COMPASS coaching quality? (b) Does implementation quality, operationalized as coaching quality, relate to intervention quality? and (c) Does implementation quality, operationalized as coaching quality, also impact student outcomes? In order to answer our questions, we needed to assess coaching quality or how well teachers perceived coaching sessions. The coaching quality measures we used were based on COMPASS and were not general measures of good coaching quality. Also, our measures were crude. Ideally we would have liked to have videotaped consultant coaching and then rate the session against a quality checklist.

Accordingly, for our second RCT, we developed two measures of coaching quality—a questionnaire to obtain teacher feedback on their satisfaction with the

Table 5.1 Coaching quality measures

Quality measures	Description	Method
COMPASS coaching fidelity checklist	A 16-item, 4-pt Likert scale (1 "not very much", 4 "very much") questionnaire that asks whether particular activities occurred during the coaching session and about the tone set by the consultant	Teacher ratings
COMPASS coaching feedback form	An 11-item, 4-pt Likert scale (1 "not very much", 4 "very much") that asks about the helpfulness and supportiveness of the coaching session	Teacher ratings

coaching sessions (i.e., COMPASS Coaching Feedback Form), and a fidelity checklist that was completed by teachers about what the coach did during the coaching session (i.e., COMPASS Coaching Fidelity Checklist). Table 5.1 briefly explains the measures, which are also available in the COMPASS manual (Ruble et al. 2012). Using these measures, quality was indicated by means above a minimum value as rated by teacher perceptions of quality. Specifically, item and scale mean scores of 3 or better were viewed as indicative of quality, corresponding to teacher ratings of "agreement" that high quality coaching occurred.

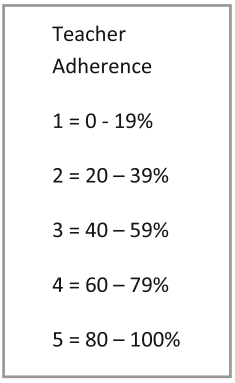
Can we demonstrate good coaching quality for both face-to-face and web-based approaches? Teachers completed ratings of coaching quality at the end of the sessions. However, to ensure that we did not overburden teachers, at each coaching time point, we randomly selected teachers to complete the COMPASS Coaching Fidelity Checklist. Although not every teacher provided feedback after each session, we were able to obtain feedback from every teacher for at least one session. The COMPASS Coaching Feedback Form, our measure of overall satisfaction, was collected once at the end of the year. As mentioned, ideally it would have been preferable to conduct independent observations of the coaching sessions and obtain ratings that could be compared with teacher ratings. Moreover, teachers may have provided higher ratings because the data collector was the consultant.

The COMPASS Coaching Fidelity Checklist ratings from our second RCT showed that we were able to demonstrate good quality coaching. The overall mean score based on teacher reports from both groups (i.e., face-to-face and web-based coaching) was 3.7 out of 4.0. When we compared our ratings by group assignment (face-to-face teacher coaching vs. web-based teacher coaching) using a t-test, we found no differences between the two groups ($t = -0.0$ (20), $p = 0.63$). Thus, the modality of coaching did not impact teachers' perceptions of quality with our fidelity form.

Similar results were obtained for satisfaction, using the COMPASS Coaching Feedback Form. The overall mean satisfaction rating was 3.6 out of a possible 4 points. When we analyzed ratings based on group assigned (face-to-face vs. WEB), no differences were found ($t\ (26) = -0.17$, $p = 0.87$). That is, similar to our previous findings, the modality of coaching had no impact on teacher perception of helpfulness of the coaching sessions. Thus, we were able to demonstrate that we were able to implement high quality coaching, as indicated both by good adherence to

the COMPASS model and by teacher satisfaction, and that web-based coaching was perceived similarly to face-to-face coaching.

Does coaching quality relate to teaching quality? We faced challenges similar to those discussed above with respect to coaching quality, when assessing teaching quality or fidelity. We will discuss those issues in more detail in the following chapter. Initially, we developed a simple tool because we wanted a measure that could be useful for actual use by consultants who are in the schools and that did not require extensive time or personnel to complete. We also wanted a tool that would be sensitive to detecting differences between teachers and changes in teaching quality over time. If so, then this would lend to the usability of the measure outside the research context. Essentially, we scored teachers using the 5-point scale below in terms of how much of the overall teaching plans were implemented.



Recall that one of our key questions was whether implementation quality (consultant behavior) was related to intervention quality (teacher behavior). To examine this question, we correlated teacher adherence scores with coaching quality scores (fidelity and satisfaction) for each of the coaching sessions. The results showed that the quality of coaching, measured with the COMPASS Coaching Fidelity Checklist, did indeed impact teaching quality. Significant correlations were obtained between coaching and teaching quality for three of the four time points. Specifically, significant associations between coaching and teaching quality were found for coaching session 1 ($r = 0.54$, $p < 0.01$), coaching session two ($r = 0.41$, $p < 0.05$), and for coaching session four ($r = 0.47$, $p < 0.05$; Ruble et al. 2013).

We then correlated teacher adherence with scores for our second implementation quality measure, the COMPASS Coaching Feedback Form, or teacher satisfaction with the coaching sessions. Unlike our prior findings, we observed *no* relationship between teacher adherence and teacher feedback/satisfaction with COMPASS coaching ($r = 0.09$; Ruble and McGrew 2013). One problem was that there was a limited range of satisfaction scores, which can hamper the ability to detect associations between variables due to restriction of range. The restricted range may have been due to the fact that only two consultants provided all of the consultation and both were original developers of COMPASS. That is, both were experts in

COMPASS. Perhaps an association would be detected when the intervention is delivered by a larger number of consultants who were not the developers of COMPASS.

Taken together, our analyses suggest that the COMPASS Coaching Fidelity Checklist was useful for understanding teacher adherence whereas satisfaction was not. Although possible restriction of range remains an issue, one plausible interpretation of these results is that what is critical to intervention or teacher adherence is consultant adherence to specific implementation elements, or coaching behaviors, rather than general satisfaction with the consultant in areas that are not clearly tied to specific implementation elements. That is, adherence or fidelity, not ratings of satisfaction, are more critical indicators of quality. These findings and this interpretation are consistent with the larger literature on the critical importance of fidelity as a measure of intervention quality and as a predictor of intervention outcomes for psychosocial and educational treatments generally (Breitenstein et al. 2010; Carroll et al. 2007; Dusenbury et al. 2003). These results led us to our next and most important question.

Does coaching quality also relate to student outcomes directly and/or is any potential relationship mediated by the association with teaching quality? To answer this question, we correlated scores from the COMPASS Coaching Fidelity Checklist and from the COMPASS Coaching Feedback Form (satisfaction) with our primary student outcome measure, GAS scores collected at each coaching session and at the end of the year by an observer who was unaware of group assignment. We found that neither coaching fidelity nor coaching feedback/satisfaction had a *direct* impact on GAS scores (Ruble et al. 2013). That is, implementation quality was not directly related to student outcomes. However, in the following chapter on the active ingredients of COMPASS we will learn that teacher adherence is associated with student outcomes. Moreover, recall that teacher adherence was also associated with coaching quality for three of the four coaching time points. That is, coaching fidelity is related to teaching quality which in turn is related to student outcomes. Therefore, in answer to the second part of our question, conceptually, one possible interpretation of these results is that implementation quality is indirectly related to student outcomes. Specifically, teaching quality may act as a mediator between coaching quality and child goal attainment outcomes because teaching quality is correlated with both coaching quality and child outcomes. However, a confident conclusion requires further confirmation using formal statistical testing of mediation.

Up to this point, we reviewed what we have learned about COMPASS as an implementation practice. For the initial consultation, we learned that fidelity and satisfaction based on teacher report are correlated. We also learned that the quality of the communication exchanges was overall very positive and correlated with IEP quality (to be discussed more in the next chapter). For quality of our coaching sessions, we were able to show that COMPASS provided high quality coaching and that teachers were satisfied. We also learned that coaching quality correlated with our intervention practice of teacher adherence, but not with our practice outcome of child goal attainment change. Table 5.2 provides a summary of the constructs we

Table 5.2 General and specific implementation quality measures

Assessment time point	General	COMPASS specific
Initial COMPASS consultation quality	Speech exchanges	1. Fidelity
		2. Satisfaction
COMPASS coaching quality	None	1. Fidelity
		2. Satisfaction

have measured thus far, and which ones are reflective of consultation quality in general and which ones represent COMPASS implementation quality in particular.

In summary, we have focused on the implementation science framework described by Dunst and colleagues (Fig. 1.3, Chap. 1) as included in our Integrated Model and have presented data on what we have learned about consultation and coaching quality. We discussed in this section some of the elements that have been tested that help explain outcomes and other elements that showed an indirect relationship to outcomes. This includes elements specific to the initial COMPASS consultation as well as to COMPASS coaching. But it also includes general elements of good consultation (e.g., alliance, empathy). Thus, our model includes both consultant and teacher behavior that is specific to COMPASS and also general to good consultation and good teaching. In the next chapter, our focus turns to teacher behavior and what we have learned about teaching quality and its impact on student behavior and outcomes.

References

Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in nursing & health*, *33*(2), 164–173.

Brown, D., Pryzwansky, W. B., & Schulte, A. C. (2011). *Psychological consultation and collaboration : Introduction to theory and practice* (7th ed.). Upper Saddle River: University of North Carolina.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement Sci*, *2*(1), 40.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American journal of community psychology*, *41*(3–4), 327–350.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health education research*, *18*(2), 237–256.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional children*, *71*(2), 149–164.

Leaper, C. (1991). Influence and involvement in children's discourse: Age, gender, and partner effects. *Child development*, *62*(4), 797–811.

Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical psychology review*, *11*(3), 247–266.

Odom, S. L. (2009). The tie that binds evidence-based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education*, *29*(1), 53–61.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, *78*(1), 33–84.

Ruble, L., Birdwhistell, J., Toland, M., & McGrew, J. (2011). Analysis of parent, teacher, and consultant speech exchanges and educational outcomes of students with autism during COMPASS consultation. *Journal of Educational and Psychological Consultation, 21*(4), 259–283.

Ruble, L. A., & Dalrymple, N. J. (1996). An alternative view of outcome in autism. *Focus on Autism and Other Developmental Disabilities, 11*(1), 3–14.

Ruble, L. A., & Dalrymple, N. J. (2002). COMPASS: a parent-teacher collaborative model for students with autism. *Focus on Autism and Other Developmental Disabilities, 17*(2), 76–83.

Ruble, L. A., Dalrymple, N. J., & McGrew, J. H. (2010). 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, 32*(4), 286–301.

Ruble, L. A., Dalrymple, N. J., & McGrew, J. H. (2012). *Collaborative model for promoting competence and success for students with ASD*. New York: Springer Science + Business Media.

Ruble, L., & McGrew, J. H. (2013). Teacher and child predictors of achieving iep goals of children with autism. *Journal of Autism and Developmental Disorders*. doi: 10.1007/s10803-013-1884-x.

Ruble, L. A., McGrew, J. H., & Toland, M. D. (2013). *Mechanisms of change in compass consultation for students with autism* (Vol. 35). USA: Sage Publications.

Sheridan, S. M., Meegan, S. P., & Eagle, J. W. (2002). Assessing the social context in initial conjoint behavioral consultation interviews: An exploratory analysis investigating processes and outcomes. *School Psychology Quarterly, 17*, 299–324.

Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: strength, integrity, and effectiveness. Journal of consulting and clinical psychology, 49(2), 156.

Chapter 6
COMPASS Intervention Quality and Active Ingredients

As before, we used the Integrated Model (Fig. 6.1) shown below to guide our research questions. In this chapter, we focus on the second block, the quality of the intervention practice or teacher behavior more specifically. In prior chapters we demonstrated the relationship between implementation elements (i.e., consultant behavior) and intervention elements (i.e., teacher behaviors). In this chapter we were most interested in exploring the relationship among teacher behaviors and their relationships to outcomes. One of the key questions we had concerned our measurement of intervention quality or how well the teacher implemented and followed through with COMPASS consultation and coaching recommendations and its relationship to outcomes. Specifically, good versus poor outcomes may be explained by differences in terms of intervention quality. However good outcomes together with low intervention quality may indicate the influence of non-intervention factors such as naturally occurring changes in the student (e.g., maturation), variables related to a revised implementation of the intervention, or other concurrent medical, educational or psychosocial interventions (Harn et al. 2013). Moreover, the measurement of fidelity is even more crucial in the largely uncontrolled community setting, when service providers, practitioners, and teachers are asked to implement evidence-based practices. For COMPASS to be a feasible consultation model implemented by school-based consultants, we needed to develop methods to insure the quality of implementation through the use of simple, real time, assessment of fidelity of both the implementation (what the consultant does) and the intervention (what the teacher does) (Fig. 6.1).

Recall in the previous chapter our description of two types of quality measures—structural and procedural. We examined both structural and process fidelity in both RCTs. Table 6.1 describes our measures of intervention quality, the criterion or purpose of the measure, the specific construct represented by the measure, and the description of the measure. With respect to structural fidelity, we first applied a simple measure of structural fidelity and tested whether it was stable, able to be measured reliably, and sensitive to detecting change associated with child educational outcomes. We also applied a second structural fidelity measure of IEP quality. Specifically, we examined how well teachers followed through and updated their student's IEP following COMPASS. We developed an IEP quality measure and assessed this change pre and post-COMPASS and compared to the control groups.

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*,
SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2_6

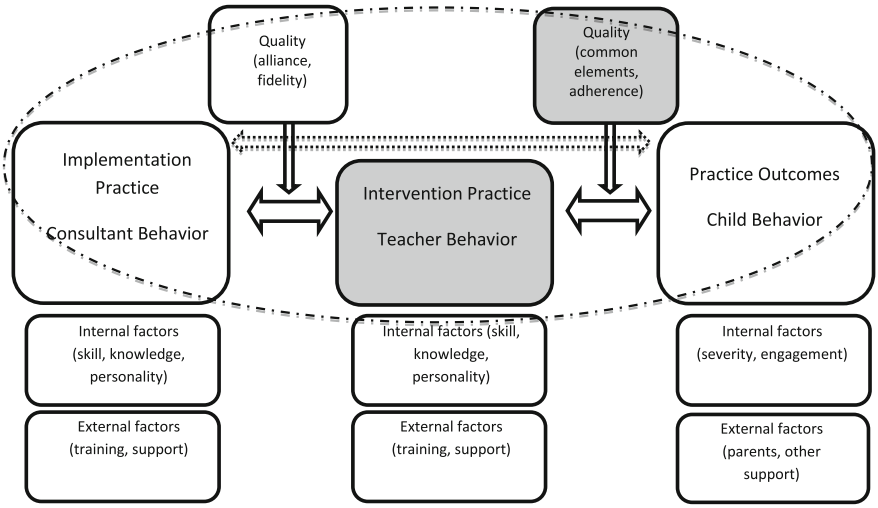


Fig. 6.1 Integrated model with focus on intervention practice quality

Table 6.1 Intervention fidelity measures

Type of fidelity	Criterion	Construct	Description
Structural	Adherence	Teacher adherence	The percentage of elements implemented from the intervention plans
Structural	Program differentiation	Targeted IEP quality	Areas of IEP improvement as a function of COMPASS
Procedural	Participant responsiveness	Teacher engagement	Quality of teacher engagement during instruction with child
Procedural	Participant responsiveness	Student engagement	Quality of student engagement during teacher instruction
Procedural	Quality of delivery	Common elements of teaching sequence	Quality of implementation of the intervention plan

These two measures will be described in detail as we discuss mechanisms of change or explanations for the success of COMPASS.

For procedural fidelity, we applied three measures that assessed the quality of the intervention as delivered by the teacher. Two tapped into the quality of the teacher's instruction while with the student, and a third tapped into the quality of engagement of the student during interaction with the teacher. This last measure was developed post hoc, after our RCTs. We explain the rationale for this measure using a common elements approach later in the chapter. We will begin with a discussion of our

measures of structural fidelity. As we will show, based on our findings, both measures index critical aspects of the COMPASS model, that is, they are mechanisms of action.

Mechanisms of Action

Mechanisms of action typically are thought to reflect the underlying theory of change or active ingredients that explain why a treatment works. When we consider the model of psychotherapy, for example, in psychoanalytic theory, a key mechanism of change is catharsis or the process of releasing emotions. Thus one would expect that a person experiencing successful change would also undergo catharsis. Similarly, according to Beck's cognitive theory of depression (Beck 1995), the cognitive triad (negative view of self, other, and future) is thought to underlie depression. Thus, therapeutic change should result from and produce change in proportion to the degree to which an individual identifies with the three elements of the cognitive triad.

When asked to generate reasons why COMPASS works, we had to think carefully about the goals of COMPASS and what outcomes we expected to observe following the different activities. With respect to an intervention such as COMPASS, the articulation of mechanisms is complicated by the fact that COMPASS is an implementation strategy that is intended to alter an intervention strategy through teacher behavior change, that then should impact a child outcome. Thus, we had to think about mechanisms at two different levels—one associated with the consultant and the other associated with the teacher.

In our paper *Mechanisms of Change in COMPASS Consultation for Students with Autism* (Ruble et al. 2013), we examined the underlying factors that help explain why COMPASS works. We wanted to know what implementation variables, intervention variables (teacher variables), and child variables impacted outcomes. We were guided by both the National Research Council (2001) recommendations for effective programs, the Individuals with Disability Education Act (IDEA 2004) mandates for research supported educational intervention, and also by the frameworks described in Chap. 1. In the next section, we describe those critical elements that have been identified in our two RCTs and the hypothesized elements that need to be evaluated more thoroughly.

Recall that the two steps of COMPASS included (a) the initial, parent-teacher goal setting and treatment planning session and (b) the follow-up teacher coaching and performance-based assessment activities. At the level of implementation strategy, we had several potential elements that we thought would be important for creating change in the teacher and the child, such as how well the consultant implemented the consultation and the coaching sessions (reviewed in Chap. 5). At the level of intervention practice, what the teacher did as a result of the consultation had to be considered. At this level, we hypothesized that at least two teacher variables would be critical for positive child goal attainment outcomes—IEP

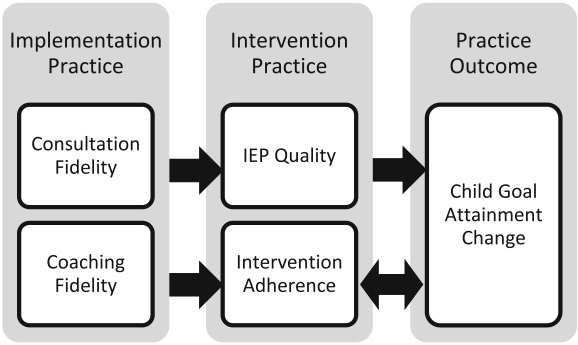


Fig. 6.2 Dunst et al. (2013) framework for understanding active ingredients

Quality, an expected outcome of the initial consultation, and implementation fidelity of teaching plans, an expected outcome of the coaching sessions. Figure 6.2 using the Dunst et al. framework (2013) specifies the two active ingredients as part of the intervention practice that we tested. This framework is familiar because we presented it in Chap. 1 and it is embedded in our Integrated Model.

Intervention practice: IEP quality. COMPASS explicitly incorporates the concept of social validity. Social validity refers to the "accurate and representative sample of the consumers' opinions" which results in information that is "used to sustain satisfactory practices or effect changes in the program to enhance its viability" (Schwartz and Baer 1991). That first contact with the parent and teacher is critical for creating a process for a shared understanding of the child from both the parent and teacher viewpoints. We believe clear understanding of the entire set of challenges and supports facing the child are critical for sound goal setting and strategy selection and that teachers with this level of understanding will do a better job and will have more confidence in their choices; this is one reason why teachers follow through with the plans that result from COMPASS.

We also believe that the discussions of the COMPASS Profile (e.g., see Fig. 5.2), which summarize and assess the challenges and resources impacting the child and family, results in better and appropriate IEP goals for that specific child. After all, we know that children with ASD need targeted instruction in social communication skills and self-direction, although what specific goals for the areas are not readily apparent. Two children may both share a need for social skills instruction, however, one may need instruction on having peer-appropriate conversation while another child may need instruction on how to share toys back and forth. For each child, the specific goals and teaching methods are unique. The information needed to identify the unique challenges for each child comes from the discussion of the COMPASS Profile assessment forms which helps create a picture of the whole child, at home, at school, and in the community. This complete representation helps specify the individualized goal, as well as personal and environmental challenges and supports necessary for understanding and developing the intervention plan.

Table 6.2 NRC and IDEA quality indicators

NRC indicators
1. Includes goals/objectives for social skills to improve involvement in school and family activities
2. Includes goals/objectives for expressive, receptive, and non-verbal communication skills
3. Includes goals/objectives for organizational skills and other behaviors that underlie success in a general education classroom (independently completing a task, following instructions, asking for help, etc.)
IDEA indicators
4. This objective is able to be measured in behavioral terms
5. The conditions under which the behavior is to occur is provided i.e. when, where, with whom
6. The criterion for goal acquisition is described i.e. rate, frequency, percentage, latency, duration as well as a timeline for goal attainment is described specifically for objective

Thus, one expectation for completing step one of COMPASS, the initial consultation, was ecologically valid, personalized goals, and that these goals would be reflected in a better IEP. Specifically, we expected that the IEPs would have teaching goals that were well written and personalized to the child and reflected the needs of students with ASD based on both NRC (2001) best practices for educating students with autism and IDEA (2004) federal mandates for special education programs. Both guidelines are incorporated into COMPASS. The elements that came from these two sources resulted in an IEP evaluation tool (Ruble et al. 2010) that we used to test our prediction of better quality IEPs. The IDEA elements that were incorporated into our tool were based on the quality of the IEP goals. Well written goals have at least three features: (a) they are measurable; (b) they are observable; and (c) they have a criterion or expected attainment level described. The NRC elements of quality that were incorporated focused on the nature of the goals and their sensitivity to the needs of students with ASD. An NRC informed IEP contains goals in key areas identified as critical for ASD including: social goals, communication goals, and goals that reflect skills necessary for independent or self-directed learning skills. Table 6.2 shows examples of our evaluation tool that was used for quality determination.

Following the initial consultation, we ask teachers to update the IEPs with the new goals developed from the consultation. To test whether the features described above actually were incorporated and changed as a function of COMPASS, we analyzed IEPs from teachers who received COMPASS in both RCTs. We had access to the original IEPs and to the revised IEPs, which reflected recommendations from the COMPASS consultation. A rater unaware of group assignment evaluated the quality of all IEPs before COMPASS and again for those who received COMPASS using the updated IEP. To make this comparison, we used the evaluation tool to score the quality of the IEPs based on NRC and IDEA standards.

Three Features of Well-Written IEP Goals:

1. Measurability
2. Observational
3. Clear Criterion Level

Our basic aim was to investigate whether the initial COMPASS consultation impacted IEP quality. To be able to confidently attribute obtained changes to COMPASS, we asked three interrelated research questions. First, we asked if IEP quality changed, i.e., improved, for students whose teachers received COMPASS? Second, we asked whether IEP quality was higher for the experimental group compared to the control group after receiving COMPASS? Third, we asked whether changes in IEP quality were restricted to IEP elements targeted by COMPASS or were broad and related to overall IEP quality.

For those receiving COMPASS, did IEP quality improve after the consultation? To answer the first question about whether IEP quality changed for students whose teachers received COMPASS, we scored IEP quality for teachers in the experimental condition both before and after COMPASS and compared scores. We repeated this analysis in both RCTs. Overall, we found that IEP quality improved significantly after COMPASS when compared to baseline IEP quality scores for study one ($t(13) = -2.7$, $p = 0.02$) and study two ($t(27) = -8.6$, $p = 0.000$). That is, as expected, we were able to confirm the IEP quality increased in both RCTs. However, improvement in IEP quality was stronger in the second RCT. One possible reason for this is that we realized during the first study that we needed to spend a lot of time on helping teachers to create high quality goals. Teachers had a difficult time generating high quality measurable goals for social, communication, and independent work skills. Thus, we created a template to make the process easier and more efficient in the second RCT. The Fig. 6.3 shows what we used during the consultation to ensure a high quality goal. Thus, we believe that the use of the template helped teachers to create better goals, and that this was one explanation for the better IEP quality scores in the second RCT. More explanation of the use of the template is provided in our manual.

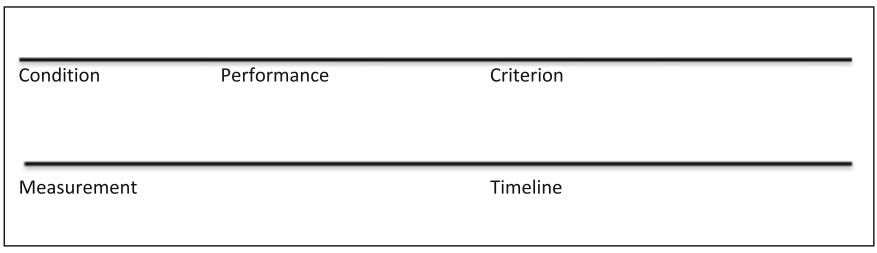


Fig. 6.3 Template for creating high quality IEP goal

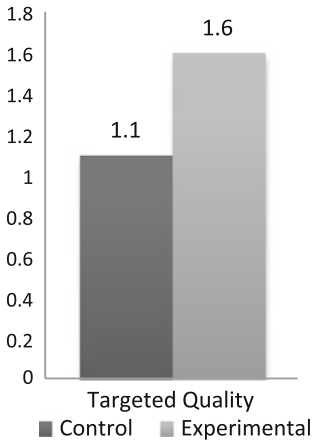
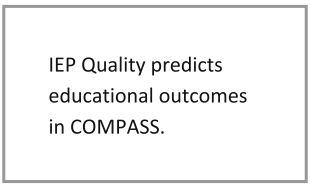


Fig. 6.4 Between Group Differences of IEP Quality

Was IEP quality good or at least better for the COMPASS group compared to the teachers who did not receive COMPASS? For this analysis, we compared IEP quality of the control and experimental conditions. As before we repeated the analysis in both RCTs. Compared to the control condition, targeted IEP quality was greater in the group receiving COMPASS ($t(47) = -5.7$, $p = 0.000$) (see Fig. 6.4). This suggests that COMPASS does result in changes in specific quality elements that can be measured. Importantly, IEP quality prior to the COMPASS intervention was similar for the control and experimental groups, suggesting that the obtained improvements after COMPASS were related to the intervention.

Were improvements in IEP quality made across the board (which might indicate the influence of some factor that was not specific to COMPASS) or were improvements specific and related to those elements expected to change as a result of COMPASS? As mentioned earlier our measure of IEP quality assesses elements of IEP quality identified by both NRC recommendations and IDEA standards. COMPASS, however, did not target all aspects of IEP quality, but focused specifically on identifying and crafting measurable goals for the three critical areas identified by the NRC task force. Thus, we did not expect improvements in IEP quality generally but only in those areas specifically targeted by COMPASS. Accordingly, we divided our IEP quality assessment into targeted and non-targeted IEP quality elements. As before we repeated the analyses for both RCTs. As expected, improvements in IEP quality were found only for IEP areas targeted by COMPASS ($t(62) = 7.2$, $p = .000$, two-tailed) and were absent in areas that were not targeted ($t(58) = -0.44$, $p = 0.66$) in our combined sample. In summary, as expected, COMPASS resulted in improved IEP goals and quality compared to before COMPASS and to comparison conditions, and was limited to areas of IEP quality specifically targeted for improvement. It is important to emphasize that these results were obtained in two separate and independent studies, considerably strengthening our confidence in the findings.

Taken together these results indicate that COMPASS is able to improve goal quality, as measured by IEP quality, that the improvement reflects a level of quality greater than found for teachers not receiving COMPASS, and that the improvements are not general, but specific to areas targeted by COMPASS. These are important results validating our hypothesized theory of change within COMPASS. A further aspect of our hypothesized theory of change is that good, high quality goals are important, in and of themselves, as drivers of teacher behavior, and that the predicted changes in teacher behavior should align with the goals, which should, in turn, positively impact student outcomes.



Thus, our next question was whether IEP quality correlated with child goal attainment change scores. As already noted, we reasoned that if we improved the quality of the IEP, this should result in improved child educational outcomes. We based our hypothesis on several factors. First, improved IEP quality after COMPASS was thought to reflect the careful selection of goals and development of intervention plans with parent and teacher input and shared-decision making and understanding of best practice IEPs for students with ASD. That is, the goals reflect ecologically and empirically valid goals that should align with the shared understanding and interest of all parties and, further should produce increased commitment to and confidence in the goals. Second, goals that reflected the personal needs of the student and were written to be measurable and observable, with a clear criterion level would make progress monitoring easier to conduct. Thus, skill attainment could be more readily observed. Moreover, clear goals are more motivating, because individuals are better able to visualize and understand the specified targets. These might seem like obvious features that would represent all IEP goals, but we found in our paper "*Examining the Quality of IEPs for Young Children with Autism*" (Ruble et al. 2010) that **only 40 % of goals were described in behavioral terms with clear conditions under which the behavior was to occur**. In other words, most IEP goals are not clear enough to measure, or to motivate behavior. If they are not able to be measured, then there is no way of knowing how much progress the child has made on their goals. Third, we believe that the activity of teachers who followed through with our suggestion to update the IEP with the new goals reflected adherence and agreement with the process. That is, we believe that teachers with better quality IEPs were not only more likely to embrace the goals, but also more likely to embrace and implement the strategies to achieve them.

Based on the above rationale, there was good reason to expect IEP quality to be related to student outcomes. As a first step, we tested the relationship between IEP quality and child goal attainment change. As we hoped, IEP quality was

significantly and robustly correlated with GAS change (r = 0.53, p = 0.025) for study 1. In other words, as predicted, students of teachers whose IEPs were rated of higher quality also achieved greater progress on IEP goals. Moreover, we also found that the nontargeted IEP quality elements were not correlated with goal attainment change scores (r = 0.001, p = 1.0). That is, the helpful impact of IEP quality was only for those aspects specifically targeted by COMPASS. As mentioned earlier, this helps to rule out other non-COMPASS influences on IEP (which likely would produce a general impact on quality) as instrumental in affecting outcomes. Together, these are potentially important findings, since IEPs are a central feature of special education and the elements targeted by COMPASS were also the same features sensitive to positive child educational outcomes.

But there was one critical concern. Our result was from a single small sample study. Could we replicate this relationship in our second study? Thus, it was critical to attempt to replicate the finding in a new study with a new sample. When we did this in our second study, the result was similar (r = 0.64, p = 0.000). Because we were able to replicate this result independently in study 2, this added considerable confidence to our initial findings. We also replicated the finding that nontargeted IEP quality elements were not correlated with child educational outcomes (r = 0.23, p = 0.17). Moreover, using combined data from both studies, the overall Pearson correlation was similarly strong ($r = 0.58$, $p = 0.000$, two-tailed). Figure 6.5 visually displays this effect by graphing the GAS change scores for students with IEP quality above and below the sample median. Thus, we have robust evidence that IEP quality is one of the explanations for why COMPASS works.

Goal-Directed Behaviors. One possibility for the success of COMPASS is that the activities needed to develop high-quality IEP goals that are measurable and objective, also may be essential for establishing *clarity* in teacher goal-directed behaviors. Special educators have challenging work. They often have insufficient time to plan and help meet student needs. Demands for accountability and paperwork may interfere with classroom teaching. And feelings of loss of control in designing and implementing curricular practices and innovations may lead to a loss of focus, stress, and burnout (Wisniewski and Gargiulo 1997). The COMPASS process may help counter these challenges by helping teacher's feel a sense of efficacy and influence in educational decision-making. The goals that result from

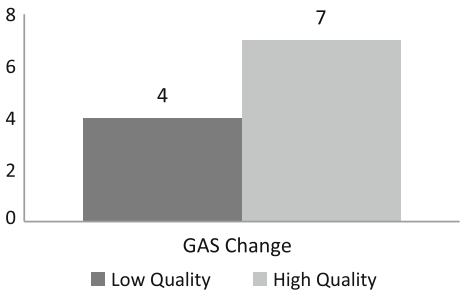


Fig. 6.5 GAS change by IEP quality

decision-making are integral for focused behavior; that is, as posited by goal setting theory (Ryan 1970), goals affect action. In fact, the more that goals are clearly defined, well-specified, and time-limited, the better the stage is set for task performance. Locke and Latham (2002) refer to four ways in which goals affect action: (a) by drawing and maintaining effort toward activities associated with the goals, (b) by increasing effort, (c) by influencing persistence, and (d) by affecting indirect behaviors of excitement, discovery, and use of task-relevant knowledge and strategies.

The goal setting activities within the initial COMPASS consultation embed these important actions of goal development and goal measurement. Moreover, the coaching sessions provide additional features critical for goal attainment. That is, for goals to be obtained, feedback toward progress is essential (Locke and Latham 2002). Coaching sessions include performance feedback within the set activities of progress monitoring. However, future research is needed to carefully assess and test the degree to which COMPASS actually includes and promotes these aspects of goal setting and their impact on outcomes. For example, to test this in a future RCT, we would need to assess and analyze the relationships among the following variables: teacher goal attainment self-efficacy, teacher ratings of effort toward each goal, time spent on each goal, perception of receiving feedback on each goal, and helpfulness of that feedback.

Although we now have very clear evidence that IEP quality is related to GAS change, and as outlined above, have identified several possible reasons why goal setting may be critical to success, several questions remain. One question concerns whether the results pertain to each subdomain of GAS change. Our current results are based on an overall GAS score that reflects mean GAS change score across three different IEP objectives: social skill, communication skill, and independent work behavior skill. Does IEP quality predict GAS improvement equally for all three subdomains? Another question is whether there are particular elements/items of the IEP quality scale that are particularly predictive? For example, does the measurability of a goal lend more importance than the type of goal, criterion level set for goal attainment, or the observability of the goal? That is, what are the key elements of IEP quality that seem to be related to GAS improvement? A further question concerns the possible influence of child and teacher factors. Are the results obtained equally for different subgroups of students with ASD (e.g., those with differing levels of autism severity) or for different subgroups of teachers (e.g., those with differing levels of experience or job burnout)? Clearly, many questions remain for future research on understanding the importance and key elements of IEP quality as it impacts GAS change. In the next section, we discuss the other intervention practice variable we believe is critical to the effectiveness of COMPASS.

Teacher Adherence. Recall in our Integrated Model (Fig. 6.1) that the intervention is carried out by the teacher and that intervention quality refers to teaching quality, i.e., the quality of teacher behaviors. A key indicator of quality is intervention adherence. Fidelity or adherence has been shown to be a robust predictor of treatment outcomes across a variety of disciplines, interventions, and diagnostic groups. Recall that in COMPASS there are multiple levels of adherence or fidelity.

That is, COMPASS is the implementation strategy, with the consultant as the change agent; and the teaching plans are the intervention strategies, with the teacher as the change agent. Thus, we hypothesized that adherence or fidelity to both the implementation and intervention strategies would be necessary if COMPASS was to be effective. In particular, both were expected to be related to student change, i.e., GAS change. However, implementation quality should only impact student outcomes indirectly through the intermediary of teacher behavior or quality, since only teachers interact directly with the students. We often assume that fidelity leads to better outcomes, but rarely is this tested directly in consultation research. As will be shown below, in addition to IEP quality, we have evidence that teacher adherence, or how much of the teaching plans were implemented by the teacher, partially explains the effectiveness of COMPASS.

However, first we had to create a measure of teacher adherence to the teaching plans. We chose to measure teacher adherence based on behavior that was either observed during coaching or videotaped for coaching. Recall that the coaching activity comprises a series of actions that include reviewing or observing the teacher implementing the teaching plan, asking the teacher to reflect on the observation, comparing the observed implementation of the teaching plan to the written plan, discussing any discrepancies between the written and observed instruction, problem-solving any issues, and planning for future activities. Thus, teacher adherence was measured by determining the percentage of components from the written intervention plans that were implemented during the coaching observation (see box).

Teacher
Adherence

1 = 0 - 19%

2 = 20 – 39%

3 = 40 – 59%

4 = 60 – 79%

5 = 80 – 100%

However, we assessed teacher adherence slightly differently in the two RCTs. In study 1, the observation was conducted live and in the classroom. In study 2, the observation was made from a video that the teacher made. Thus, as a preliminary step we first verified that the two methods of observation were comparable in providing valid data to assess adherence. In our paper *Goal Attainment Scaling as an Outcome Measure in Randomized Controlled Trials of Psychosocial Interventions in Autism* (Ruble et al. 2012), we evaluated whether teachers who made tapes of their implementation of teaching plans were providing tapes that were artifactually of higher quality and biased toward score inflation. If so, then the tapes would not be representative of the teacher's general instruction and teacher-made tapes would not be

consistent or reliable with tapes made by the researcher when conducting a live observation in the classroom. To test whether or not teachers who provided tapes were biased in showing the best case scenario, we compared GAS scores from teacher-made tapes to GAS scores from researcher-collected tapes. We found that there was no difference between the two sources ($t(24) = -1.6$, $p = 0.11$). Given this preliminary result, we felt confident that videotape samples of instruction provided by teachers were representative of their teaching overall. Thus we treated adherence ratings based on live observation and taped observation as equivalent.

To determine how much of the intervention plans were implemented, overall adherence was measured by taking the number of strategies demonstrated by the teacher that were consistent with the written intervention plan divided by the total number of strategies described in the plan. Teachers who implemented 80 % or more of the strategies received the highest score for implementation. We then tested whether our measure of teacher adherence correlated with student outcomes. As predicted, in study 1, our simple teacher adherence measure correlated strongly with child goal attainment change ($r = 0.59$, $p = 0.01$). That is, students of teachers who reliably implemented the COMPASS teaching plans showed greater improvements in goal attainment over the course of the study. However, as with our other results, it was important to replicate the finding in an independent sample. Happily, in study 2, we found very similar results. Moreover, the results appeared to be even stronger, the correlation between overall GAS change and overall teacher adherence was very large ($r = 0.83$, $p = 0.000$), explaining over 68 % of the variance in final student outcomes.

We next asked whether teacher adherence also was associated with interim changes in GAS across coaching time points. Thus, as a follow-up, in study 2, we also examined the correlation between interim measures of teacher adherence obtained over the course of the study and GAS scores obtained at the same time. Similar to our finding for final GAS score, with one exception, adherence for each coaching session correlated with the interim GAS score obtained for that session (session 1 ($r = 0.54$, $p < 0.01$); session 2 ($r = 0.63$, $p < 0.001$); and session 4 ($r = 0.44$, $p < 0.05$). Interestingly, fidelity and GAS scores collected at coaching session 3 were not correlated. As discussed more thoroughly in the next section, one possible explanation was the negative impact on adherence of the winter break. Overall, we now had very strong evidence, both from independent studies and at multiple time points, that teacher adherence was related to student outcomes.

As before with IEP quality, these results greatly increase our confidence that adherence is a critical ingredient. That is, children whose teachers more closely followed the teaching plans as outlined using COMPASS, achieved better outcomes as measured by progress toward IEP goals (see Fig. 6.6). Moreover the size of the correlation indicated a moderate to strong relationship. That is, teaching quality has an important impact on outcomes. Further, this is one of the few consultation studies that confirm this relationship. Also, it is important to note that this is the second indicator of the critical importance of teaching quality for student outcomes,

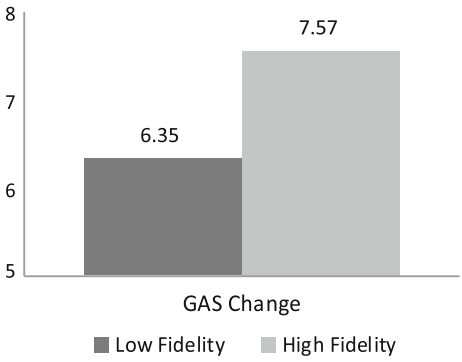


Fig. 6.6 GAS change by low and high teaching fidelity

in addition to IEP quality. Together these findings strongly confirm the potential importance of quality in explaining variations in student outcomes. However, these findings also led to further questions: Is teaching quality consistent over time?

Does teaching quality improve over time? To answer this question, we examined scores from our single adherence item that came from the COMPASS Coaching Impressions Questionnaire. Recall that this was a single question that asked what percent of the intervention plans was implemented overall by the teacher. Figure 6.7 displays adherence scores over time using the mean scores from both RCTs. The findings demonstrate that teacher fidelity significantly improved over time. Also interesting is that the lowest gain in improvement was between coaching 2 and 3, when winter break occurred. Thus the data suggest that coaching session 3 fidelity was impacted by this break. But by coaching session 4, teachers made up for the poorer performance from the previous coaching session and implemented more elements from the teaching plans. In fact visual analysis of the slope of change between coaching sessions 3 and 4 indicates a slope similar to that between coaching sessions 1 and 2, suggesting that competence and adherence in the implementation of intervention plans develop and improve over time. Moreover, these data provide supporting evidence for the need of coaching or booster sessions because fidelity is enhanced with each subsequent coaching activity throughout the school year.

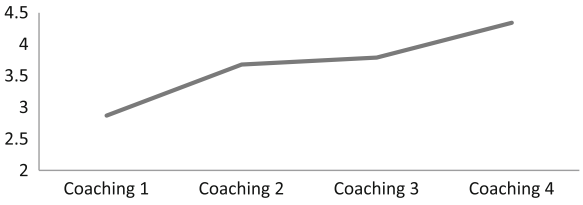


Fig. 6.7 Teacher adherence by coaching session

Other Measures of Teaching Quality: Procedural Fidelity

Recall that we also examined teaching quality in terms of procedural /process fidelity. Now that we had evidence that teaching quality was related to outcomes, we wanted to explore this further using additional measures to try to understand how quality impacts outcomes. Initially, we restricted our analyses to the two engagement measure of process fidelity. We had several questions. Our first question concerned the correlation between the two engagement measures. That is, did the measures show evidence of both overlap (convergent validity) and independence (discriminant validity). Other questions were whether engagement was consistent over time, and whether the measures were sensitive enough to detect changes in one another. A fourth question was whether they also could predict the practice outcome, child goal attainment change. Below we briefly describe the three measures and explore them in detail in the following section (Table 6.3).

The literature uses various ways to describe high quality programs. However, one consistently important aspect associated with high quality programs for students with ASD is engagement. The NRC identified engagement as an essential ingredient in educational programs. They further stipulated that a minimum of 25 h of week of active engaged time was necessary for effective teaching. Moreover, they noted that it was important to differentiate between number of hours of intervention and number of hours of engaged time. Unfortunately, the NRC's definition of engagement is vaguely defined (Ruble and Robson 2007) as "sustained attention to an activity or person" (p. 160). We developed two measures of engagement in hopes of providing clarity to the construct. Because engagement is bidirectional and is a property of the teacher-student dyad, we used two measures—teacher engagement and student engagement.

Table 6.3 Intervention process quality measures

Procedural fidelity measure	Description	Source and target
Autism engagement rating scale	A 6-item, 5-pt Likert scale that measures the child with autism's degree of cooperation, functional use of objects, productivity, independence, and attention as well as goal consistency between child and teacher	Coach ratings about child
Teacher engagement rating scale	A 6-item, 5-pt Likert scale that measures the teacher's level of affect, maintenance of interaction, directedness, responsiveness, initiation, and level of movement with the child	Coach ratings about teacher
Common elements of teaching sequences	A 6-item scale that measures how well the teacher set up a meaningful activity, maintained the child's attention, provided a clear initial request, allowed sufficient response time following prompts, and provided reinforcement	Coach ratings about teacher

In addition to engagement, we wanted to measure the extent to which the teacher exhibited good teaching practice. To do so, we developed another measure of teaching quality, adopting the Common Elements of Effective Teaching model of Grisham and Ruble (2014). This questionnaire assessed how well the teacher applied good teaching using a common elements approach. The engagement measures are available in the COMPASS manual (Ruble et al. 2012); the measure of common elements is available from the first author. All three quality ratings were based on consultant observations of the teacher implementing the intervention plans with the student. As discussed previously, in our first study, the observations were conducted live in the classroom. In our second study, our ratings were based on videotapes provided by the teachers of their instructional interactions with students.

Our student engagement measure, called the Autism Engagement Rating Scale (AES), (Ruble et al. 2012; Ruble and McGrew 2013) assesses six areas of child behavior: (a) cooperation, (b) functionality, (c) productivity, (d) independence, (e) consistency, and (f) attention. Teacher engagement was assessed using the Social Interaction Rating Scale for Autism (SIRS; Ruble et al. 2008), which was originally developed to measure the quality of the parent interaction during free play with the child. We adapted it to measure the quality of the teacher interaction. Like the AES, the SIRS measures six features of teacher behavior during an instructional interaction with the student: (a) level of affect; (b) maintenance of interaction; (c) directedness; (d) responsiveness; (e) initiation, and (f) level of movement. Teacher and Student Engagement were rated for each of the four coaching sessions. We will describe the Common Elements measure and scoring later.

First we present data on the engagement scales. Recall that our key questions asked about the correlation between our quality measures (in this case the engagement scales), their consistency over time, and the degree to which the quality measures predicted student outcomes. The intercorrelation matrix below displays the child engagement and teacher engagement collected at each coaching session and correlated against goal attainment scores for each coaching time-point and also at the end of the year. The ratings of child and teacher engagement and child goal attainment scores are listed on the left column. The first four rows represent the child engagement scores for each coaching session, the second four rows represent the teachers' engagement ratings for each coaching session and the last five rows represent the goal attainment scores collected at each of the four coaching sessions and the end of the year, about 6 weeks following the last coaching session (Table 6.4).

In answer to our first question, the child engagement and teacher engagement measures were consistently positively related, with the largest, and most consistently significant correlations between engagement measured at the same time point. The results indicate that the measures both overlap to some degree, convergent validity, and assess independent and unique aspects of engagement, discriminant validity, and that the two measures reflect related aspects of engagement that appear to reinforce one another, i.e., they co-occur or covary.

Table 6.4 Intercorrelation matrix of child and teacher engagement and GAS outcomes

	1	2	3	4	5	6	7	8	9	10	11	12
1. Child 1	–											
2. Child 2	0.29	–										
3. Child 3	0.39*	0.60**	–									
4. Child 4	0.36*	0.48**	0.52**	–								
5. Teacher 1	0.27	0.40*	0.30	0.12	–							
6. Teacher 2	0.00	0.47**	0.42*	0.34*	0.54**	–						
7. Teacher 3	0.40*	0.26	0.50**	0.28	0.25	0.43*	–					
8. Teacher 4	0.26	0.09	0.26	0.50**	0.33	0.44**	0.58**	–				
9. GAS 1	0.19	0.36*	0.25	0.31*	0.21	0.26	0.22	0.14	–			
10. GAS 2	0.21	0.11	0.16	0.07	0.46**	0.34*	0.32	0.27	−0.03	–		
11. GAS 3	0.33	0.29	0.40*	0.30	0.13	0.33*	0.39*	0.17	0.05	0.79**	–	
12. GAS 4	0.30	0.24	0.50**	0.50**	0.27	0.42**	0.34	0.41*	0.20	0.56**	0.74**	–
13. Change GAS	0.22	0.51**	0.37*	0.51**	0.15	0.35*	0.15	−0.09	0.34*	0.07	0.39*	0.25

Note Child = child engagement; teacher = teacher engagement; Base on combined experimental groups from RCTs 1 and 2
*$p < 0.05$; **$p < 0.01$

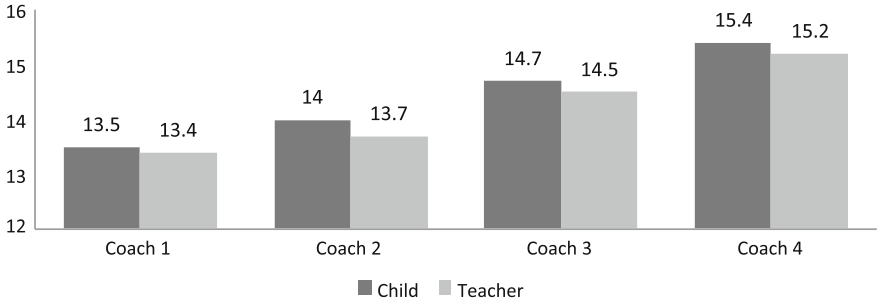


Fig. 6.8 Mean child and teacher engagement over time

A second question was how consistent is engagement over time? Analysis of the intercorrelation matrix indicated that teacher engagement at different time points correlates with teacher engagement positively and ranges from 0.25 to 0.58. The same pattern was observed for child engagement intercorrelations that ranged from 0.29 to 0.60. We plotted the mean scores of teacher and child engagement for each coaching session in Fig. 6.8. We also see an increase in both over the series of coaching sessions. Overall, the data indicate that the measures are sensitive to change, and that engagement tends both to be somewhat consistent over time and to be a changing feature of the teacher-child dyad that varies across time (coaching sessions).

Our third question concerned whether teaching quality and child engagement impacted child outcomes. We calculated the overall mean of child and teacher engagement and correlated these scores with the GAS change score. A significant correlation was obtained between GAS change and both child ($r = 0.55$, $p = 0.001$) and teacher engagement ($r = 0.43$, $p = 0.03$).

A more indepth analysis of the intercorrelation matrix of the predictive patterns of engagement provides additional information. Analysis of the correlation between child engagement and GAS suggests a somewhat inconsistent pattern. That is, child engagement at both Time 1 and Time 2 predicts GAS at other times better than when measured concurrently with engagement. Similarly, child engagement at Time 3 predicted GAS at both Time 3 and Time 4, but less strongly at Time 3. The exception to this pattern, was child engagement at Time 4, which was most strongly correlated to concurrent GAS change. For teacher engagement, a more consistent predictive pattern emerged. Teacher engagement collected at Times 2, 3, and 4, correlated with child engagement at each time point and also with the GAS score at each time point. Thus, for point predictions at each coaching session, teacher engagement is the more sensitive measures. However, when we look at overall GAS change scores, child engagement was more consistently and strongly correlated (ranging from 0.22 to 0.51) compared to teacher engagement (ranging from −0.09 to 0.35).

Although both measures provide useful information, overall, child engagement is generally predictive of overall GAS change, and the last two time points show some tendency to be specific to a time point. Teacher engagement, however, is slightly less consistent and shows greater tendency to be most predictive

concurrently. In both cases the concurrent predictions show up stronger for the later time points.

As mentioned above, there was a lack of concurrent correlations for the earlier time periods, and especially with Time 1 scores, for both teacher and child engagement. One explanation for the lack of findings for the first coaching session is that it was not unusual for teachers to have only implemented one of the three teaching plans resulting in poor fidelity overall prior to the first coaching sessions. We observed that it took teachers at least 2–3 months before they were implementing all three teaching plans. In contrast, our significant findings with outcomes for the other three coaching sessions highlight the potentially important relationships between teacher and child engagement and outcomes. Also, as mentioned above, the results show that our measures of teacher and child engagement are sensitive to changes in child goal attainment outcomes.

Common Elements of Effective Teaching

One major challenge for special education teachers is the expectation that they be prepared and able to successfully instruct children with a wide variety of needs and disabilities. Because implementation science is concerned with the factors that hinder or facilitate the use of evidence-based practices (EBP) in real world settings, including our public school classrooms (Odom 2009) a critical question is how best to prepare and support teachers' skills for designing, implementing, and evaluating EBPs, especially for those with complex learning needs such as children with ASD (Gisham-Brown et al. 2005). One problem is the sheer number of potential EBPs. That is, how realistic is it for every teacher to be expert in all EBPs for children with learning challenges and especially those with ASD? For example, there are at least 24 EBPs for students with ASD (Odom et al. 2010) and expecting a teacher to be competent in the delivery of all methods is unrealistic because classroom special educators receive cross-categorical training. An alternate is to ensure that teachers have high quality skills in the techniques that are common across a variety of EBPs.

For decades, researchers of psychotherapy outcomes have focused a great deal of attention to the study of common elements of effective psychotherapy. Psychotherapy research has revealed that when compared against viable alternatives, there is little difference in effectiveness between bona fide therapeutic approaches (Lambert and Ogles 1997). In contrast, across a variety of disorders, large-scale comparative studies of psychotherapy outcome research typically shows that individual client characteristics affects outcome more than the type of treatment (Clarkin and Levy 2004; Lambert 1992). The same findings are also reported in ASD treatment research. For example, many studies have replicated the finding that 50 % of children with ASD show a marked treatment response to high-quality early intervention (Lovaas 1987; McEachin et al. 1993; Sallows and Graupner 2005; Smith et al. 2004). Further, like psychotherapy outcome research, ASD treatment outcome research suggests that the variability in treatment response is associated

with individual rather than treatment factors or treatment philosophy. Pretreatment child factors such as level of IQ, language, and autism severity predict treatment responsiveness (Reichow et al. 2012) Thus, it might be more helpful to focus on the features common across effective treatment programs and use these indicators as a way to improve teaching quality. These key elements of instruction should be those that (a) teachers can implement with fidelity and (b) support young children to achieve important learning outcomes.

As noted earlier, we adopted a model of common elements of good teaching and developed an instructional quality measure to capture common elements of teaching sequences (Grisham-Brown and Ruble 2014). This preliminary work is guided by the premise that teachers are more likely to implement consistent instruction with fidelity if they can focus on common elements.

Purposes of the measure include designing intervention plans, and measuring implementation fidelity in inclusive early childhood settings. This measure assumes that regardless of the EBP employed, the elements outlined in Fig. 6.9 must be present for the child to have an opportunity to demonstrate the targeted skill. As listed in the figure, the teacher must set up the learning environment so that the child has the opportunity to demonstrate the skill. The materials and activities should be based on the interest of the child in a manner that engages and induces the child to respond or initiate the targeted skill. The teacher must allow enough time for the child to respond. Once the child responds, the teacher must make a decision about how to close the teaching sequence if the child (1) demonstrates the expected response, or (2) needs further support to demonstrate the expected response.

The items in Fig. 6.9 were included in a new measure called the Common Elements of Teaching Sequences (CETS) developed by Grisham-Brown and Ruble (2014). As described earlier, prior to each coaching session, teachers prepared a video recording of their implementation of the three intervention plans that focused on a social skill, a communication skill, and a learning skill. Analyses of the first and fourth coaching sessions were completed using the CETS measure. Research

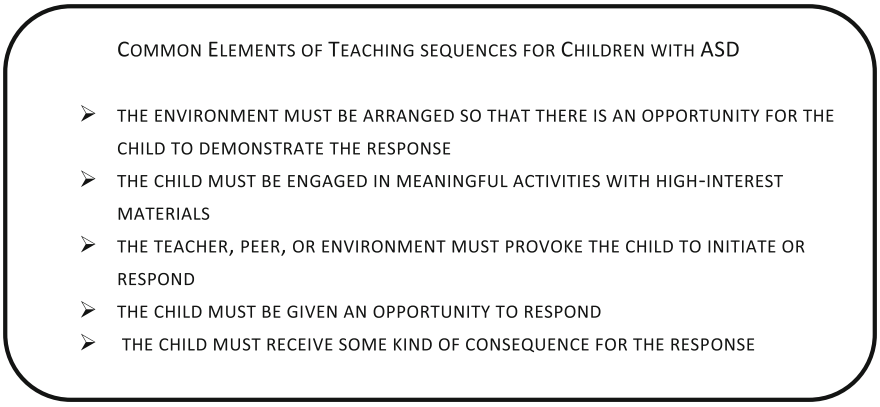


Fig. 6.9 Common elements of effective teaching

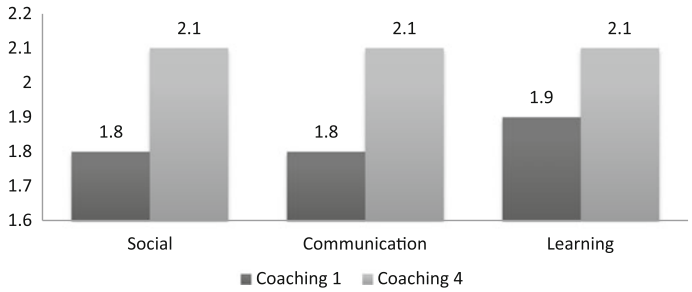


Fig. 6.10 Pre and post ratings from the common elements of teaching sequences

assistants who were not part of the study reviewed and scored tapes until they reached 80 % agreement for each CETS item. After satisfactory agreement was reached, they individually coded the videotapes.

We had two key questions similar to those we asked earlier for engagement: (1) Is teaching quality consistent over time or does it show improvement? and (2) Does teaching quality as measured using CETS predict student outcomes? To answer question one, we analyzed whether the overall mean score of teaching quality assessed with the CETS improved from coaching session 1 to coaching session 4. Figure 6.10 shows that there was a significant improvement in mean scores over time for each of the three different teaching domains (social, communication, and learning skills). Moreover, teaching quality improved for all three learning domains equally ($p < 0.00$). To answer question two, we correlated CETS scores with concurrent and overall GAS scores. Importantly, teaching quality measured at coaching session 1 correlated with child goal attainment outcomes collected at the same time point ($r = 0.56$, $p = 0.00$), and also future scores collected at coaching session 4 ($r = 0.61$, $p = 0.00$) and the final GAS change score ($r = 0.37$, $p = 0.05$). In contrast, CETS scores collected at coaching session 4 were unrelated to the concurrent GAS scores observed for coaching 4 ($r = 0.06$, $p = 0.74$) and the overall GAS change score ($r = 0.06$, $p = 0.77$). One explanation is the lack of range in scores observed at the later coaching sessions. Although there is promise for the use of the CETS, more research is necessary for understanding the lack of sensitivity for detecting changes in child outcomes for later coaching sessions.

In summary, we have described what we believe are active ingredients of COMPASS in this chapter. We also have presented several approaches for measuring teaching quality and now have indications that a variety of teacher behaviors impact outcomes. One measure was specific to COMPASS (e.g., consultant ratings of adherence), other measures pertained to the quality of the teacher-student relationship (i.e., engagement), and another focused on general elements of good teaching (i.e., common elements). Thus we have consistent overlapping evidence using various conceptual and methodological approaches of the critical importance of teaching quality. Overall, the chapter has shown consistently that teaching quality, as measured using a variety of procedural and structural methods, is a

critical and sensitive marker of teacher behavior and of student outcomes. Moreover, as we have shown in earlier chapters, teacher quality is also modifiable based on consultant behavior (COMPASS coaching and initial consultation). Thus, there are clear links between consultant behavior, teacher behavior, and outcomes.

We do not year clearly understand some of the differences in predictiveness between quality measures. However, one possible theme is that teacher quality focusing on common elements may be a better marker of progress early, before teachers have fully absorbed and implemented the specific elements of COMPASS, such that basic good teaching is explaining more of the variance, and that teacher quality reflective of adherence to COMPASS/implementing teaching plans, may be a better marker for progress later, once the teachers have had an opportunity to fully learn and practice COMPASS specific elements.

We have several questions important for future research, such as (a) Are the quality elements independent? (b) Do some build on each other? That is are the common elements or some aspects of them necessary to have good engagement? and (c) Is consultant behavior or particular aspects of consultant behavior more related to particular measures/aspects of teaching quality? In the next chapter, we return to our Integrated Model and focus more specifically on teacher internal factors that impact child outcomes.

References

Beck, J. S. (1995). *Cognitive therapy*. New York: Wiley.

Clarkin, J. F., & Levy, K. N. (2004). The influence of client variables on psychotherapy. *Bergin and Garfield's handbook of psychotherapy and behavior change, 5*, 194–226.

Dunst, C. J., Trivette, C. M., & Raab, M. (2013). An implementation science framework for conceptualizing and operationalizing fidelity in early childhood intervention studies. *Journal of Early Intervention, 35*(2), 85–101.

Grisham-Brown, J., Hemmeter, M. L., & Pretti-Frontczak, K. (2005). *Blended practices for teaching young children in inclusive settings*. Baltimore: Brookes Publishing Company, MD 21285.

Grisham-Brown, J., & Ruble, L. (Feb, 2014). *Common elements teaching sequences for assessing fidelity for implementing interventions to young children with autism and other developmental delay*. Paper presented at the Conference on Research Innovations in Early Intervention, San Diego, CA.

Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and Fit: What do we really know about fidelity of implementation in schools? *Exceptional Children, 79*(2), 181–193.

Individuals with Disabilities Education Act. (2004). http://idea.ed.gov/explore/view/p/%2Croot%2C.

Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists. *Handbook of psychotherapy integration, 1*, 94–129.

Lambert, M. J., & Ogles, B. M. (1997). The effectiveness of psychotherapy supervision. Handbook of psychotherapy supervision, pp. 421–446.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist, 57*(9), 705.

Lovaas, O. Ivar. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*(1), 3–9.

McEachin, J., Smith, T., & Lovaas, O. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97*(4), 359–372.

National Research Council. (2001). *Educating children with autism*. Washington, DC: National Academy Press.

Odom, S. L. (2009). The tie that binds: evidence-based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education, 29*(1), 53–61.

Odom, S. L., Collet-Klingenberg, L., Rogers, S. J., & Hatton, D. D. (2010). Evidence-based practices in interventions for children and youth with autism spectrum disorders. *Preventing school failure: Alternative education for children and youth, 54*(4), 275–282.

Reichow, B., Barton, E. E., Boyd, B. A., & Hume, K. (2012). Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews, 10*, CD009260.

Ruble, L. A., Dalrymple, N. J., & McGrew, J. H. (2012a). *Collaborative model for promoting competence and success for students with ASD*. New York, NY: Springer Science+Business Media.

Ruble, L., McDuffie, King, A., & Lorenz, D. (2008). Caregiver responsiveness and social interaction behaviors of young children with autism (in revision).

Ruble, L., McGrew, J., & Toland, M. (2012). Goal attainment scaling as outcome measurement for randomized controlled trials. *Journal of Autism and Developmental Disorder* (manuscript in press).

Ruble, L. A., McGrew, J. H., & Toland, M. D. (2013). *Mechanisms of change in compass consultation for students with autism* (Vol. 35). US: Sage Publications.

Ruble, L., & McGrew, J. H. (2013). Teacher and child predictors of achieving iep goals of children with autism. *Journal of Autism and Developmental Disorders*. doi: 10.1007/s10803-013-1884-x.

Ruble, L. A., McGrew, J., Dalrymple, N., & Jung, L. A. (2010). Examining the quality of IEPS for young children with autism. *Journal of Autism and Developmental Disorders, 40*(12), 1459–1470. doi:10.1007/s10803-010-1003-1.

Ruble, L. A., & Robson, D. M. (2007). Individual and environmental determinants of engagement in autism. *Journal of Autism and Developmental Disorders, 37*(8), 1457–1468.

Ryan, Thomas A. (1970). *Intentional behavior: An approach to human motivation*. Oxford, England: Ronald Press.

Sallows, G., & Graupner, T. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation, 110*(6), 417–438. doi:10.1352/0895-8017(2005)110[417:ibtfcw]2.0.co;2.

Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: is current practice state of the art? *Journal of Applied Behavior Analysis, 24*(2), 189–204.

Smith, T., Groen, A., & Wynn, J. (2004). Randomized Trial of Intensive Early Intervention for Children with Pervasive Developmental Disorder, p. 153.

Wisniewski, L., & Gargiulo, R. M. (1997). Occupational stress and burnout among special educators: A review of the literature. *The Journal of Special Education, 31*(3), 325–346.

Chapter 7
Teacher Internal and External Factors

So far, we have suggested that teacher, rather than child variables explain the success of COMPASS. Up to this point, we have reviewed critical teacher behaviors related to quality, including IEP quality, adherence, engagement, and effective teaching principles (common elements) and their impact on student outcomes. These teacher behaviors are aspects that can be observed. Besides these observable aspects of teacher behavior, we have also examined teacher internal factors, aspects of teaching we cannot directly observe but are important for child outcomes.

Given the strong impact of teaching quality on outcomes, we also wanted to understand potential influences that might impact teachers' abilities to deliver quality instruction. As illustrated in our Integrated Model (Fig. 7.1), we proposed two primary influences on teacher behavior, internal and external. Internal factors include knowledge and skill in special education, experience, knowledge and skills in COMPASS and personal characteristics, including personality factors, self-efficacy and overall sense of well-being and burnout. Teacher internal factors that we examined specifically included self-efficacy, stress, and burnout. In addition, after we conducted our first study and noticed that many teachers did not collect data as part of their regular teaching practice, we developed a new measure and also looked at teachers' beliefs about data collection and what influenced the likelihood of intention to collect data. External factors impacting teacher behavior include such factors as teaching resources (e.g., availability of instructional support equipment, well-equipped and designed classrooms), administrative policies (e.g., time for data collection and instructional planning), and social/personal supports (e.g., instructional team support, administrative support). We had few measures of external factors, but when available, we include them in our analyses and discussion. Below we discuss our findings with respect to both internal and external factors. Most of our research has focused on internal factors that we present first.

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*,
SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2_7

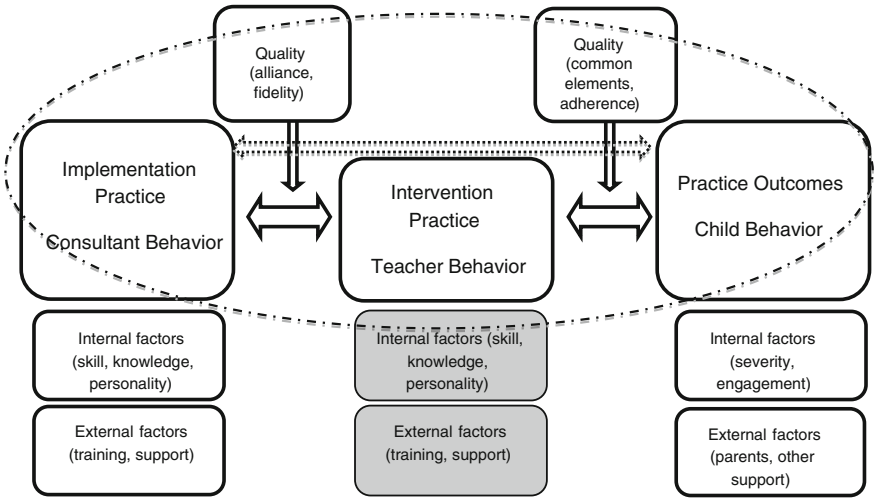


Fig. 7.1 Integrated model with focus on teacher internal and external factors

Teacher Self-efficacy

One important factor that might help us understand differences in teachers' responses to COMPASS is teacher self-efficacy. For teachers, self-efficacy refers to beliefs in one's ability to produce desired educational outcomes. Self-efficacy originates from Bandura's work (1986) that showed that when people hold perceptions that their actions can result in chosen outcomes, they are more likely to have higher motivation, effort, and persistence for achieving those outcomes in the face of obstacles and barriers (Soto and Goetz 1998). Much like goal setting theory we described earlier (Ryan 1970) that suggests the importance of clear, time-limited, and well-specified goals, self-efficacy may also play a significant mediating role in goal accomplishment. If we could document what influences self-efficacy of teachers of students with ASD, then we would be able to use this information in future research for improving educational outcomes.

Bandura identified four factors that predict self-efficacy: (a) mastery experience (perception of past performance); (b) vicarious experience (observation of others); (c) social persuasion (persuasive messages received by others); and (d) physiological and emotional states (somatic and affective reactions regarding performance). In other words, teachers who perceive themselves as performing well in the past, have the opportunity to observe competent peers, receive encouragement from others, and experience positive feelings regarding their work have high self-efficacy. Our first study on self-efficacy (Ruble et al. 2011) explored three of these four factors—mastery, social persuasion, and physiological/emotional states. We measured mastery by number of years teaching students with ASD, social persuasion by leadership

support, and emotional state by burnout. Burnout was broken down further into one of three areas: emotional exhaustion, depersonalization (treating students as if they were objects), and reduced personal accomplishments (low sense of reward from teaching). We administered a generic measure of teacher self-efficacy called the Teacher Interpersonal Self-Efficacy Scale (TISES) and examined teacher responses against the three sources. The TISES taps into self-efficacy for three general educational domains: (a) classroom management; (b) obtaining colleagues' support; and (c) obtaining principal's support.

When we correlated each of the three types of self-efficacy against the sources, we found that self-efficacy for classroom management significantly correlated with all three emotional states of burnout. In other words, teachers who reported higher ability to manage the classroom reported reduced emotional exhaustion and depersonalization, and higher sense of personal accomplishment. However, self-efficacy subscales measuring the perceived ability to obtain colleague and principal support did not correlate with any of the proposed sources of self-efficacy. Moreover, two of the sources of self-efficacy, mastery, as measured by number of years teaching students with autism, and social persuasion, as measured by administrator support, failed to correlate with any self-efficacy subscale. Perhaps when judging their level of mastery, teachers of students with autism are unable to rely on prior experience due to the heterogeneity in symptom presentation of students. The diverse set of symptoms may create difficulties generalizing experiences with one child with ASD to a different child.

Another possible explanation is that teachers may simply feel underprepared, regardless of prior levels of experience. Most current teacher training in autism is unsuccessful and is not based on methods that result in changes in classroom practice. Effective training in autism-specific interventions (Jennett et al. 2003), such as those that are hands-on and classroom-based is limited. However, the failure to find an association between self-efficacy and social persuasion also could be an artifact of low power due to low sample size ($n = 35$). Specifically, the correlation coefficient for classroom management was of moderate size ($r = 0.28$) and would have been significant with a slightly larger sample. Alternately, special education teachers may not rely on the same sources of support as general education teachers. Instead they may derive support from individuals with direct involvement in supporting their teaching (e.g., autism specialist or special education director) rather than general administrators such as principals, because persuasive messages are most influential when the persuader is intimately familiar with the task and the teacher holds confidence in their opinions (Bandura 1977).

Overall, our findings failed to find strong support for the relationship between self-efficacy measured by the TISES and putative sources of self-efficacy other than the role burnout might play in child educational outcomes. One possible explanation for these generally negative results is that our measure of self-efficacy was too broad, since it was designed to be used with general education teachers. Consistent with this concern, Bandura has emphasized that self-efficacy is not a general or broad construct, but narrowly focuses on very specific tasks and behaviors.

Thus, we decided to re-examine these hypotheses using a specific measure that related to the teaching tasks associated with learners with ASD.

Accordingly, we created a new measure to assess teacher self-efficacy directly related to the tasks required to teach young students with ASD and tested it in our second RCT (Ruble et al. 2013). The new measure, the Autism Self-Efficacy Scale for Teachers (ASSET), consists of 30 questions that represent various tasks associated with effective teaching (e.g., confidence assessing social interaction, using visual structure to increase independence) and was completed by 44 teachers of students with ASD. Items were rated using a 100-point scale (1 = cannot do at all to 100 = highly certain can do). Teachers reported an overall mean score across items of 74.5. After initial inspection of responses indicated a general failure to use the lower part of the scale, we converted the 100-point scale to a 6-point scale using the following metric: 0–50 to 0, 51–60 to 1, 61–70 to 2, 71–80 to 3, 81–90 to 4, and 91–100 to 5. The mean score on the new 6-point scale was 2.58. The high internal consistency of the total scale (0.96) indicated that the ASSET was unidimensional and represented by a single construct.

As before, we analyzed the associations between self-efficacy as measured with the ASSET and the three potential sources of self-efficacy—administrative support, mastery, and emotional responses. Similar to our prior study, self-efficacy was unrelated to administrator support and mastery. Moreover, unlike our prior study using the TISES, the ASSET failed to correlate to burnout. However, ASSET scores were related to teacher stress instead. Specifically, low self-efficacy on the ASSET was related to increased teacher stress as measured by subscales indexing higher self-doubt and need for support and higher teaching disruption. We also observed significant correlations between all subscales related to teacher stress (self-doubt, loss of satisfaction, disruption of teaching, frustration with parents) and all subscales related to burnout (exhaustion, depersonalization, lack of personal accomplishment). This is a potentially important finding. The difference between stress and burnout is burnout is the long-term consequence of chronic stress. Perhaps the ASSET may be helpful as an early indicator of stress that as it accumulates, potentially leads to teacher burnout. If identified early, the long-term consequences of stress may be able to be reduced through teacher stress/burnout interventions.

A critical question was whether self-efficacy, an internal personal characteristic of teachers, also impacted teaching quality and child outcomes (Fig. 7.1). When we correlated ASSET scores collected at the beginning of the year with child goal attainment change scores and IEP quality, no significant associations were found. But when we examined ASSET scores obtained at the end of the school year, findings were different. Teacher self-efficacy collected at the end of the school year was associated with child goal attainment change scores ($r = 0.31$, $p = 0.04$) and with IEP quality ($r = 0.44$, $p = 0.00$). That is, as predicted by Self-efficacy Theory (Bandura 1977), teacher self-efficacy both impacted teacher behavior and the outcome of that behavior, child goal improvement. Thus, these findings suggest that self-efficacy as measured with the ASSET appears to be related to outcomes through its impact on IEP quality and perhaps indirectly through its impact on stress and burnout. Moreover, the findings are consistent with our overall model positing

the potential impact of teacher internal factors on student outcomes. Finally, the fact that self-efficacy impacted outcomes, even within a well conducted clinical trial, suggests that more research is needed on understanding ways to enhance and support teacher self-efficacy.

Teacher Burnout

The most critical teacher variables are those that directly impact child educational success. So far, we have discussed three important factors—observable teacher behaviors associated with IEP quality, teaching adherence/quality, and engagement, and internal teacher reports of self-efficacy. But are there other internal teacher factors that impact child educational outcomes? We alluded to the importance of stress and burnout above because there is now a considerable body of research identifying stress and burnout as major contributors to teacher turnover (Awa et al. 2010; Carlson and Thompson 1995). For example, the annual attrition rate for special educators (13 %) is twice that of general educators and the 3 year attrition rate is approximately 25 %, i.e., one in four teachers is lost every 3 years (Boe et al. 2008; Cook and Boe 2007; McLeskey et al. 2004; McLeskey and Billingsley 2008; Nichols et al. 2008). Furthermore, stress and dissatisfaction with their teaching positions lead an additional 20 % of special educators each year to transfer to general education or seek another position in special education (Boe et al. 2008; Boe 2014). These high turnover rates are further exacerbating the national shortage in special educators, estimated at 11.2 % (U.S. Department of Education 2008), with the unfortunate result that many special education teachers are hired without adequate preparation (Boe 2014; McLeskey and Billingsley 2008). Pairing our most challenging learners, especially students with complex needs such as those with ASD, with our least prepared educators is clearly undesirable.

Moreover, instability in our teacher workforce has negative consequences for schools and most importantly for students. Although there is a large amount of attention given to teacher burnout, surprisingly, to our knowledge, no study has directly examined or linked burnout to child educational outcomes. We wanted to understand what impact, if any, burnout has on student outcomes. However, to provide context, we also wanted to understand the potential predictive strength of burnout relative to other factors from our model. Accordingly, in this analysis we focused on several measures of the internal factors (stress, burnout, years teaching, experience with ASD), a single measure of external factors (administrator support), and teacher and student engagement impacting teacher behavior and teaching quality. In Table 7.1 we show the correlations between teacher burnout and stress and teacher internal and external factors for our sample of teachers who participated in both RCTs—the control and experimental group participants. All the measures were obtained at Time 2, the end of the school year, with the exception of administrator support. It should be noted that we were unable to include self-efficacy in our analysis because we used different measures in the two RCTs.

Table 7.1 Correlations between burnout and teacher variables

	GAS change	Emotional exhaustion	Depersonalization	Personal accomplishments
Emotional exhaustion	−0.03			
Depersonalization	−0.07	0.44**		
Personal accomplishments	0.36**	−0.37**	−0.32**	
Administrator support	−0.18	−0.25*	−0.17	0.36**
Years teaching ASD	0.05	0.05	0.20	−0.16
Number taught	0.18	−0.12	0.03	0.03
Stress	−0.29*	0.36**	0.36**	−0.43**
Teacher engagement	0.05	0.18	0.15	0.11
Student engagement	0.32**	0.04	0.12	0.04

*$p <$ 0.05; **$p < 0.01$

The results suggest that teacher burnout has a direct impact on child outcomes. Students of teachers who reported a higher sense of personal accomplishments showed greater improvement on IEP outcomes. Further, the external factor of administrator support was also related to less burnout. That is, teachers perceiving greater administrator support reported less emotional exhaustion and a greater sense of personal accomplishment. Not surprisingly, teacher stress was also related to burnout in expected ways. Moreover, similar to burnout, stress was also directly related to child outcomes. However, teaching quality measured as engagement was not impacted by burnout. Thus, both burnout and stress have important implications for how much benefit students with ASD receive from their educational program.

Child goal attainment outcomes are directly associated with IEP quality, teaching adherence, teacher self-efficacy, and burnout/stress.

Thus, in addition to the five variables we have shown previously to directly impact student educational outcomes—IEP quality, teaching adherence, teaching quality measured as child engagement and use of common elements, and self-efficacy, we have now identified another variable, teacher burnout/stress as critical to positive outcomes. Referring to our Integrated Model, overall, we have identified five teacher quality elements and two internal teacher factors related to child outcomes. In the following section, we describe research on an additional potential internal factor, critical attitudes that influence best practices in educating students with disabilities.

Teacher Data Collection

As mentioned, implementation science is concerned with the variables that hinder or facilitate the adoption of evidence based practices. Obviously, it takes individuals to adopt new practices. But what are some of the ways of understanding whether or not an individual teacher applies best practices in the classroom? As mentioned above, one variable became of special interest to us after we conducted our first RCT of COMPASS. During the first study it was not unusual for us to observe that the large majority of teachers did not collect data on child IEP goals. The omission of data collection as part of an overall strategy for educating students with disabilities is a potentially critical failure in an intervention focused on the achievement of individualized goals. Regular tracking of progress toward goal attainment is needed to ensure that the teaching plans are actually being effective and provides the ability to respond quickly to adjust teaching plans when progress is insufficient. For those teachers who received COMPASS, we spent time as part of the coaching sessions discussing strategies for collecting data, including how, when, and by whom. The data we used during the coaching sessions helped validate progress measured with the GAS. And while GAS was not meant to replace weekly data collection, it often was the sole measure used for progress monitoring despite our recommendations for more frequent data collection. Because data collection is a critical activity teachers are expected to implement, we wanted to understand more about teacher's beliefs and attitudes about data collection and why it was not being done. To do this, we created a measure for assessing teacher's attitudes toward data collection called the TIDE or **T**eachers **I**ntentions toward **D**ata Collection **E**fforts (Ruble et al. 2015). The measure was based on the Theory of Planned Behavior (Ajzen 1991) and identified influences on special education teachers' reports of their intention to collect data. The TIDE included items that assessed attitude toward the behavior (AB), social norms (i.e., attitudes toward data collection of one's peer group; SN), and perceived behavioral control of barriers that might hinder the behavior (PBC). We found that all three subscales demonstrated good internal consistency (Ruble et al. 2015) and correlated with teachers' report of intention to collect data (see Table 7.2). That is, consistent with the theory, having positive attitudes toward collecting data, having peers with positive attitudes toward collecting data and feeling one had control over potential barriers to data collection all strongly related to ones' intention to collect data.

Teacher self-efficacy is the strongest predictor of teacher's intention to collect data.

Table 7.2 Correlations between TIDE and teacher variables

	1	2	3	4
Intention	–			
Subjective norm (SN)	0.47**	–		
Attitude (AB)	0.44**	0.66***	–	
Behavioral control (PBC)	0.39**	0.46**	0.26	–
Self-efficacy	0.57***	0.47**	0.50***	0.45**
Administrative support	0.35*	0.51***	0.15	0.47**
Emotional exhaustion	−0.25	−0.45**	−0.47***	−0.37*
Depersonalization	−0.16	−0.43**	−0.53***	−0.16
Personal accomplishments	0.38*	0.37*	0.42**	0.16
Stress	−0.21	−0.40**	−0.38*	−0.25

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

Going back to our Integrated Model, we also wanted to understand how teacher stress, burnout, self-efficacy and administrative support influenced AB, SN, PBC and intention. As shown above, the results indicated that all three factors—stress, burnout, and self-efficacy correlated with subscales of the TIDE.

Teachers who reported increased stress and burnout reported poorer attitudes toward data collection and lower perceived social norms in support of data collection. Those teachers with higher emotional exhaustion also reported lower perceived behavioral control of data collection. Thus, similar to our findings reported earlier, the internal experiences of stress and burnout are important for understanding influences on the evidence based practice of data collection.

In addition to the negative aspects associated with data collection, we also identified facilitators or supports for data collection. Teachers who reported greater personal accomplishments, self-efficacy to collect data successfully, and administrator support reported increased intention to collect data. In fact, self-efficacy was the strongest predictor of behavioral intention of any of the assessed variables ($r = 0.57$, $p = 0.00$), including the three TPB subscales.

A final question concerned the potential impact of scores from the TPB on child goal attainment change. We found no direct effect. Goal attainment change was unrelated to intention to collect data and to the TPB subscales. However, there was an indirect effect of the TIDE subscales. Specifically, teachers' attitudes toward data collection (AB) correlated with IEP quality; and IEP quality, as reported earlier, correlated directly with child goal attainment outcomes. Although we have not yet formally tested a mediational model, attitudes toward data collection impacted the overall quality of the IEP, including the validity and viability of the IEP goals, and this in turn impacted student progress as measured by goal attainment.

External Factors. Up to this point we have primarily focused on the internal experiences of teachers, their sense of self-efficacy, burnout, stress, and experience associated with teaching. Compared to internal factors, we have limited information on external factors and this area needs much further development both in terms of

Table 7.3 Administrator support items

Items
Participating in training/continuing education
Having flexibility to be creative or develop innovative programs
Having adequate classroom staffing
Having reasonable class or caseload size
Having a range of classroom placements in which students with autism can participate
Having adequate planning time
Participating in professional organization activities
Developing an atmosphere of understanding and acceptance within the school
Providing a process for collaboration and teaming with school personnel regarding specific students
Fostering communication with parents
Having adequate space and materials

conceptualization, measures, and empirical investigation. The one external factor we have included is administrator support. Our measure of administrator support is a broad measure that asks about several areas of teaching support (see Table 7.3). We asked teachers to report how much support they received in the following areas using a 4-point Likert scale (1 'not much support' to 4 'very much support').

We applied this measure in several of our analyses described thus far. We found that administrator support correlated with important teacher internal factors of burnout (Table 7.1) and intention toward and attitudes about collecting data (Table 7.2). As a reminder, we found burnout to be directly associated both with GAS change scores and with administrator support. Thus, administrator support appears to be associated with child outcomes by its impact on teacher burnout. If teachers feel more supported, they report higher personal accomplishment. Those teachers with a higher sense of personal accomplishment are more likely to have students achieve greater GAS change.

In this chapter, we focused on what we have learned about teacher internal and external factors. The next chapter covers child internal and external factors and their impact on outcomes.

References

Ajzen, I. (1991). *The theory of planned behavior* (Vol. 50). Netherlands: Elsevier Science.

Awa, W. L., Plaumann, M., & Walter, U. (2010). Burnout prevention: A review of intervention programs. *Patient Education and Counseling, 78*(2), 184–190. doi:10.1016/j.pec.2009.04.008.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.

Bandura, A. (1986). *The explanatory and predictive scope of self-efficacy theory* (Vol. 4). US: Guilford Publications.

Boe, V. R. M. (2014). *Hearts of hope: experiences of ebd teachers and factors contributing to career longevity* (Vol. 74). US: ProQuest Information & Learning.

Boe, E. E., Cook, L. H., & Sunderland, R. J. (2008). Teacher turnover: Examining exit attrition, teaching area transfer, and school migration. *Exceptional Children, 75*(1), 7–31.

Carlson, B. C., & Thompson, J. A. (1995). Job burnout and job leaving in public school teachers: Implications for stress management. *International Journal of Stress Management, 2*(1), 15–29. doi:10.1007/bf01701948.

Cook, L. H., & Boe, E. E. (2007). National trends in the sources of supply of teachers in special and general education. *Teacher Education and Special Education, 30*(4), 217–232.

Jennett, H. K., Harris, S. L., & Mesibov, G. B. (2003). Commitment to philosophy, teacher efficacy, and burnout among teachers of children with autism. *Journal of Autism and Developmental Disorders, 33*(6), 583–593.

McLeskey, J., & Billingsley, B. S. (2008). How does the quality and stability of the teaching force influence the research-to-practice gap? A perspective on the teacher shortage in special education. *Remedial and Special Education, 29*(5), 293–305.

McLeskey, J., Tyler, N. C., & Saunders Flippin, S. (2004). The supply of and demand for special education teachers: A review of research regarding the chronic shortage of special education teachers. *Journal of Special Education, 38*(1), 5–21.

Nichols, S. M. C., Bicard, S. C., Bicard, D. F., & Casey, L. B. (2008). A field at risk: the teacher shortage in special education. *Phi Delta Kappan, 89*(8), 597–600.

Ruble, L. A., Usher, E. L., & McGrew, J. H. (2011). Preliminary investigation of the sources of self-efficacy among teachers of students with autism. *Focus on Autism and Other Developmental Disabilities, 26*(2), 67–74. doi:10.1177/1088357610397345.

Ruble, L. A., Toland, M. D., Birdwhistell, J. L., McGrew, J. H., & Usher, E. L. (2013). *Preliminary study of the autism self-efficacy scale for teachers (ASSET)* (Vol. 7). Netherlands: Elsevier Science.

Ruble, L. A., McGrew, J. H., Wong, W. H., Missall, K. N., & Guskey, T. R. (2015). Understanding influences on special education teachers' intention to collect data. Submitted.

Ryan, T. A. (1970). *Intentional behavior: An approach to human motivation*. Oxford, England: Ronald Press.

Soto, G., & Goetz, L. (1998). *Self-efficacy beliefs and the education of students with severe disabilities* (Vol. 23). US: Assn for Persons with Severe Handicaps.

U.S. Department of Education July 15, 2008. U.S. Department of Education, Office of Special Education Programs, Data Analysis System, OMB #1820–0518 (available on Department. of Education, Office of Special Education Programs website).

Chapter 8
Child Internal and External Factors

In our prior chapters, we described what we have learned about our implementation practice and intervention practice in terms of quality and internal and external factors and investigated the potential impact of other variables in explaining variability in COMPASS outcomes. In this chapter we focus on the potential impact of child internal and external factors on outcomes, as described in our Integrated Model (Fig. 8.1).

The decision-making that is central to the EBPP approach for client care is particularly relevant for the treatment of ASD (see Fig. 1.1, Chap. 1). Early intervention treatment studies demonstrate repeatedly that about 50 % of children with ASD show a marked response to high-quality early intensive behavioral intervention (Lovaas 1987; McEachin et al. 1993; Sallows and Graupner 2005; Smith et al. 2004). While this is good news, this also means that that the other half do not show significant change. Some of the factors that help explain the variability in treatment response include cognitive level, language ability, and severity of autism symptoms (Reichow 2012). For ASD, one of the most commonly investigated client variables is autism severity. That is, children with more severe autism are often viewed as being more difficult to help. Another issue is that many interventions available today are not necessarily developed to help all subgroups of those with ASD. For example, cognitive based treatments are less useful and effective for nonverbal children with lower IQs. Thus, a critical question was the extent to which COMPASS was able to help all children with ASD, regardless of severity or symptom profile. A subsidiary question was whether teacher behavior, as measured by adherence, varied in terms of child factors.

Question 1: To what extent is COMPASS able to help all children with ASD, regardless of severity or symptom profile? Recall our study design. For both RCTs, our inclusion criteria were that the children had an IEP that was based on their eligibility of having autism. They also needed to have a confirmed diagnosis of autism based on an evaluation conducted by us. Thus, children in both RCTs received a comprehensive assessment that included measures of autism severity, IQ, language, adaptive behavior, and interfering behavior. These baseline evaluations were conducted prior to group assignment. To answer the first question about the

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*,
SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2_8

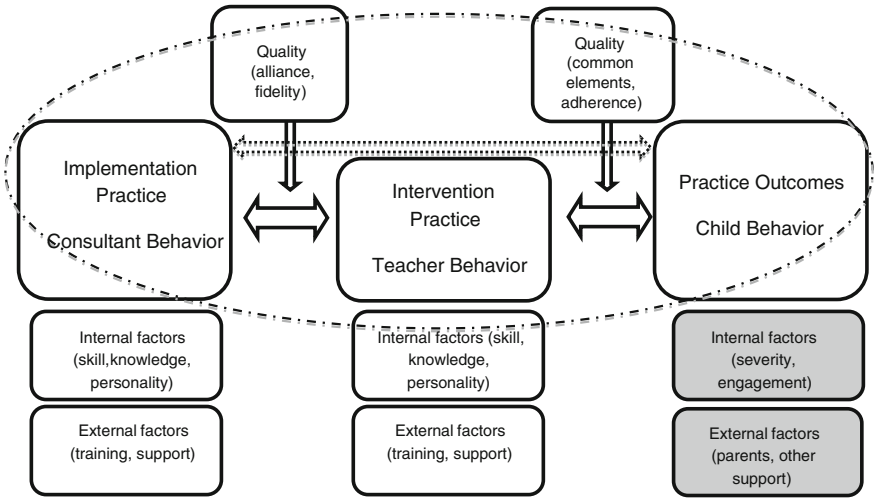


Fig. 8.1 Integrated model with focus on child internal and external factors

impact of child factors on outcomes, we correlated GAS change with all four child characteristics: autism severity, IQ, language, and adaptive skills (see Table 8.1). Bivariately, consistent with other intervention research in ASD, we found that treatment response was greater for students with less severe symptoms, higher IQ, greater language ability, and greater adaptive functioning.

To help us understand if these factors indexing severity were associated with and potentially explained by / mediated by other child factors or by more mutable intervention factors, such as engagement, we conducted further analyses. Accordingly, we used multiple regression to examine the independent impact on outcomes of these

Table 8.1 Intercorrelations between child and teacher variables

	1	2	3	4	5	6	7	8
1. GAS change	–							
2. Coaching 1 adherence	0.25	–						
3. Coaching 2 adherence	0.32*	0.31*	–					
4. Coaching 3 adherence	0.24	0.29*	0.42**	–				
5. Coaching 4 adherence	0.34*	0.30*	0.45**	0.56**	–			
6. Autism severity	−0.41**	−0.06	−0.06	0.09	0.03	–		
7. IQ	0.32*	−0.05	−0.00	−0.08	−0.03	−0.41**	–	
8. Language	0.31*	0.04	−0.01	−0.06	−0.07	−0.53**	0.73**	–
9. Adaptive	0.33*	−0.01	−0.05	−0.22	−0.22	−0.50**	0.74**	0.76**

*p < 0.05; **p < 0.01

severity indicators when entered together with other significant child predictors including engagement. Specifically, we regressed six child behaviors (IQ, language, adaptive behavior, problem behavior, age, and engagement) against GAS change. We found that engagement emerged as the only significant child predictor of goal attainment outcomes multivariately. Thus, the impact of autism severity was no longer significant when engagement was considered (Ruble and McGrew 2013). This is potentially important, because it implies that these relatively stable and immutable severity factors, do not necessarily limit COMPASS impact. Instead, child engagement may be the critical factor impacting outcomes. We discussed child engagement in our chapter on teaching quality. We presented evidence that engagement is correlated with teaching quality, which in turn is correlated with child goal attainment outcomes. Thus, engagement, unlike IQ or autism severity, can be manipulated through high quality teaching.

The greater the need for help, as indexed by child severity ratings, the higher the ratings of helpfulness of the intervention.

Another commonly used measure of treatment impact is satisfaction. We assessed teacher satisfaction with COMPASS at both the initial consultation and for coaching contacts. When we examined the association between satisfaction and child severity, we did identify one interesting finding. Using our combined data from both RCTs, there was a negative association between child IQ and teacher satisfaction with the initial consultation ($r = -0.34$, $p = 0.03$) and follow up coaching sessions ($r = -0.49$, $p = 0.00$), indicating that teachers whose students were rated as more severe tended to be more satisfied with COMPASS. In other words, the greater the need for help, as indexed by child severity ratings, the higher the ratings of helpfulness of the intervention. Thus, despite showing less improvement, consultees (teachers) reported COMPASS was of greater benefit to them for severe students.

Question 2: Are teachers able to achieve similar levels of teaching quality as measured by adherence, for all children, irrespective of severity or other child variables? To test the second question, we correlated teacher adherence ratings with child characteristics. Importantly, child factors did not relate to how well teachers were able to implement the intervention programs. Table 8.1 shows the correlations between teacher adherence at each coaching session and child characteristics. The data show that there were no relationships between a child's level of autism severity, IQ, language, or adaptive behavior and teacher adherence ratings. We also analyzed the data using a t-test after first dividing teachers into low and high performing groups. As with the correlation analyses, there were no differences

between the low and high performing teachers on autism symptomatology ($t = -0.17$, $p = 0.87$), IQ ($r = -1.6$, $p = 0.12$), adaptive behavior scores ($r = -1.4$, $p = 0.18$) and language abilities ($t = -0.67$, $p = 0.51$).

Thus, differences in teacher adherence are not due to differences in child factors. This is good news because it suggests that teacher behaviors, which are more amenable to change, are not associated with child characteristics that often explain differential response to treatment. These findings together with the findings that, for more severe students, treatment satisfaction is greatest and child engagement underlies obtained outcome differences, lend confidence that COMPASS is helpful for those across the autism spectrum.

Our last analysis related to child factors was trying to understand other reasons for children's differential responsiveness to COMPASS. To begin to understand if we could identify children whose teachers were not performing, we separated our sample from study 2 into two groups: (a) a high performing group—students who fell within the upper half of GAS improvement scores and (b) a low performing group—students who fell in the bottom half of GAS improvement scores. We then examined teacher adherence as measured by how many of the components of the teaching plan were implemented for each coaching session for each group. Recall that the adherence scores reflected a percentage of components implemented from the teaching plans using a scale from 1 to 5. The highest score, 5, meant that 80 % or more of the teaching plans were implemented. The Fig. 8.2 shows the results of teacher adherence data over time based on GAS scores that fell in the upper half and GAS scores that were in the bottom half. Visual inspection of the data show that students who obtained GAS scores in the upper half had teachers who were implementing a larger number of elements of the teaching plans at each time point, and this difference was evident early in the coaching process, with at least 10 % more of the elements implemented compared to the low performing group. Also notable is the irregularity of implementation of teaching plans of teachers whose students who obtained GAS scores in the lower half. In fact, even fewer elements were implemented by coaching session 2, compared to coaching session 1, and no difference was noted between coaching sessions 3 and 4. So while the high performing group made steady progress after the second coaching session, the low

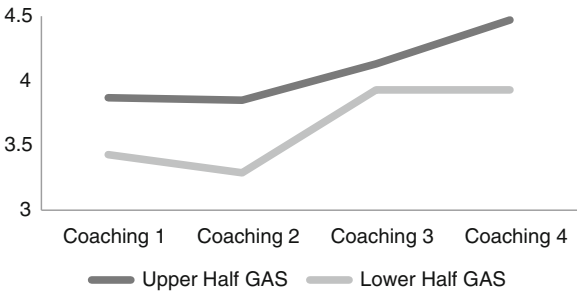


Fig. 8.2 Teacher adherence by coaching session and GAS outcome

performing group declined after session 1, made progress during session 3, and remained stable for session 4. Also important was that a statistically significant difference in teacher adherence was observed at coaching session 4 between the two groups ($t = -2.1$, $p = 0.04$). Clearly more research is needed to understand teacher adherence and how to individualize the coaching process to help all teachers achieve good adherence. Whether coaching sessions need to be offered more frequently, involve direct modeling of the implementation of the teaching plans, or some combination of the two remains for future empirical studies.

References

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*(1), 3–9.

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97*(4), 359–372.

Reichow, B. (2012). *Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders* (Vol. 42). Germany: Springer.

Ruble, Lisa, & McGrew, John H. (2013). Teacher and child predictors of achieving iep goals of children with autism. *Journal of Autism and Developmental Disorders,*. doi:10.1007/s10803-013-1884-x.

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation, 110*(6), 417-438. doi: 10.1352/0895-8017(2005)110[417:ibtfcw]2.0.co;2

Smith, T., Groen, A. D., & Wynn, J. W. (2004). Randomized Trial of Intensive Early Intervention for Children with Pervasive Developmental Disorder, p. 153.

Chapter 9
Conclusion

In our concluding chapter, we review what we have learned using our Integrated Model below as a guide. Recall that our framework incorporates the Implementation Science Model from Dunst et al. (2013). The implementation practice refers to the methods used by the COMPASS consultant/coach to support the transfer of the intervention practice to the teacher, which should result in improved child educational outcomes. As we have continued to emphasize throughout, consultation research is a complex task because we need to study both the implementation practice and the intervention practice (Fig. 9.1).

Additional predictive factors in the model are the quality of consultant and teacher behavior as well as the internal and external factors associated with the consultant, the teacher, and the student. In our work, COMPASS has served as the evidence-based implementation strategy proven to result in better educational outcomes for children with ASD. The bulk of our work has focused on the intervention practice that is the link between COMPASS (what the consultant does with the teacher) and child outcomes (what the teacher does with the child).

Each of the three primary practice areas in our Integrated Model represents interdependent activities that are both distinct and also linked to each other. In other words, the quality of the implementation practice (COMPASS consultant fidelity) should be associated with the quality of the intervention practice (teacher fidelity), which in turn should be associated with the effectiveness of the practice outcomes (child educational goal attainment). In this chapter, we will describe our main findings for each of the three areas and also areas for future research.

So far we have described several frameworks that help explain the need for COMPASS and our study of its effectiveness (Chaps. 1 and 2). We have described our approach for the development and testing of an appropriate outcome measure (Chap. 3). We also presented evidence that COMPASS is an effective implementation practice (Chap. 4) and even works using web-based technology for improving child educational outcomes (Chap. 5). We also described our process for establishing evidence of at least two active ingredients that help explain why

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*,
SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2_9

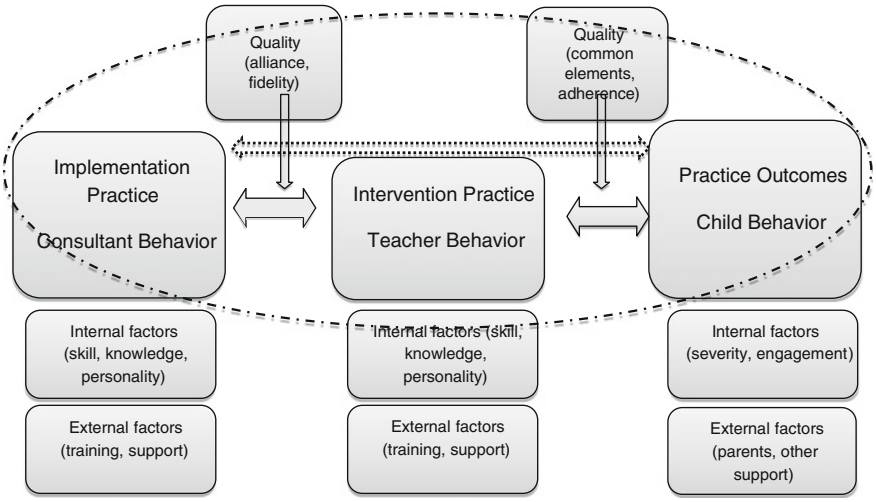


Fig. 9.1 Integrated model

COMPASS works as well as other important quality measures (Chap. 6). Teacher and child internal and external factors that are associated with outcomes were also reviewed (Chaps. 7 and 8). Our main findings are listed in the Fig. 9.2.

- COMPASS is Effective
- Web-based Coaching is Effective
- Teachers are Satisfied with COMPASS
- Goal Attainment Scaling is a Valid and Reliable Method for Measuring Outcome
- Teacher adherence and IEP Quality are Active Ingredients
- Teacher Engagement and Use of Common Elements are Effective
- Coaching Improves Teacher Adherence
- Teacher Internal Factors of Stress, Burnout, and Self-Efficacy Are Important for Improving Child Outcomes
- COMPASS is Helpful for Children Across the Spectrum
- Differences in COMPASS Effectiveness are Best Explained by Child Engagement, Rather than IQ or Autism Severity

Fig. 9.2 COMPASS main findings

Despite all the information we have learned, there are several questions that remain. With regard to our Integrated Model, there are still several gaps in our knowledge. We start with the Implementation Practice. One of our most urgent needs is to demonstrate that we can transfer our skills as a COMPASS consultant to school-based consultants. We have preliminary evidence that COMPASS can be taught. We developed a training package using direct and online methods and our manualized protocol to teach graduate students seeking a certification in ASD. The training package builds upon materials piloted by five graduate students (three of whom were special education classroom teachers of children with autism) enrolled in a course on COMPASS. The final course assignment was to implement a COMPASS consultation with a child they did not know, the child's parent and teacher, as well as to collect fidelity assessments. Adherence and satisfaction ratings completed by teachers and parents showed promising results. Parent and teacher satisfaction ranged from 3.0 to 4.0 on a 4-point Likert scale (4 is highest). Adherence ratings ranged from 88 to 100 %. However, these results are preliminary and have not been tested rigorously. We are currently seeking funding to test a method of training COMPASS consultants and to demonstrate that consultants who are independent from the primary research team can be trained to deliver COMPASS with high fidelity and with outcomes similar to those reported in our RCTs.

Another important gap in knowledge concerns treatment responsiveness as measured by the intervention practice. Understanding what factors might explain low and high performing results is essential for understanding how to help as many children benefit from COMPASS as possible. We have thought about treatment responsiveness in several ways. For example, we examined whether success was equal across schools, across studies, across delivery method (face to face versus internet), and across client factors (autism severity). We found that these factors were not ultimately important for predicting success. Nevertheless, we did find that not all teachers were able to achieve success following COMPASS intervention. Despite having received COMPASS delivered by highly trained consultants, teachers had variable responses as measured by adherence / teaching quality scores, and as a result, so did their student. This suggests that for some teachers, the standard delivery of COMPASS requires an adapted or modified approach to ensure all children benefit as much as possible. Instead of waiting until the end of the school year to determine children's outcomes, we wondered if there were early warning red flags that could help us identify and predict those teachers who were in trouble and not likely to make progress early in the coaching process.

In conclusion, we have presented the evidence for the effectiveness of COMPASS and factors that help explain outcomes. Although we are pleased with our progress in understanding COMPASS, clearly more work is necessary to ensure that COMPASS can be disseminated reliably and that all children can benefit.

Reference

Dunst, C. J., Trivette, C. M., & Raab, M. (2013). An implementation science framework for conceptualizing and operationalizing fidelity in early childhood intervention studies. *Journal of Early Intervention, 35*(2), 85–101.

Index

Note: Page numbers followed by "f" and "t" indicate figures and tables respectively

L.A. Ruble and J.H. McGrew, *COMPASS and Implementation Science*, SpringerBriefs in Child Development, DOI 10.1007/978-3-319-18555-2