social sciences

MDPI

Review

Meta-Analysis of Parent Training Programs Utilizing Behavior Intervention Technologies

Kimberly B. Bausback [1,2,*] and Eduardo L. Bunge [1,2,3]

[1] Department of Clinical Psychology, Pacific Graduate School of Psychology, Palo Alto University, Palo Alto, CA 94304, USA; ebunge@paloaltou.edu
[2] The Children and Adolescent Psychotherapy (CAPT) Research Lab, Palo Alto University, Palo Alto, CA 94304, USA
[3] *i*Health Institute for International Internet Interventions for Health, Palo Alto University, Palo Alto, CA 94304, USA
* Correspondence: kbausback@paloaltou.edu

Abstract: Behavioral Parent Training (BPT) traditionally occurs in face-to-face (FTF BPT). Recently, Behavioral Intervention Technology (BIT) has been developed to deliver BPT in lieu of or as an adjunct to FTF BPT using websites, computer software, smartphone applications, podcasts, pre-recorded sessions, and teletherapy. The present meta-analysis reviews BIT BPT randomized control and comparison studies to determine the overall efficacy of BITs, if the level of human support significantly effects BIT BPT treatment outcomes, and which populations BIT BPT are effective for, by analyzing the following study variables: socioeconomic status, race, and clinical population. The analyses indicated that, overall, BIT BPT is an effective treatment ($g = 0.62$), and did not indicate a significant difference between levels of human support (χ^2 (3) = 4.94, $p = 0.18$). Analysis did indicate a significant difference between studies that used waitlist or education control groups, compared to studies that used active treatment controls (χ^2 (1) = 12.90, $p = 0.00$). The analyses did not indicate a significant difference between clinical population, low socioeconomic status, and racial minority studies. These findings provide preliminary evidence that BIT BPT is effective for treating child and adolescent externalizing behavior in a variety of populations.

Keywords: Behavioral Intervention Technology; Behavioral Parent Training; child and adolescent

Citation: Bausback, Kimberly B., and Eduardo L. Bunge. 2021. Meta-Analysis of Parent Training Programs Utilizing Behavior Intervention Technologies. *Social Sciences* 10: 367. https://doi.org/10.3390/socsci10100367

Academic Editor: Nigel Parton

Received: 1 September 2021
Accepted: 17 September 2021
Published: 29 September 2021

1. Introduction

Behavioral disorders are one of the most common reasons youth are referred to psychotherapy (Egger and Angold 2006; Zisser and Eyberg 2010) and account 19.1% of psychological disorders in youth, making it one of the most common youth psychological disorders (Comer et al. 2013; Merikangas et al. 2010). These behaviors tend to occur in adolescence (Fuentes et al. 2020) and may continue into young adulthood (Steinberg 2007). If untreated, externalizing behavioral disorders lead to increased high school dropout rates, higher rates of incarceration, increased unemployment, higher rates of substance abuse, and lack of psychosocial maturity leading to increased emotional and interpersonal problems (Able et al. 2007; Fuentes et al. 2020; Liu 2004; Steinberg 2007). Externalizing behavior disorders account for significant medical and social financial costs. The cost per year, per child is estimated to be $14,576 (Pelham et al. 2007) for attention deficit/hyperactivity disorder and $12,547 and $6630 for conduct disorder and oppositional defiant disorder, respectively (Foster et al. 2005).

Currently, the gold standard evidence-based treatment for externalizing disorders is behavioral therapy conducted with the parents: Behavioral Parent Training (BPT) (Anil Chacko et al. 2017; Chorpita et al. 2011; Farmer et al. 2002). Many children have primary caregivers that are not their biological parents; however, based on the nomenclature utilized in evidence-based treatment manuals and extant literature, all caregivers will be referred

to as "parents". Parenting and its relationship to child behavior is a long-researched topic. BPT improves child behavior by providing parenting skills and knowledge to parents of children, adolescents, and teenagers. BPT is based on the principle that parents are a main source of influence on their children, particularly during childhood and adolescence, but even once parent socialization is over (Gimenez-Serrano et al. 2021). Parenting research has determined that factors, such as parental warmth versus strictness influence child and adolescent behavior (Garcia et al. 2020; Musitu Ochoa et al. 2012). Parental warmth is related to greater adjustment and competence (Garcia et al. 2020; Fuentes et al. 2015). Subsequently, BPT interventions aim to increase parental warmth as a source of social support for child (Desatnik et al. 2021; Grolnick et al. 2021), while also combining principles of operant conditioning to promote desired behaviors.

There are a variety of existing face-to-face (FTF) BPT manuals with strong research support, such as The Incredible Years (Webster-Stratton and Reid 2003), Parent Management Training-Oregon Model (Forgatch and Patterson 2010), Parent-Child Interaction Therapy (Eyberg et al. 1995), Triple-P Positive Parenting (Sanders 1999), and Helping the Noncompliant Child (Forehand and McMahon 1981). Several meta-analyses on FTF BPT for disruptive behavior disorders found effect sizes in the small to large range (e.g., Comer et al. 2013; Lee et al. 2012; Leijten et al. 2013; Lundahl et al. 2006; Mingebach et al. 2018; Maughan et al. 2005; Serketich and Dumas 1996).

Despite the strong evidence base and empirical evidence supporting BPT, there are several barriers to accessing and participating in these interventions. Weisenmuller and Hilton (2020) note that there are several systemic, cultural, and individual factors challenges that decrease access to BPT, particularly for underserved populations, such as low socioeconomic status families and rural populations. Systemic barriers include both lack of insurance that provides mental health care and an insufficient number of mental health providers. Further, participants are reluctant to engage in treatment due to stigma and conflicts with cultural and religious beliefs. Research found 25% of individuals referred to BPT do not enroll; for those that do enroll, dropout rates are estimated to be 26%, resulting in an approximate attrition rate of 51% (Chacko et al. 2016; Nock and Ferriter 2005). Research suggests that stigma (e.g., being perceived as a "bad parent"), gender factors (e.g., being male in a primarily female group), low socioeconomic status, and lack of time and resources are factors that contribute to higher rates of attrition (Chacko et al. 2016; Mytton et al. 2014). Kazak et al.'s (2010) meta-systems analysis emphasized the importance of developing evidence-based practices that are efficacious that are cost-effective and easy to implement and disseminate. Therefore, while FTF BPTs are effective, there is a need to evaluate evidence-based treatment delivered through technology to meet the needs of mental health consumers, particularly for underserved groups.

Behavioral Intervention Technology (BITs) may mitigate these barriers to treatment by providing BPT in a more accessible, engaging modality. BITs are technology designed to treat psychopathology by modifying behavior (Mohr et al. 2013), and they include the use of technology as a part of psychotherapy (e.g., smartphone applications, computer programs, virtual reality, wearable technology, robots, video messaging, and electronic messaging). There are several different types of BIT, which require varying levels of human support and range from adjuncts to face-to-face (FTF) therapy to fully automated BIT. More specifically, Muñoz (2017) outlines the spectrum of BIT as traditional FTF therapy without the use of technology, traditional FTF therapy with BIT as adjuncts, guided BIT with human support as adjuncts, and fully automated BIT. Research suggests that BITs are used more frequently, are more engaging, and have larger effect sizes when they have some level of human support (Andersson and Cuijpers 2009; Baumeister et al. 2014; Day and Sanders 2017; Schueller et al. 2017). While human support may increase engagement and efficacy of BITs, the level of human support impacts the cost and scalability of treatment (Schueller et al. 2017).

Recently, BITs have been specifically designed to either complement or supplement FTF BPT interventions. Both novel BIT BPT interventions have been created and FTF

BPT have also been adapted to be delivered through BIT (e.g., Triple P was converted to an online intervention, Triple P Online (TPOL); Sanders 1999). There are a variety of randomized control trials evaluating Teletherapy, FTF BPT with BIT as adjuncts, BIT with human support, and fully automated BIT: active and passive (Baker et al. 2017; Breitenstein et al. 2016; Cefai et al. 2010; Comer et al. 2017; Dadds et al. 2019; Day and Sanders 2018; DuPaul et al. 2018; Enebrink et al. 2012; Franke et al. 2016; Ghaderi et al. 2018; Irvine et al. 2015; Jones et al. 2014; Nixon et al. 2003; Porzig-Drummond et al. 2015; Rabbitt et al. 2016; Sanders et al. 2000, 2008, 2012, 2014; Sourander et al. 2016; Stormshak et al. 2019; Wetterborg et al. 2019; Xie et al. 2013).

Additionally, several systematic reviews and meta-analysis have also provided support for the use of BITs BPT (Baumel et al. 2016; Corralejo and Rodríguez 2018; Nieuwboer et al. 2013; Spencer et al. 2019; Thongseiratch et al. 2020), with effect sizes in the small to large range (0.22 to 0.67). Previous analyses found that BIT BPT are more effective for children whose disruptive behavior was in the clinical range (Baumel et al. 2016; Spencer et al. 2019) and when the BIT is interactive (e.g., computer game versus video; Baumel et al. 2016). There were conflicting results on whether human support significantly increases the efficacy of BIT BPT (Spencer et al. 2019; Thongseiratch et al. 2020). However, many of these studies had small sample sizes or examined the effects of BIT BPT on multiple psychological disorders (Nieuwboer et al. 2013; Spencer et al. 2019). Further, extant research primarily validates BIT BPT in White American individuals. Thus, it is paramount that future studies evaluate the efficacy in BIT BPT with racial and ethnic minorities (Corralejo and Rodríguez 2018).

While current meta-analyses support the efficacy of BIT BPT, research largely does not examine the moderating effects of human support, control group, socioeconomic status, race, and clinical sample population. Additionally, previous meta-analysis study selection criteria vary from including only studies that compare BIT BPT with waitlist, education, or no treatment controls, while other meta-analyses also include studies that compare BIT BPT with other FTF or BIT treatment; it is important to understand how a control group impacts overall efficacy of BIT BPT. Further, it is important to understand how to deliver BIT BPT in the most efficient manner (Schueller et al. 2017). Extant research largely does not examine the effect of the level of human involvement in BIT, which is important for developing future BIT and determining the cost-effectiveness of fully automated BIT compared to BIT with human support. Additionally, it is important to not only measure if BIT BPT are effective but also for whom (Comer and Myers 2016). Previous research does not analyze how SES impacts the efficacy of BIT BPT (Baumel et al. 2016), which is important given that low SES individuals have higher attrition rates in FTF BPT (Chacko et al. 2016). Furthermore, BIT BPT literature largely excludes research on underserved populations and lacks racial diversity; therefore, additional analysis of the effect of BIT BPT with racial minorities is needed (Corralejo and Rodríguez 2018). Additionally, it is important to determine if BIT BPT is effective interculturally. Moderation and subgroup analysis should be conducted in order to determine what type of BIT is most effective. Through this analysis, the gap between efficacy and effectiveness can be better understood.

The aims of the current meta-analysis are to: Aim 1: To examine if BIT BPT reduces externalizing behavioral disorder symptoms by evaluating the pre- and post-effect sizes of BIT BPT studies. Aim 2: To determine the effects of varying levels of human support (i.e., teletherapy, FTF BPT with BIT as adjuncts, BIT BPT with human support, fully automated active BIT, and fully automated passive BIT) on the efficacy of BIT BPT, and if there is a moderating effect. Aim 3: To determine if study sample criteria for externalizing behavior (i.e., clinical versus nonclinical studies) has a moderating effect on BIT BPT. Aim 5: To determine if low socioeconomic (SES) status has a moderating effect on BIT BPT. Aim 6: To determine if racial minority status has a moderating effect on BIT BPT.

2. Materials and Methods

2.1. Selection Process of Articles

The articles included in this meta-analysis were found on an internet-based information search conducted in April of 2020. Computer searches of PsycINFO, PsychARTICLES, and SciELO databases were conducted for all published studies between 2000 and 2020. The following Boolean search phrase was used to search by title or abstract: *(behav* parent training OR parent management training OR parent training OR parent-child interaction therapy OR parent child interaction therapy OR parent* program* OR parent* intervention) AND (technolog* OR video OR internet OR net OR web* OR virtual reality OR augmented reality OR mobile OR text-messaging OR texting OR smartphone* OR app* OR comput* OR wearables OR artificial intelligence OR bots OR robots OR chat OR online OR digital OR tele* OR eHealth OR "e-Health" OR mhealth OR m-health)*. Additionally, this Boolean phrase was used to narrow down the search through all text: *(CBCL OR Child Behavior Checklist OR ECBI OR Eyberg Child Behavior Inventory OR SNAP-IV OR Swanson, Nolan, and Pelham Questionnaire OR Vanderbilt Assessment Scale OR NICHQ Vanderbilt Assessment Scale OR Conners 3 OR Conners Comprehensive Behavior Rating Scales OR Conners CBRS OR CERS OR Conners Early Childhood Rating Scale OR SDQ OR Strengths and Difficulties Questionnaire OR BASC OR Behavior Assessment System for Children)*. See Figure 1 for the selection process of the articles.

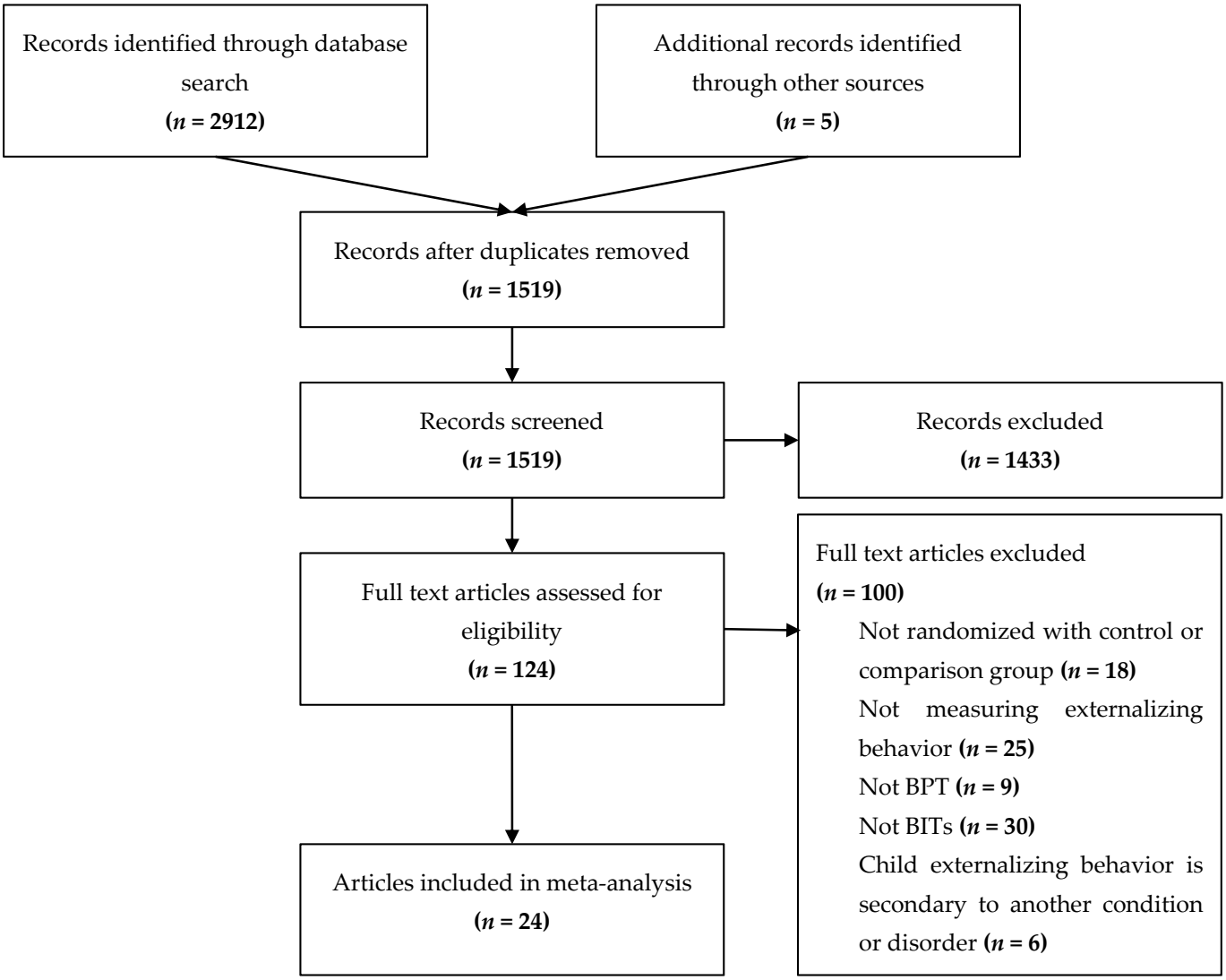

Figure 1. Selection process of articles.

Criteria for inclusion of studies in this meta-analysis were studies that evaluated behavioral parent training for parents of children 18 years old or younger with disruptive/externalizing behaviors. The search included studies reporting on a parent training intervention targeting their child's disruptive behavior problems as measured by pre-/post-intervention parent-report on a well-validated assessment measure. Only randomized control or comparison trials published in English and peer-reviewed were included. Programs or interventions needed to use technology as a primary mode of delivery to be included.

Studies were excluded if the primary intervention did not target externalizing behavior. Specifically, studies were excluded if the externalizing behavior was caused by a medical condition or traumatic brain injury. Studies in which the externalizing behavior was secondary to a physical, mental, or neurodevelopmental disability were excluded. However, studies in which participants had comorbid psychological disorders (e.g., mood disorders) were not excluded, as long as they were secondary to externalizing behavior and the primary target of the intervention was child behavior. For articles that only included technology as a component of in session, FTF treatment (i.e., brief video vignettes used in session) were excluded. Studies that analyzed BPT with video vignettes as a *minor* component to self-directed treatment (e.g., a brief skills video that accompanies a self-directed treatment manual) were not considered BIT and, thus, not included in this study. Telephone calls and self-directed treatment using manuals (i.e., bibliotherapy) were excluded.

This search yielded results with 2912 articles; five additional articles were included that were cited in articles or previous BIT BPT meta-analyses. From the initial search, 1398 duplicate articles were removed. Of the remaining 1519 articles, 1433 were screened and were removed due to inclusion and exclusion criteria. Lastly, 124 full text articles were assessed for eligibility, and 100 were excluded due to inclusion and exclusion criteria, thus resulting in 24 articles that met inclusion criteria. The final review included 24 randomized control or comparison trials (RCTs) published between 2000 and 2020 (total n = 3957).

2.1.1. Classification of Behavioral Parent Trainings Programs

Based on the recommendations by Muñoz (2017), a BPT was considered a traditional FTF with BIT as adjuncts when the primary source of treatment occurred FTF, and technology was only a complement to FTF treatment (e.g., FTF sessions with videos or smartphone applications accessed between FTF sessions). BPT were considered guided BIT with human support as adjuncts when the intervention was primarily delivered through technology, and humans were only supporting BIT in a facilitative capacity (e.g., self-directed treatment delivered online with weekly telephone check-ins from a therapist). A BPT was categorized as a fully automated BIT when the treatment occurred only online or via software without any human support. Fully automated BIT was considered active when the intervention included an interactive component (e.g., an interactive website), and it was considered passive (fully automated passive BIT) when the intervention only included passive viewing or listening (e.g., videos and podcasts). An article was considered telemental health when the treatment was traditional FTF treatment but delivered using an online video-conferencing service.

Additionally, consistent with previous research on disadvantaged populations (Chacko et al. 2016; Leijten et al. 2013; Lundahl et al. 2006), articles were dichotomously categorized as low SES or non-low SES studies. Studies were considered low SES when they indicated most of the participants were low SES, or when the majority of the participants (i.e., over 50%) were below the poverty line (i.e., based on Federal Income Guidelines), had a Hollingshead index of 30 or less, or had a National Statistics Socio-economic Classification (NS-SEC) of five or less. Similarly, studies were considered racial minority studies when the study indicated that the majority of participants were members of a racial minority or based on the summed percentage (i.e., over 50%) of racial minorities included in the study. Dichotomous categorization was also used to classify study samples as clinical or nonclinical. Studies were considered clinical when their inclusion criteria required that participants met clinical cut-offs for externalizing behavior on a well-validated externalizing behavior measure or either DSM-IV or DSM-5 diagnosis. Lastly, studies were categorized

as active control or waitlist or education control studies. Studies were considered active control studies when the control group was another BIT or FTF BPT.

2.1.2. Methodological Quality

To ensure the 24 studies included in this meta-analysis were of sound quality, the methodological quality of the articles were evaluated using the 26-item Single-Case Reporting Guideline in BEhavioral Interventions (SCRIBE; Tate et al. 2016) (Appendix A) guidelines for study design. Due to the unique nature of BIT BPT, including that participants cannot be blind to condition and high rates of attrition, SCRIBE was selected as the methodological quality rating system because it is a comprehensive list of study components with research support (Lobo et al. 2017). Two blinded, independent raters assessed the methodological quality of each article (KB and ES). For each study, raters assigned a score of one (yes) or zero (no) for each of the 26 SCRIBE items, and the total score was summed for a numerical quality score.

2.2. Measures

This meta-analysis included studies that utilized well-validated measures of child externalizing behavior. The following outcome measures use similar, parent-report questionnaires regarding child problem behaviors (i.e., defiant, aggressive, and oppositional behavior). Measures that did not utilize parent-report of child behavior were not included in order to minimize error. Further, only the externalizing *behavior* measure scales were included. Measure scales that examined non-behavioral symptoms (e.g., inattention/hyperactivity, mood disorder symptoms) were not included in order to minimize potential error introduced from including multiple constructs.

The Child Behavior Checklist (CBCL; Achenbach 1991; Achenbach and Rescorla 2000) is a parent-, teacher-, and self-report measure that analyzes internalizing and externalizing disorders and social functioning utilizing a 3-point Likert scale. The preschool age parent-report form is a 100-item questionnaire for children between 1.5 years old to 5 years old, and it has high test-retest reliability (CBCL Externalizing = 0.87; Achenbach and Rescorla 2000). The school-age parent-report is for parents of children between 6 and 18 years old and is 120 items, and it also has excellent test-retest reliability (CBCL Externalizing = 0.94; Achenbach and Rescorla 2000).

The Conners Early Childhood Rating Scale (CERs; EC-BEH; Conners 2009) is a 190-item parent-report, 184-item teacher-report for children between the ages of 2 and 6 years old. This form gathers information regarding developmental milestones in addition to its behavior scales. The behavior scales include mood symptoms, social functioning, inattention/hyperactivity, and defiant/aggressive behavior. A copy of the technical manual was not accessible; thus, information regarding the reliability of each individual scale was not able to be obtained. Instead, the median test-retest reliability of all the scales was used for the Defiant/Aggressive scale. The Conners Early Childhood Rating Scales have excellent test-retest reliability (CERs = 0.92; Conners 2018).

The Conners' Parent Rating Scale—Revised (CPRS–R; Conners et al. 1998) is a 27-item, parent-report ADHD assessment tool for children between the ages of 3 and 17 years old. It has four subscales: Oppositional, Cognitive Problems, Hyperactive-Impulsive, and ADHD Index. The Oppositional subscale has questionable test-retest reliability (CPRS-R Oppositional = 0.60; Conners et al. 1998).

The Eyberg Child Behavior Inventory (ECBI; Eyberg and Ross 1978; Eyberg and Ross 1999) is a parent-report form that measures problem behavior and conduct and its frequency in children and adolescents. It is a 36-item report for parents of children between the age of 2 and 16 years old. Test-retest reliability is in the good range (ECBI Problem Score = 0.88; Robinson et al. 1983).

The National Institute for Children's Health Quality (NICHQ) Vanderbilt Assessment Scales (Wolraich 2002) include a 55-item parent report and a 43-item teacher report to assess children between 6 and 12 years old. This measure assesses ADHD symptoms

and includes screening questions for ODD, CD, and mood disorders. This measure has excellent test-retest reliability (Vanderbilt Oppositional Defiant Disorder/Conduct = 0.95; Bard et al. 2013).

The Strengths and Difficulties Questionnaire (SDQ; Goodman 1997) is a 25-item parent-, teacher-, and self-report assessment for children 4–17 years old that measures emotional symptoms, conduct problems, hyperactivity-inattention, peer problems, and prosocial behavior. The test-retest reliability on the parent measure is in the questionable range (SDQ Conduct Problems = 0.64; Goodman 2001).

2.3. Data Synthesis

Comprehensive Meta-Analysis Version 3.3.07 (CMA; Borenstein et al. 2014) was used to perform a pretest-posttest control group design (PPC; Morris 2008). The PPC compares the pre- and post-test assessment scores for the treatment and control or comparison conditions, allowing an evaluation of overall change compared to the non-treatment group. In order to calculate overall effect size, the standardized mean difference between groups' pre- and post-assessment difference with a corrected inverse-variance weighted effect size was used (Hedges's *g*; Hedges and Olkin 1985). Where possible, this was calculated using the pre- and post-assessment means, standard deviations, and sample sizes reported in the included articles (Hunter and Schmidt 2004). In articles that did not report standard deviation (Baker et al. 2017; Ghaderi et al. 2018; Sourander et al. 2016; Wetterborg et al. 2019), it was calculated from the standard error as recommended by Higgins et al. (2019). One article (Porzig-Drummond et al. 2015) did not report the pre- or post-test mean or standard deviation; thus, the change statistic (F-value) and sample size was used.

Additionally, pre- and post-test correlations were used to address change due to measurement reliability using the test-retest statistic for each measure. Five studies included more than one report of the included measures for social skills and the grand mean was used for each study. Specifically, two studies (Comer et al. 2017; Enebrink et al. 2012) included more than one of the selected measures for problem behavior (ECBI and CBCL and ECBI and SDQ, respectively). One study (Xie et al. 2013) included two outcome variables from one measure (Vanderbilt ODD and Conduct subscales). Two studies (Sanders et al. 2012, 2014) collected parent-report data from both mothers and father and reported the results separately. In these five cases, the mean effect size was combined into a grand mean, so that each study only included one effect size. Mean effect size was calculated using CMA and accounted for the test-retest performance of each measure and subscale.

Five studies (Cefai et al. 2010; Day and Sanders 2018; DuPaul et al. 2018; Nixon et al. 2003; Stormshak et al. 2019) included two treatment groups compared to one control group. Of those studies, two (DuPaul et al. 2018; Nixon et al. 2003) included a BIT treatment group and a FTF treatment group, and, due to meta-analysis inclusion criteria, only the BIT treatment group was analyzed. Three studies (Cefai et al. 2010; Day and Sanders 2018; Stormshak et al. 2019) included two different BIT treatment groups compared to one control group (e.g., Fully Automated BIT versus FTF with BIT as adjuncts versus WLC). In order to include both BIT treatment groups to compare the difference between BIT subgroups, the studies were analyzed independently, and the control group sample size was divided into equal segments for each treatment group to account for the same control group appearing more than once in the data (Borenstein et al. 2009, 2015). One study (Dadds et al. 2019) reported the results of two independent studies (i.e., urban population and rural population) that were evaluated with independent, matched control groups. In this case, the groups were not pooled but were analyzed independently based on Borenstein et al. (2009) recommendations.

Several moderation analyses were performed to investigate the potential effect of study variables. In order to calculate subgroups effect size and perform moderation analysis on human support, a subgroup analysis was conducted in CMA to evaluate the between group effects of variables of interest using a random effects model, as recommended by Borenstein et al. (2009). Moderation is present when the between subjects effects is significant (Aiken

et al. 1991). Due to the number of studies that met inclusion criteria, there were not enough studies present in each category to conduct subgroup analyses on all five levels of human support (Fu et al. 2011). Therefore, traditional FTF with BIT as adjuncts and BIT with human support were combined into the subgroup "Supported BIT". This process was repeated for the 19 observations that compared BIT BPT to waitlist or education controls and for the nine observations that compared BIT BPT to active controls. In the 19 observations waitlist control study analysis, similar to the 28 observations analysis, FTF with BIT as adjuncts and BIT with human support were combined together into the subgroup Supported BIT due to a paucity of studies that met criteria for FTF with BIT as adjuncts. Additionally, fully automated active BIT and fully automated passive BIT were combined into the subgroup "Fully Automated BIT" because there was not a sufficient number of fully automated passive BIT studies to analyze on the subgroup level. In the nine observation active control study analysis, fully automated active BIT, fully automated passive BIT, FTF with BIT as adjunct, and BIT with human support were combined into the subgroup "BIT" due to an insufficient number of studies; the teletherapy articles were combined into a separate subgroup. For both the waitlist control and active control studies, subgroup analysis was performed on CMA to evaluate the between group effects of subgroups using the same process as previously described for the all studies analysis. Additionally, subgroup analysis was also conducted on CMA to determine if study comparison group (i.e., waitlist control or active comparison) had a moderating effect on BIT BPT.

The other moderation analyses were performed utilizing SPSS (IBM Corporation, Armonk, NY, USA, 2019). One-way ANOVAs were performed for each moderation analysis, and moderation was considered present when the between subjects effects were significant (Aiken et al. 1991). For moderation analyses comparing the effect size of each study, the corrected inverse-variance weighted effect size (Hedges's g) calculated on CMA was used. Moderation analysis were conducted for: levels of human support (Independent Variable (IV) = Human support; Dependent Variable (DV) = Effect Size), racial minority group membership (IV = racial minority group membership, DV = effect size), low SES group membership (IV = low SES group membership, DV = effect size), and clinical sample (IV = clinical sample, DV = effect size). For all moderation analysis, moderation was deemed present when the between subjects effects were significant (Aiken et al. 1991). All moderation analyses were performed using the 28 observations comparing BIT to both active and waitlist controls, as well as the 19 observations that compared BIT to waitlist controls and the nine observations that compared BIT to active controls. See Supplementary Table S1 for a list of study variables included in the moderation analyses.

3. Results

A total of 24 studies were included in the analysis; see Table 1 for a full description of the studies. Of these, 20 were entered as individual studies, and four studies included multiple independent subgroups and were entered independently. This resulted in 28 observations with 3957 participants. Fully automated active BIT included nine observations and 1756 participants, and fully automated passive BIT included four observations and 453 participants. BIT with Human support included eight observations and 1308 participants, FTF with BIT as adjuncts included three observations and 194 participants, and Teletherapy included four observations with 246 participants. The interventions ranged from 2–12 sessions and were delivered in various formats, including Podcast, videos, smartphone application, video-conferencing, computer software, and websites.

Table 1. Selected Studies for Behavioral Parent Training for Youth With Externalizing Behavioral Disorders.

#	Authors (Date)	Age Range	% Male	RCT Conditions	*n*	Components	Follow-Up	Primary Outcome Measures	Effect Size (Pre-Post Intervention)	Methodological Quality Rating (Average Score)	BITs Level
1	Baker et al. (2017)	2–9	55	Triple P-Positive Parenting Program (TPOL) Brief vs. WLC	200	TPOL Brief is a five-module interactive, self-directed computer program with video modeling and downloadable resources.	9-month	ECBI, CAPES, PS, Behavior Concerns and Parent Confidence Scale, PCPTOS, PAI, PPC, DASS-2, CSQ	-	20.5	FAA-BIT
2	Breitenstein et al. (2016)	2–5	43	*ez*PARENT vs. Attention Control Condition (Health Promotion Group)	79	The *ez*PARENT program is a 12-week, six module self-administered, tablet-based application. It is an adaptation of the Chicago Parent Program.	6-month	ECBI; TCQ; PQ; PSI-SF	ECBI Problem Scale: $d = -0.18$	20.5	FAA-BIT
3	Cefai et al. (2010)	9–15	50.86	Individual Parenting Wisely CD-Rom vs. Group Parenting Wisely CD-Rom vs. WLC	125	One to three session self-administered individual intervention using the Parenting Wisely CD-Rom. The group condition completed Parenting Wisely program FTF as a group during two-sessions with clinician facilitated discussion.	3-month	PSOC, ECBI	ECBI Problem Scale: Individual $d = 0.45$ Group $d = 0.69$	15	FAA-BIT FTF with BITs as adjuncts
4	Comer et al. (2017)	3–5	82.5	Internet Parent-Child Interaction Training (iPCIT) vs. FTF PCIT	40	Video teleconferencing with a therapist who provides live coaching through a webcam and Bluetooth earpiece. On average, treatment length is 20 sessions and is titrated based on family needs.	6-month	CBCL; ECBI; K-DBDs diagnostic interview; CGI-S/I; CGAS; BTPS; CSQ-8; TAI	ECBI Problem Scale: $d = -1.15$ CBCL Externalizing: $d = -1.10$	21	Teletherapy
5	Dadds et al. (2019) Study 1: Urban Study 2: Rural	3–9 3–14	79.7 79.6	AccessEI vs. FTF BPT	133 73	AccessEI is an online intervention that includes six to 10 60- to 70-min video conferencing sessions with a clinician paired with six video modules (a total of one-hour and 14-min). FTF is an intensive PMT. In study one, FTF took place over one week (four 1.5-h sessions) with one follow-up call. In study 2, FTF took place in six-10 weekly 1-h sessions.	3-month	CPRS-R; BSI; SDQ	Conners-Oppositional: $\eta^2 = 0.579$ $\eta^2 = 0.569$	20	Teletherapy

Table 1. *Cont.*

#	Authors (Date)	Age Range	% Male	RCT Conditions	*n*	Components	Follow-Up	Primary Outcome Measures	Effect Size (Pre-Post Intervention)	Methodological Quality Rating (Average Score)	BITs Level
6	Day and Sanders (2018)	1–8	46.5	TPOL vs. Telephone Supported TPOL (TPOLe) vs. WLC	183	TPOL is an eight module parenting intervention that utilizes video, interactive activities, and downloadable resources with optional text reminders. TPOLe included up to eight practitioner support sessions in which participants were able to ask questions, the practitioner reviewed module content and participant goals, and the practitioners created adherence plans if the participant was not engaging with the program.	5-month	ECBI; PS; DASS; PTC; PPC; RQI; PAI	ECBI Problem Scale: WLC vs. TPOL $d = 0.66$ WLC vs. TPOLe $d = 0.93$ TPOL vs. TPOLe $d = 0.26$	21.5	FAA-BIT BITs with Human Support
7	DuPaul et al. (2018)	3–5	63.8	Online BPT vs. FTF BPT vs. WLC	47	The online program was a 10 session internet intervention. The first session in the BIT BPT occurred in person, and parents received an overview of the program. The rest of the intervention was delivered online, and parents received weekly calls from research assistants to check on intervention implementation and answer any questions. FTF was a 10 session therapist led manualized BPT program.	-	CERS, PSI-SF, Test of Parent Knowledge	CERS Defiant/Aggressive: $\eta_p^2 = 0.07$	19.5	BITs with Human Support
8	Enebrink et al. (2012)	3–13	57.7	Internet Parent Management Training (PMT) vs. WLC	104	The internet-PMT is a seven-session program that is delivered online with feedback from research assistants. It is based off of the Swedish BPT program *Comet* (Kling et al. 2010).	6-month	ECBI, SDQ, PPI	ECBI Problem Scale: $d = 0.72$ SDQ Conduct Problems: $d = 0.30$	21	BITs with Human Support
9	Franke et al. (2016)	3–4	-	TPOL vs. Delayed Intervention	53	TPOL is a self-directed, eight module, internet positive parenting intervention. Participants received two telephone consultations with Triple P facilitators.	6-month	EC-BEH; EC-BEH-s; CBS; SDQ; PS; PSDQ; DASS-21; PSOC; CSQ	EC-BEH Defiance/Aggression: $d = 0.45$	20.5	BITs with Human Support

Table 1. *Cont.*

#	Authors (Date)	Age Range	% Male	RCT Conditions	*n*	Components	Follow-Up	Primary Outcome Measures	Effect Size (Pre-Post Intervention)	Methodological Quality Rating (Average Score)	BITs Level
10	Ghaderi et al. (2018)	10–13	-	iComet vs. Family Check-Up (FCU)	231	iComet is a seven session parent training program delivered through a secure website. FCU is a parent training model that is catered to the parent's needs.	1-year 2-years	SDQ; DBD	SDQ Conduct Problems: $d = 0.06$	21	FAA-BIT
11	Irvine et al. (2015)	11–14	52.9	Parenting Toolkit vs. WLC	307	Parenting Toolkit is a nine module online intervention completed entirely on the computer.	-	ECBI, Parenting Scale	ECBI Problem Scale: $\eta^2 = 0.009$	19.5	FAA-BIT
12	Jones et al. (2014)	3–8	53	Standard Helping the Noncompliant Child (HNC) vs. Technology Enhanced HNC (TE-HNC)	15	TE-HNC consists of eight to 12 standard, in-person sessions and access to a phone application with video examples, reminders, surveys, and home practice.	-	ECBI; consumer satisfaction scale	ECBI Problem Scale: $d = 1.59$	18.5	FTF with BITs as Adjuncts
13	Morawska et al. (2014)	2–10	61.9	TPOL Podcast vs. WLC	139	The TPOL podcast consists of seven episodes that range from nine- to 14-minutes. These podcasts present parent training topics in a conversational manner.	6-month	ECBI, CAPES, PS, PTC	ECBI Problem Scale: $d = 0.39$	17	FAP-BIT
14	Nixon et al. (2003)	3–5	70.4	Modified Parent-Child Interaction Therapy (PCIT) vs. FTF PCIT vs. WLC	54	The modified PCIT condition included videotapes in which PCIT skills were discussed and modelled, along with five face-to-face sessions and five 30-minute telephone consultations. The standard condition included 12 one- to two-hour weekly PCIT sessions.	6-month	ECBI; CBCL; HSQ-M; PSI; PSOC; PLOC; PS; DPICS-II	CBCL Externalizing: FTF PCIT vs. modified PCIT $d = 0.01$ Modified PCIT vs. WLC $d = 0.59$	17.5	FTF with BITs as Adjuncts
15	Porzig-Drummond et al. (2015)	2–10	50	1–2-3 Parenting vs. WLC	84	1–2-3 Parenting is a self-directed, online parenting program where parents learn parenting strategies by watching two videos (totaling four hours) and receive email reminders to complete the lesson and practice.	6-month	ECBI, PSI, DASS	ECBI Problem Scale: $d = 0.70$	21	FAP-BIT

Table 1. *Cont.*

#	Authors (Date)	Age Range	% Male	RCT Conditions	*n*	Components	Follow-Up	Primary Outcome Measures	Effect Size (Pre-Post Intervention)	Methodological Quality Rating (Average Score)	BITs Level
16	Rabbitt et al. (2016)	6–13	67.5	Full Contact Webcam PMT vs. Reduced Contact PMT	60	Full contact PMT included eight 50-minute teletherapy. Reduced contact PMT included 12 weekly prerecorded web sessions and 15- to 20-minute phone calls every 2 weeks with a therapist to address questions or concerns.	-	CBCL; IAB; CGAS; RDI	CBCL $d = 0.79$	18.5	BITs with Human Support
17	Sanders et al. (2012)	2–9	67.2	TPOL vs. Internet As Usual Control Group	116	TPOL is an eight-module self-directed, interactive, internet intervention. It includes video modeling, personalized goal setting, content reviews and answer feedback, interactive exercises, downloadable worksheets and podcasts, and automated text and email prompts.	6-month	ECBI, SDQ, PS, PTC, DASS-21, PAI, PPC, CSQ	ECBI Problem Scale: $d = 0.71$ SDQ Conduct: $d = 0.58$	18.5	FAA-BIT
18	Sanders et al. (2008)	2–9	64.9	*Driving Mum and Dad Mad* with Triple P workbook and website vs. *Driving Mum and Dad Mad*	174	*Driving Mum and Dad Mad* is a six-episode show about parents with young children. The enhanced condition also included Triple P self-directed workbook, weekly emails on parenting topics, reminders, access tip sheets, videos, and the option to email for assistance.	6-month	ECBI, PS, DASS1, PAI, PPC, RQI	ECBI Problem Scale $d = 0.63$	15	FAP-BIT
19	Sanders et al. (2014)	3–8	67	TPOL vs. Self-Help Workbook	193	TPOL is an eight-module self-directed, interactive, internet intervention. It includes video modeling, personalized goal setting, content reviews and answer feedback, interactive exercises, downloadable worksheets and podcasts, and automated text and email prompts. The self-help workbook consists of the same core content as TPOL, but is delivered through a workbook divided into 10 weekly sessions with reading, activities, and homework tasks.	6-month	ECBI, DASS-21, PSQ	Mother Report ECBI Problem $d = 1.44$ Father Report ECBI Problem $d = 0.73$	14	FAA-BIT

Table 1. *Cont.*

#	Authors (Date)	Age Range	% Male	RCT Conditions	*n*	Components	Follow-Up	Primary Outcome Measures	Effect Size (Pre-Post Intervention)	Methodological Quality Rating (Average Score)	BITs Level
20	Sanders et al. (2000)	2–8	58.9	Triple P Television Series *Families* vs. WLC	56	*Families* consists of 12 20- to 30-minute episodes which feature a story regarding family issues along with Triple P guidelines and instructions. Participants also had 12 written self-help Triple P information sheets.	6-month	ECBI; PS; PSOC; DASS; PPC; AARP	ECBI Problem Scale: $p = 0.09$	17	FAP-BIT
21	Sourander et al. (2016)	4	61.9	Strongest Families Smart Website (SFSW) vs. Education Control Group	464	SFSW is a 11 session internet-assisted BPT with weekly telephone coaching. Education control included access to a website on positive parenting strategies with 45-minute weekly coaching calls.	12-month	CBCL; SDQ	CBCL Externalizing: $d = 0.34$	19	BITs with Human Support
22	Stormshak et al. (2019)	6th–7th grade students	47.9	FCU Online (FCU) vs. FCU Online Plus Coach vs. WLC	322	FCU online includes at least three online sessions and is adapted to fit participant needs and goals. In the online only version, feedback was provided online. In the online plus coach version, feedback was provided over the telephone or video-conferencing.	-	SDQ	SDQ Conduct Problems: WLC vs. web-only $d = -0.13$ WLC vs. web + coach $d = -0.102$ Web-only vs. web + coach $d = 0.020$	18	FAA-BIT BITs with Human Support
23	Wetterborg et al. (2019)	12–17	41	Parent Web vs. WLC	75	Parent Web is a six- to nine-week parenting intervention delivered through the internet with five core modules and six optional modules. Each module has text, illustrations and movie clips. A practitioner provides reminders, feedback, and answers questions.	6-month 9-month	DBD, SDQ, APQ, PSS, HADS	SDQ Conduct: $d = 0.34$	19.5	BITs with Human Support

Table 1. *Cont.*

#	Authors (Date)	Age Range	% Male	RCT Conditions	*n*	Components	Follow-Up	Primary Outcome Measures	Effect Size (Pre-Post Intervention)	Methodological Quality Rating (Average Score)	BITs Level
24	Xie et al. (2013)	6–14	68.2	Videoconference BPT vs. FTF BPT	22	Both groups received 10 weekly sessions manualized parent training; however, the videoconference group never met FTF.	-	Vanderbilt Assessment Scales, SSRS, PRQ-CA; CGAS; CGI-s; CGI-I	Vanderbilt Conduct $p = 0.33$ Vanderbilt ODD $p = 0.66$	18.5	Teletherapy

Abbreviated Acceptability Rating Profile (AARP); Alabama Parenting Questionnaire (APQ); Barriers to Treatment Participation Scale (BTPS); Brief Symptom Inventory (BSI); Child Adjustment and Parent Efficacy Scale (CAPES); Child Behavior Checklist (CBCL); Child Behavior Scale (CBS); Children's Global Assessment Scale (CGAS); Client Satisfaction Questionnaire (CSQ); Clinical Global Impression-Severity and Improvement Scales (CGI-S/I); Conners Early Childhood Behavior Scale (Conners EC-BEH); Conners Early Childhood Behavior Scale-Short Form (Conners EC-BEH-s); Conners Early Childhood Rating Scale (CERS); Conners' Parent Rating Scale—Revised (CPRS–R); Depression Anxiety Stress Scales (DASS); Depression Anxiety Stress Scales-21 (DASS-21); Disruptive Behavior Disorders Rating Scale (DBD); Dyadic Parent-Interaction Coding Systems-II (DPICS-II); Eyberg Child Behaviour Inventory (ECBI); Home Situations Questionnaire–Modified (HSQ-M); Hospital Anxiety and Depression Scale (HADS); Kiddie-Disruptive Behavior Disorders Schedule (K-DBDS); Modified Parents' Consumer Satisfaction Questionnaire (PCSQ); Parent-Child Play Task Observation System (PCPTOS); Parent-Child Relationship Inventory (PCRI); Parent Child Relationship Questionnaire for Child and Adolescents (PRQ-CA); Parent Locus of Control Scale (PLOC); Parent Problem Checklist (PPC); Parental Anger Inventory (PAI); Parenting Questionnaire (PQ); Parenting Scale (PS); Parenting Sense of Competence (PSOC); Parenting Stress Index (PSI); Parenting Stress Index-Short Form (PSI-SF); Parenting Styles and Dimensions Questionnaire (PSDQ); Parenting Tasks Checklist; Perceived Stress Scale (PSS); Relationship Quality Inventory (RQI); Strengths and Difficulties Questionnaire (SDQ); Social Skills Rating System (SSRS); Test of Parenting Competence (TOPC); Therapy Attitude Inventory (TAI); Toddler Care Questionnaire (TCQ).

3.1. Methodological Quality

Based on SCRIBE (Tate et al. 2016) study guidelines, the 24 included articles were scored by two independent raters (KB and ES). Study scores ranged from 14 to 23, for an average rating of 18.84 out of a possible total of 26. Interrater reliability was evaluated utilizing Statistical Package for the Social Sciences (SPSS) version 26.0 (IBM Corporation 2019). A two-way mixed intraclass correlation was conducted using the recommendations of Hallgren (2012). Results of these analyses indicated acceptable agreement (ICC = 0.74).

Figure 2 provides a forest plot of the effect sizes of each study, as well as the grand mean effect sizes. The overall model had significant heterogeneity ($\chi^2(27) = 188.81$, $p < 0.01$), indicating significant variability across study effect sizes. Additionally, all BIT subgroups had significant heterogeneity: fully automated active BIT (χ^2 (8) = 77.24, $p < 0.01$), fully automated passive BIT (χ^2 (3) = 8.38, $p < 0.04$), supported BIT (χ^2 (10) = 68.28, $p < 0.01$), and teletherapy (χ^2 (10) = 10.72, $p < 0.01$). Due to the significant heterogeneity, a random effects model was utilized to account for the variation in true effect size between the included studies. The inconsistency estimate (I^2) was calculated to examine how the heterogeneity between studies contributed to the inconsistency in effect estimates (Borenstein et al. 2009; Card 2012). The inconsistency estimate was high for the overall model, fully automated active BIT, fully automated passive BIT, supported BIT, and teletherapy ($I^2 = 85.7$, $I^2 = 89.64$, $I^2 = 64.18$, $I^2 = 85.36$, $I^2 = 72.01$; Higgins et al. 2003).

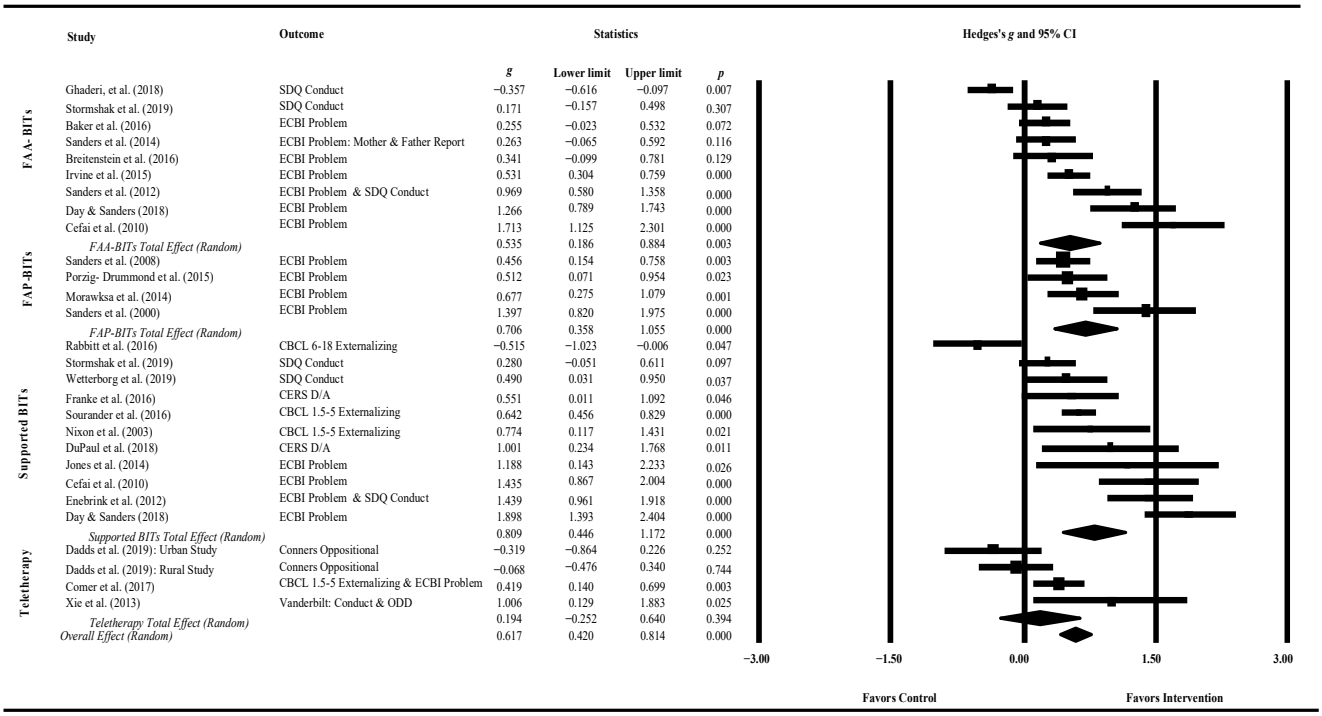

Figure 2. Forest plot: All studies (active and waitlist control).

Risk of bias was analyzed by evaluating funnel plot symmetry and calculating Rosenthal's Fail-Safe N (Rosenthal 1979). The funnel plot in Figure 3 displays the studies' effect estimates (Hedges's g; *x*-axis) against their precision, as measured by their standard errors (*y*-axis). Studies that are more precise with smaller standard errors cluster around the central line, which represents the overall effect size of the studies (Hedges's g; vertical line). Studies without bias and between study heterogeneity form a symmetrical funnel with studies clustered around the center line. After a collaborative review, the funnel symmetry was determined to be asymmetric (Borenstein 2005; Sterne et al. 2011). For all studies (i.e., BIT compared to active and waitlist controls), 14 of the studies fell within the funnel plot and, thus, clustered around the overall effect size, but 14 studies pulled away

from the mean effect size of the funnel plot. Eight studies with relatively low effect sizes pulled away to the left, and six studies with relative high effect sizes pulled away to the right. Egger's regression (Egger et al. 1997; t (26) = 1.92, $p < 0.05$) confirmed asymmetry. These results suggest that the effect may be biased. The studies that fell outside of the funnel plot were evaluated for commonalities that may have introduced systematic error (e.g., level of human support, participant characteristics) and explained the resultant bias. Five of the studies (Dadds et al. 2019; Ghaderi et al. 2018; Rabbitt et al. 2016; Sanders et al. 2014) that fell outside of the funnel plot pulling towards the left were studies that compared BIT to active controls. Studies that compared BIT to active controls may have introduced systematic error. Additionally, Rosenthal's Fail-Safe N was calculated in order to estimate risk of bias. The Fail-Safe N calculation resulted in 1492 "file drawer" studies, i.e., 1492 studies with an effect size of zero would need to be added to the analysis to render the overall effect insignificant (Borenstein et al. 2009).

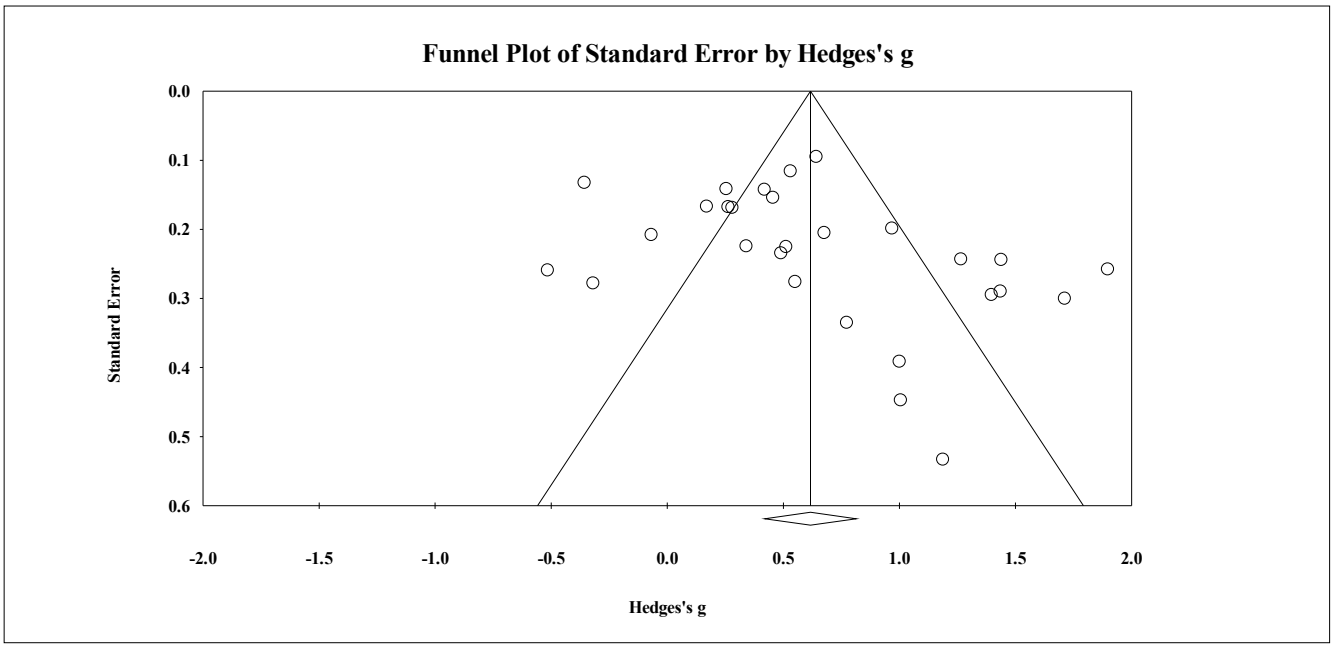

Figure 3. Funnel plot: All studies.

3.2. Primary Comparison

3.2.1. All Studies

CMA (Borenstein et al. 2014) was used to perform a pre-post control design (Morris 2008) to calculate overall and subgroup effect size. The overall effect size for all studies was in the medium range (g = 0.62, 95% CI (0.42, 0.81)) under the random effect model (Cohen 1998). Subgroup effect sizes were compared using a mixed effects analysis. Both the fully automated active BIT and fully automated passive BIT subgroups indicated a medium effect size (g = 0.54, 95% CI (0.19, 0.88) and g = 0.71, 95% CI (0.36, 1.06), respectively), and the Supported BIT subgroup demonstrated a large effect size (g = 0.81, 95% CI (0.45, 1.17)). This indicates that both fully automated BIT BPT (active and passive) and supported BIT BPT significantly improve child externalizing behaviors when compared to active (i.e., another BIT or FTF treatment) or waitlist control. Teletherapy did not indicate a significant effect size (g = 0.19, 95% CI (−0.25, 0.64)). Of note, all four studies in the teletherapy subgroup compared teletherapy to an active control (i.e., another BIT BPT or FTF BPT), and two of the four study results favored the active control. The overall between group heterogeneity in the mixed effects analysis did not indicate a significant difference between fully automated active BIT, fully automated passive BIT, supported BIT,

nor teletherapy (χ^2 (3) = 4.94, p = 0.18). The primary comparison results for all studies are displayed in Table 2.

In order to compare all five levels of human support (i.e., fully automated active BIT, fully automated passive BIT, BIT with human support, FTF treatment with BIT as adjunct, and teletherapy), a one-way ANOVA was performed on SPSS (IBM Corporation 2019). There was no significant difference between effect size and BIT levels of human support (F(4,23) = 0.947, p = 0.455).

Two additional moderation analyses were conducted using one-way ANOVAs on SPSS (IBM Corporation 2019). The between group differences were examined to compare studies where the majority of participants were members of a racial minority group or low SES. Of note, there were only three studies where the majority of participants were members of a racial minority group (racial minority studies; Breitenstein et al. 2016; Jones et al. 2014; Irvine et al. 2015). Similarly, there were only three studies where the majority of participants were low SES (low SES studies; Breitenstein et al. 2016; Jones et al. 2014; Stormshak et al. 2019). Therefore, the moderation analysis should be interpreted with caution. The between group differences did not indicate a significant difference between the results of racial minority studies and non-racial minority studies (F(1, 26) = 0.007, p = 0.933). Additionally, results indicated that there was no significant difference between low SES studies and non-low SES studies (F(1, 26) = 0.031, p = 0.861). Results of a one-way ANOVA comparing clinical and nonclinical studies did not reveal a significant difference between the groups (F(1, 26) = 1.130, p = 0.298).

Table 2. Effect size outcomes for BITs subgroups: all studies.

Problem Behavior Outcomes	n_{obs}	g	95% Confidence Interval		Z	Test of Homogeneity
			Lower	Upper		
Overall Effect	28	0.62	0.42	0.81	6.14 **	χ^2 (27) = 188.81 **
Subgroup: Treatment Type						
Fully Automated Active BIT	9	0.54	0.19	0.88	3.01 **	χ^2 (8) = 77.23 **
Fully Automated Passive BIT	4	0.71	0.36	1.06	3.97 **	χ^2 (3) = 8.38 **
Supported BITs	11	0.81	0.45	1.17	4.37 **	χ^2 (10) = 68.28 **
Teletherapy	4	0.19	−0.25	0.64	0.85	χ^2 (3) = 10.72
					Subgroup χ^2 (3) = 4.94, p = 0.18 (ns)	

** $p < 0.01$; g: Hedges's g.

3.2.2. Waitlist Control Studies

The overall effect size for all waitlist control studies was in the large range (g = 0.82, 95% CI (0.61, 1.03)) under the random effect model (Cohen 1998). The fully automated BIT subgroup indicated a medium effect size (g = 0.73, 95% CI (0.46, 1.01)), and the human supported BIT subgroup demonstrated a large effect size (g = 0.92, 95% CI (0.58, 1.27)) when compared using a mixed effects analysis. This indicates that both fully automated BIT BPT and supported BIT BPT significantly improve child externalizing behaviors when compared to waitlist control. The overall between group heterogeneity in the mixed effects analysis did not indicate a significant difference between fully automated BIT and supported BIT (χ^2 (1) = 0.72, p = 0.40). Primary comparison results for waitlist or education control studies are displayed in Table 3.

Consistent with the moderation analyses conducted for all studies, there were no significant differences between levels of human support for BIT BPT effect size. Further, racial minority, low SES, and clinical study status did not affect the significance of difference (average Hedges's g) between BIT BPT and waitlist control ($p > 0.05$). Please see Supplementary Table S2 for the results of these moderation analyses.

Table 3. Effect size outcomes for waitlist versus active control studies.

Problem Behavior Outcomes	n_{obs}	g	95% Confidence Interval		Z	Test of Homogeneity
			Lower	Upper		
Overall Effect	28	0.62	0.42	0.81	6.14 **	χ^2 (27) = 188.81 **
Subgroup: Control Group Type						
Active Control	9	0.14	−0.17	0.45	0.90	χ^2 (8) = 40.00 **
Waitlist or Education Control	19	0.82	0.61	1.03	7.68 **	χ^2 (18) = 94.80 **
						Subgroup χ^2 (1) = 12.90, p = 0.000 (s)

** $p < 0.01$; g: Hedges's g.

3.2.3. Active Control Studies

The overall effect size for active control studies was not significant (g = 0.14, 95% CI (−0.17, 0.45)) under the random effect model (Cohen 1998). This suggests that BIT BPT does not significantly improve child externalizing behavior when compared to an active treatment control group (i.e., FTF BPT or another BIT BPT). When compared using a mixed effects analysis, both the BIT (i.e., combined fully automated active BIT, fully automated passive BIT, FTF with BIT as adjuncts, and BIT with human support) and teletherapy subgroups did not indicate a significant effect size (g = 0.11, 95% CI (−0.35, 0.57) and g = 0.19, 95% CI (−0.25, 0.64), respectively). The overall between group heterogeneity in the mixed effects analysis did not indicate a significant difference between BIT and teletherapy subgroups in active control studies (χ^2 (1) = 0.067 p = 0.80). Primary comparison results for active control studies are displayed in Table 3.

Consistent with the moderation analyses conducted for all studies, there were no significant differences between levels of human support for BIT BPT effect size, racial minority, low SES, and clinical study status did not affect the significance of difference (average Hedges's g) between BIT BPT and active control ($p > 0.05$). Please see Supplemental Table S2 for the results of these moderation analyses

4. Discussion

The most widely supported evidence-based treatment is BPT usually delivered in face-to-face format (Anil Chacko et al. 2017; Chorpita et al. 2011). BPT uses behavioral principles to change child behavior by teaching positive parenting strategies (e.g., one-on-one time, positive reinforcement, timeout). There are several well-established FTF BPT manualized interventions (e.g., Incredible Years, PMTO, PCIT, Triple P, HNC). Previous meta-analyses on BPT found significant small to large effect sizes on child disruptive behavior (d = 0.42 to 0.88; Comer et al. 2013; Lee et al. 2012; Leijten et al. 2013; Lundahl et al. 2006; Mingebach et al. 2018; Maughan et al. 2005; Serketich and Dumas 1996). Recently, researchers have partially examined the effects of BIT BPT compared to FTF BPT. The present study examined the efficacy of BIT BPT, the effect of different levels of human support, and for which populations are BIT BPT effective.

In total, 24 studies met inclusion criteria to be included in the meta-analysis which included 28 observations with a total of 3957 participants. The combined effect size of all 24 of the studies indicated that BIT BPT is effective in reducing child externalizing behavior from pre- to post-treatment compared to control (i.e., both active and waitlist control; g = 0.62). Additionally, when BIT BPT was compared to waitlist control groups alone, there was a large combined effect size of the 16 waitlist control studies (g = 0.82). These results provide further support that BIT BPT is an effective intervention for treating externalizing behavior in children and adolescents. These results are comparable to previous FTF meta-analyses (ES = 0.42 to 0.88; Comer et al. 2013; Lee et al. 2012; Leijten et al. 2013; Lundahl et al. 2006; Mingebach et al. 2018; Maughan et al. 2005; Serketich and Dumas 1996), as well as slightly higher than previous BIT BPT meta-analyses (ES = 0.22 to 0.67; Baumel et al. 2016; Corralejo and Rodríguez 2018; Nieuwboer et al. 2013; Spencer et al. 2019; Thongseiratch et al. 2020). The results may be slightly higher than previous BIT BPT

meta-analyses, because the present analysis limited error by only including studies that specifically address child externalizing behavior as measured by the problem behavior subscales of well validated measures. Overall, these findings show that parents of children and adolescents with externalizing behaviors have a wider range of efficacious treatment options for BPT beyond traditional FTF interventions.

The overall effect size of the nine observations that compared BIT BPT to active control groups (i.e., other FTF or BIT interventions) did not indicate that BIT BPT is effective in reducing child externalizing behavior from pre- to post-intervention compared to active control ($g = 0.14$). Thus, BIT BPT interventions may not be more efficacious when compared to other FTF or BIT interventions. When comparing these more rigorous active control groups, three studies (Dadds et al. 2019; Ghaderi et al. 2018; Rabbitt et al. 2016) favored the active control group. However, of these studies, only one (Ghaderi et al. 2018) reported a significant between group difference between treatment and control. There was a significant difference between studies that compared BIT to active controls and studies that compared BIT to waitlist controls (χ^2 (1) = 12.90, $p < 0.05$). In summary, while BPT are efficacious compared to waitlist or education control, there was no significant difference with active control groups that included evidence based interventions.

When comparing the levels of human support between BITs subgroups in all 28 observations, fully automated active BIT, fully automated passive BIT and supported BIT demonstrated medium to large effect sizes ($g = 0.54$, $g = 0.71$, and $g = 0.81$, respectively). Teletherapy approached a small effect size and was the only BIT that produced an effect size that was not significant ($g = 0.19$, $p > 0.05$). It is of note that teletherapy was the only BIT subgroup in which all studies compared BIT to an active control (i.e., another BIT or FTF intervention), and two observations in that subgroup favored the active control, FTF PMT. Therefore, the overall effect of teletherapy may be lower due to the fact that the studies in this subgroup utilized more rigorous control groups, rather than the efficacy of the intervention itself. Despite teletherapy's trivial effect size, analysis of between group differences indicated that there was no significant difference between fully automated active BIT, fully automated passive BIT, supported BIT, and teletherapy (χ^2 (3) = 4.94, $p = 0.18$). In other words, there is no significant difference between BIT BPT by levels of human support. These results provide preliminary evidence that human support does not significantly impact the efficacy of BPT BIT. Furthermore, when looking at the intervention itself, there is no significant difference between passive viewing and listening technology (i.e., fully automated passive BITs), compared to interactive technology (fully automated active BITs). This is promising as passive BIT BPT, such as podcasts, TV shows, and pre-recorded sessions, are easily scalable, inexpensive once developed, and, thus, an option for universal intervention. This helps address the question proposed by Schueller et al. (2017): What level of human support is most efficient? While BITs that are supported have larger effect size than fully automated BIT, the difference between levels of human support is not significant. Fully automated BITs that do not require human support are less expensive (Bolier et al. 2014; Muñoz 2017) and have wider reach. Therefore, fully automated BIT BPT may be an efficient treatment option.

Additionally, when comparing levels of human support between BITs and waitlist control (i.e., excluding active control studies), fully automated and supported BIT indicated large effect sizes ($g = 0.73$ and $g = 0.92$, respectively) and no significant difference between groups (χ^2 (1) = 0.72, $p = 0.40$). When comparing levels of human support between BITs subgroups in the nine observations that compared BIT to active control studies (i.e., excluding waitlist control studies), neither the BITs nor the teletherapy subgroups were significant ($g = 0.11$ and $g = 0.19$, respectively). Additionally, there were no significant differences between BITs and active control interventions (χ^2 (1) = 0.067 $p = 0.80$). This provides additional evidence that human support does not significantly impact the efficacy of BIT BPT.

Moderation analyses comparing clinical and nonclinical studies did not indicate a significant difference between the two group's mean effect sizes; although the average

mean was higher for clinical studies. While this finding was consistent with previous research that clinical samples experience larger effect sizes of treatment (Baumel et al. 2016), it was not at a significant level. This preliminarily suggests that BIT BPT are efficacious for children whose disruptive behavior is at a clinical and non-clinical level and, thus, could be a universal treatment. BIT BPT reduces child and adolescent externalizing behavior for different levels of symptom severity and in different populations. The average effect size of low-SES studies and racial minority studies were lower than in non-racial minority studies; however, results did not indicate that there was a statistically significant difference. This analysis should be interpreted with caution because there were only three studies that met criteria for low SES study (Breitenstein et al. 2016; Jones et al. 2014; Stormshak et al. 2019) or racial minority study (Breitenstein et al. 2016; Jones et al. 2014; Irvine et al. 2015). While extant literature agrees that low SES and racial minority group membership are risk factors for externalizing behavior disorders, it largely does not evaluate these issues. Apart from these three articles, only one other article (Day and Sanders 2018) aimed to evaluate families with socioeconomic or demographic risk factors for externalizing behaviors, although it did not meet criteria to be included in the low SES or racial minority studies in the analyses. The remaining articles' aims were all related to general efficacy of the intervention, rather its efficacy in diverse, underserved, or at risk populations. These findings echo Corralejo and Rodríguez's (2018) conclusion that extant literature on BIT BPT is largely validated in white populations, rather than underserved populations who may benefit most from the increased accessibility that BIT BPT provides.

4.1. Limitations

There are several limitations of this study. First, there were only three articles that evaluated FTF treatment with BIT as adjuncts, three low SES studies, and three racial minority studies, which did not allow for meaningful analysis of those subgroups within CMA. Additionally, the number of studies within the BIT levels of human support subgroups were not equal, ranging from four to 11 studies in each subgroup. Furthermore, of the 24 studies, seven evaluated the same intervention (TPOL), which may impact the generalizability of the meta-analysis results. Additionally, although most articles included follow-up data (those that did are reported in Table 1, follow-up data was inconsistent and did not meet the threshold for meaningful analysis. Therefore, there is not any analysis to support the long term effect of BIT BPT.

Lastly, the articles included had high rates of attrition and measured treatment outcome through parent report, thus, is not blind to condition. This may impact study quality; however, these limitations are consistent throughout BPT studies. Despite these potential threats to study quality, the average methodological rating for the included studies was relatively high (18.84 out of 26), indicating that bias is minimal within the individual studies. However, the funnel plot analysis and Egger's regression indicated bias at the study level. The funnel plot analyses revealed asymmetry, and five studies with relatively low standard errors pulled to the left, and four studies with high effect sizes pulled to the right. This suggests that the overall effect may be biased, and additional studies with similar sample sizes should be performed to gain more precise overall effect estimates.

4.2. Future Directions

Future rigorous, evidence based research should be conducted to better understand the impact of BIT BPT. First, studies should evaluate the long term effects of BIT BPT to increase understanding of the impact and generalizability for BIT BPT. Additionally, future studies should directly compare BIT BPT to their FTF BPT counterparts to further understand the efficiency of BIT BPT. Network meta-analyses that evaluate multiple treatments though direct comparisons between interventions and also across individual study comparison groups would allow deeper understanding of various BIT BPT and their efficacy relative to evidence-based (i.e., FTF, other BIT) and non-evidenced-based comparisons (i.e., WLC, education control). Furthermore, while extant literature suggests

that underserved populations may benefit more from efficacious BIT BPT, there is a dearth of research that evaluates that efficacy in underserved populations. Additional BIT BPT RCTs with more diverse participants are needed. Lastly, BIT BPT treatments are still emerging, and are growing ever more important due to the reduced FTF contact associated with COVID-19 and increase in telemental health services. Thus, it is vital that future research analyzes the efficacy of BIT BPT, as well as which technology platform is most efficient (e.g., smartphone applications, computer software, videos, etc.). As the trend in psychotherapy moves towards less FTF contact and more telemental health, future research should analyze the specific factors that make BIT BPT most efficacious.

5. Conclusions

Overall, BIT BPT significantly decreases externalizing behavior in children and adolescents. All levels of human support, except for teletherapy, indicate a moderate to large effect size, and there were no significant differences between levels of human support. Additionally, analyses did not indicate significant differences between human support, low SES studies, racial minority studies, nor studies with clinical levels of child externalizing behavior. This provides preliminary support that BIT BPT are efficacious in multiple modalities and across multiple populations. While results indicate that BIT BPT may be effective for individuals in low SES or racial minority groups, there need to be rigorous empirical evaluations that address treatment for externalizing behavior disorders in underserved populations and in countries with different cultures.

This is promising because BIT BPT provides additional treatment options for parents who are unable to participate in FTF BPT due to lack of access to a mental health provider, finances, disabilities, or illness. While FTF BPT may not be accessible to every parent, many parents have access to a smartphone, tablet, or computer and, thus, are able to engage in BIT BPT. BIT BPT increases the portfolio of treatment options for child and adolescent externalizing behavior, as well as can reduce the significant personal and financial costs associated with untreated externalizing behavior disorders. BIT BPT offers a potential solution to bridge the gap between efficacy and effectiveness.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/socsci10100367/s1, Table S1: Study Variable. Table S2: Comparisons by Treatment Variables: Waitlist and Active Control Studies.

Author Contributions: Conceptualization, K.B.B. and E.L.B.; methodology, K.B.B. and E.L.B.; software, K.B.B.; validation, K.B.B.; formal analysis, K.B.B.; investigation, K.B.B. and E.L.B.; resources, K.B.B.; data curation, K.B.B.; writing—original draft preparation, K.B.B. and E.L.B.; writing—review and editing, K.B.B. and E.L.B.; visualization, K.B.B. and E.L.B.; supervision, E.L.B.; project administration, K.B.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available upon request.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Methodological Quality Rating

Table A1. The Single-Case Reporting Guideline in BEhavioural Interventions (SCRIBE) 2016 Checklist.

Item Number	Topic	Item Description
TITLE and ABSTRACT		
1	Title	Identify the research as a single-case experimental design in the title
2	Abstract	Summarize the research question, population, design, methods including intervention/s (independent variable/s) and target behavior/s and any other outcome/s (dependent variable/s), results, and conclusions
INTRODUCTION		
3	Scientific background	Describe the scientific background to identify issue/s under analysis, current scientific knowledge, and gaps in that knowledge base
4	Aims	State the purpose/aims of the study, research question/s, and, if applicable, hypotheses
METHOD		
5	Design	Identify the design (e.g., withdrawal/reversal, multiple-baseline, alternating-treatments, changing-criterion, some combination thereof, or adaptive design) and describe the phases and phase sequence (whether determined *a priori* or data-driven) and, if applicable, criteria for phase change
6	Procedural Changes	Describe any procedural changes that occurred during the course of the investigation after the start of the study
7	Replication	Describe any planned replication
8	Randomization	State whether randomization was used, and if so, describe the randomization method and the elements of the study that were randomized
9	Blinding	State whether blinding/masking was used, and if so, describe who was blinded/masked
PARTICIPANT/S or UNITS		
10	Selection criteria	State the inclusion and exclusion criteria, if applicable, and the method of recruitment
11	Participant characteristics	For each participant, describe the demographic characteristics and clinical (or other) features relevant to the research question, such that anonymity is ensured
CONTEXT		
12	Setting	Describe characteristics of the setting and location where the study was conducted
APPROVALS		
13	Ethics	State whether ethics approval was obtained and indicate if and how informed consent and/or assent were obtained
MEASURES and MATERIALS		
14	Measures	Operationally define all target behaviors and outcome measures, describe reliability and validity, state how they were selected, and how and when they were measured
15	Equipment	Clearly describe any equipment and/or materials (e.g., technological aids, biofeedback, computer programs, intervention manuals or other material resources) used to measure target behavior/s and other outcome/s or deliver the interventions
INTERVENTIONS		
16	Intervention	Describe the intervention and control condition in each phase, including how and when they were actually administered, with as much detail as possible to facilitate attempts at replication
17	Procedural fidelity	Describe how procedural fidelity was evaluated in each phase
ANALYSIS		
18	Analyses	Describe and justify all methods used to analyze data
RESULTS		

Table A1. *Cont.*

Item Number	Topic	Item Description
19	Sequence completed	For each participant, report the sequence actually completed, including the number of trials for each session for each case. For participant/s who did not complete, state when they stopped and the reasons
20	Outcomes and estimation	For each participant, report results, including raw data, for each target behavior and other outcome/s
21	Adverse events	State whether or not any adverse events occurred for any participant and the phase in which they occurred
DISCUSSION		
22	Interpretation	Summarize findings and interpret the results in the context of current evidence
23	Limitations	Discuss limitations, addressing sources of potential bias and imprecision
24	Applicability	Discuss applicability and implications of the study findings
DOCUMENTATION		
25	Protocol	If available, state where a study protocol can be accessed
26	Funding	Identify source/s of funding and other support; describe the role of funders

References

Able, Stephen L., Joseph A. Johnston, Lenard A. Adler, and Ralph W. Swindle. 2007. Functional and psychosocial impairment in adults with undiagnosed ADHD. *Psychological Medicine* 37: 97–107. [CrossRef]

Achenbach, Thomas M. 1991. *Child Behavior Checklist/4–18*. Burlington: University of Vermont.

Achenbach, Thomas M., and Leslie A. Rescorla. 2000. *Manual for the ASEBA Preschool Forms and Profiles*. Burlington: Department of Psychiatry, University of Vermont.

Aiken, Leona S., Stephen G. West, and Raymond R. Reno. 1991. *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks: Sage.

Andersson, Gerhard, and Pim Cuijpers. 2009. Internet-Based and Other Computerized Psychological Treatments for Adult Depression: A Meta-Analysis. *Cognitive Behaviour Therapy* 38: 196–205. [CrossRef]

Baker, Sabine, Matthew R. Sanders, Karen M. T. Turner, and Alina Morawska. 2017. A randomized controlled trial evaluating a low-intensity interactive online parenting intervention, Triple P Online Brief, with parents of children with early onset conduct problems. *Behaviour Research and Therapy* 91: 78–90. [CrossRef]

Bard, David E., Mark L. Wolraich, Barbara Neas, Melissa Doffing, and Laoma Beck. 2013. The psychometric properties of the Vanderbilt Attention-Deficit Hyperactivity Disorder Diagnostic Parent Rating Scale in a community population. *Journal of Developmental and Behavioral Pediatrics* 34: 72–82. [CrossRef]

Baumeister, Harald, Lars Reichler, Marie Munzinger, and Jiaxi Lin. 2014. The impact of guidance on Internet-based mental health interventions: A systematic review. *Internet Interventions* 1: 205–15. [CrossRef]

Baumel, Amit, Aditya Pawar, John M. Kane, and Christoph U. Correll. 2016. Digital parent training for children with disruptive behaviors: Systematic review and meta-analysis of randomized trials. *Journal of Child and Adolescent Psychopharmacology* 26: 740–49. [CrossRef] [PubMed]

Bolier, Linda, Cristina Majo, Filip Smit, Gerben J. Westerhof, Merel Haverman, Jan A. Walburg, Heleen Riper, and Ernst Bohlmeijer. 2014. Cost-effectiveness of online positive psychology: Randomized controlled trial. *The Journal of Positive Psychology* 9: 460–71. [CrossRef]

Borenstein, Michael. 2005. Software for publication bias. In *Publication Bias in Meta-Analysis: Prevention, Assessment, and Adjustments*. Edited by Hannah R. Rothstein, Alexander J. Sutton and Michael Borenstein. Chichester: John Wiley & Sons, pp. 193–220.

Borenstein, Michael, Harris Cooper, Larry V. Hedges, and Jeffrey C. Valentine. 2009. Effect sizes for continuous data. In *The Handbook of Research Synthesis and Meta-Analysis*. New York: Russell Sage Foundation, vol. 2, pp. 221–35.

Borenstein, Michael, Larry Hedges, Julian P. T. Higgins, and Hannah Rothstein. 2014. *Comprehensive Meta-Analysis. Computer Program*. Version 3. Englewood: Biostat.

Borenstein, Michael, Larry V. Hedges, Julian P. T. Higgins, and Hannah Rothstein. 2015. Regression in meta-analysis. In *Comprehensive Meta Analysis Manual*. Englewood: Biostat Inc.

Breitenstein, Susan M., Louis Fogg, Edith V. Ocampo, Diana I. Acosta, and Deborah Gross. 2016. Parent use and efficacy of a self-administered, tablet-based parent training intervention: A randomized controlled trial. *JMIR mHealth and uHealth* 4: e36. [CrossRef]

Card, Noel A. 2012. *Applied Meta-Analysis for Social Science Research*. New York: The Guilford Press.

Cefai, Josie, David Smith, and Robert E. Pushak. 2010. Parenting wisely: Parent training via CD-ROM with an Australian sample. *Child and Family Behavior Therapy* 32: 17–33. [CrossRef]

Chacko, Anil, Scott A. Jensen, Lynda S. Lowry, Melinda Cornwell, Alyssa Chimklis, Elizabeth Chan, Daniel Lee, and Brenda Pulgarin. 2016. Engagement in behavioral parent training: Review of the literature and implications for practice. *Clinical Child and Family Psychology Review* 19: 204–15. [CrossRef]

Anil Chacko, Carla Counts Allan, Simone S. Moody, Trista Perez Crawford, Cy Nadler, and Alyssa Chimiklis. 2017. Behavioral interventions. In *Handbook of DSM–5 Disorders in Children and Adolescents*. Edited by Sam Goldstein and Melissa DeVries. Cham: Springer, pp. 617–36.

Chorpita, Bruce F., Eric L. Daleiden, Chad Ebesutani, John Young, Kimberly D. Becker, Brad J. Nakamura, Lisa Phillips, Alyssa Ward, Roxanna Lynch, Lindsay Trent, and et al. 2011. Evidence-based treatments for children and adolescents: An updated review of indicators of efficacy and effectiveness. *Clinical Psychology: Science and Practice* 18: 154–72. [CrossRef]

Cohen, Mark A. 1998. The monetary value of saving a high-risk youth. *Journal of Quantitative Criminology* 14: 5–33. [CrossRef]

Comer, Jonathan S., and Kathleen Myers. 2016. Future directions in the use of telemental health to improve the accessibility and quality of children's mental health services. *Journal of Child and Adolescent Psychopharmacology* 26: 296–300. [CrossRef] [PubMed]

Comer, Jonathan S., Candice Chow, Priscilla T. Chan, Christine Cooper-Vince, and Lianna A. S. Wilson. 2013. Psychosocial treatment efficacy for disruptive behavior problems in very young children: A meta-analytic examination. *Journal of the American Academy of Child and Adolescent Psychiatry* 52: 26–36. [CrossRef] [PubMed]

Comer, Jonathan S., Jami M. Furr, Elizabeth M. Miguel, Christine E. Cooper-Vince, Aubrey L. Carpenter, R. Meredith Elkins, Caroline E. Kerns, Danielle Cornacchio, Tommy Chou, Stefany Coxe, and et al. 2017. Remotely delivering real-time parent training to the home: An initial randomized trial of Internet-delivered parent–child interaction therapy (I-PCIT). *Journal of Consulting and Clinical Psychology* 85: 909. [CrossRef] [PubMed]

Conners, Keith. 2009. *Conners Early Childhood*. New York: Multi-Health Systems.

Conners, Keith. 2018. *Conners Early Childhood*. Edited by Conners E. C. Brochure. New York: Multi-Health Systems.

Conners, C. Keith, Gill Sitarenios, James D. A. Parker, and Jeffery N. Epstein. 1998. The revised Conners' Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology* 26: 257–68. [CrossRef]

Corralejo, Samantha M., and Melanie M. Domenech Rodríguez. 2018. Technology in parenting programs: A systematic review of existing interventions. *Journal of Child and Family Studies* 27: 2717–31. [CrossRef]

Dadds, Mark R., Christina Thai, Antonio Mendoza Diaz, Joshua Broderick, Caroline Moul, Lucy A. Tully, David J. Hawes, Suzanne Davies, Katherine Burchfield, and Lindsay Cane. 2019. Therapist-assisted online treatment for child conduct problems in rural and urban families: Two randomized controlled trials. *Journal of Consulting and Clinical Psychology* 87: 706–19. [CrossRef]

Day, Jamin J., and Matthew R. Sanders. 2017. Mediators of parenting change within a web-based parenting program: Evidence from a randomized controlled trial of Triple P Online. *Couple and Family Psychology: Research and Practice* 6: 154. [CrossRef]

Day, Jamin J., and Matthew R. Sanders. 2018. Do parents benefit from help when completing a self-guided parenting program online? A randomized controlled trial comparing Triple P Online with and without telephone support. *Behavior Therapy* 49: 1020–38. [CrossRef]

Desatnik, Alex, Charlotte Jarvis, Nisha Hickin, Lara Taylor, David Trevatt, Pia Tohme, and Nicolas Lorenzini. 2021. Preliminary Real-World Evaluation of an Intervention for Parents of Adolescents: The Open Door Approach to Parenting Teenagers (APT). *Journal of Child and Family Studies* 30: 38–50. [CrossRef]

DuPaul, George J., Lee Kern, Georgia Belk, Beth Custer, Molly Daffner, Andrea Hatfield, and Daniel Peek. 2018. Face-to-face versus online behavioral parent training for young children at risk for ADHD: Treatment engagement and outcomes. *Journal of Clinical Child and Adolescent Psychology* 47: 369–83. [CrossRef]

Egger, Helen Link, and Adrian Angold. 2006. Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry* 47: 313–37. [CrossRef] [PubMed]

Egger, Matthias, George Davey Smith, Martin Schneider, and Christoph Minder. 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315: 629–34. [CrossRef] [PubMed]

Enebrink, Pia, Jens Högström, Martin Forster, and Ata Ghaderi. 2012. Internet-based parent management training: A randomized controlled study. *Behaviour Research and Therapy* 50: 240–49. [CrossRef]

Eyberg, Sheila M., and Arthur W. Ross. 1999. ECBI and SESBI-R: Eyberg child behavior inventory and Sutter-Eyberg student behavior inventory-revised: Professional manual. *Psychological Assessment Resources*. Available online: https://www.parinc.com/products/pkey/97 (accessed on 21 August 2021).

Eyberg, Sheila M., and Arthur W. Ross. 1978. Assessment of child behavior problems: The validation of a new inventory. *Journal of Clinical Child and Adolescent Psychology* 7: 113–16. [CrossRef]

Eyberg, Sheila M., Stephen R. Boggs, and James Algina. 1995. Parent-child interaction therapy: A psychosocial model for the treatment of young children with conduct problem behavior and their families. *Psychopharmacology Bulletin* 31: 83–91.

Farmer, Elizabeth M. Z., Scott N. Compton, J. Barbara Burns, and Elizabeth Robertson. 2002. Review of the evidence base for treatment of childhood psychopathology: Externalizing disorders. *Journal of Consulting and Clinical Psychology* 70: 1267–302. [CrossRef]

Forehand, Rex Lloyd, and Robert Joseph McMahon. 1981. *Helping the Noncompliant Child: A Clinician's Guide to Parent Training*. New York: Guilford Press.

Forgatch, Marion S., and Gerald R. Patterson. 2010. Parent Management Training—Oregon Model: An intervention for antisocial behavior in children and adolescents. In *Evidence-Based Psychotherapies for Children and Adolescents*, 2nd ed. Edited by John R. Weisz and Alan E. Kazdin. New York: Guilford, pp. 159–78.

Foster, E. Michael, Damon E. Jones, and Conduct Problems Prevention Research Group. 2005. The high costs of aggression: Public expenditures resulting from conduct disorder. *American Journal of Public Health* 95: 1767–72. [CrossRef]
Franke, Nike, Louise J. Keown, and Matthew R. Sanders. 2016. An RCT of an online parenting program for parents of preschool-aged children with ADHD symptoms. *Journal of Attention Disorders* 24: 1716–26. [CrossRef] [PubMed]
Fu, Rongwei, Gerald Gartlehner, Mark Grant, Tatyana Shamliyan, Art Sedrakyan, Timothy J. Wilt, Lauren Griffith, Mark Oremus, Parminder Raina, Afisi Ismaila, and et al. 2011. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology* 64: 1187–97. [CrossRef] [PubMed]
Fuentes, María C., Antonio Alarcón, Fernando García, and Enrique Gracia. 2015. Use of alcohol, tobacco, cannabis and other drugs in adolescence: Effects of family and neighborhood. *Anales de Psicologia* 31: 1000–7. [CrossRef]
Fuentes, María C., Oscar F. Garcia, and Fernando Garcia. 2020. Protective and risk factors for adolescent substance use in Spain: Self-esteem and other indicators of personal well-being and ill-being. *Sustainability* 12: 5962. [CrossRef]
Garcia, Oscar F., Maria C. Fuentes, Enrique Gracia, Emilia Serra, and Fernando Garcia. 2020. Parenting warmth and strictness across three generations: Parenting styles and psychosocial adjustment. *International Journal of Environmental Research and Public Health* 17: 7487. [CrossRef] [PubMed]
Ghaderi, Ata, Christina Kadesjö, Annika Björnsdotter, and Pia Enebrink. 2018. Randomized effectiveness trial of the family check-up versus internet-delivered parent training (iComet) for families of children with conduct problems. *Scientific Reports* 8: 1–15. [CrossRef] [PubMed]
Gimenez-Serrano, Sofia, Fernando Garcia, and Oscar F. Garcia. 2021. Parenting styles and its relations with personal and social adjustment beyond adolescence: Is the current evidence enough? *European Journal of Developmental Psychology*, 1–21. [CrossRef]
Goodman, Robert. 1997. The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry* 38: 581–86. [CrossRef] [PubMed]
Goodman, Robert. 2001. Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry* 40: 1337–45. [CrossRef]
Grolnick, Wendy S., Madeline R. Levitt, Alessandra J. Caruso, and Rachel E. Lerner. 2021. Effectiveness of a Brief Preventive Parenting Intervention Based in Self-Determination Theory. *Journal of Child and Family Studies* 30: 905–20. [CrossRef]
Hallgren, Kevin A. 2012. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology* 8: 23–34. [CrossRef] [PubMed]
Hedges, Larry V., and Ingram Olkin. 1985. *Statistical Methods for Meta-Analysis*. San Diego: Academic Press.
Higgins, Julian P. T., Simon G. Thompson, Jonathan J. Deeks, and Douglas G. Altman. 2003. Measuring inconsistency in meta-analyses. *BMJ* 327: 557–60. [CrossRef]
Higgins, Julian P. T., James Thomas, Jacqueline Chandler, Miranda Cumpston, Tianjing Li, Matthew J. Page, and Vivian A. Welch, eds. 2019. Chapter 6: Choosing effect measures and computing estimates of effect. In *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.0 (Updated July 2019). London: Cochrane.
Hunter, John E., and Frank L. Schmidt. 2004. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. London: Sage Publications.
IBM Corporation. 2019. *IBM SPSS Statistics for Mac*. Version 26.0. Armonk: IBM Corporation.
Irvine, A. Blair, Vicky A. Gelatt, Michael Hammond, and John R. Seeley. 2015. A randomized study of internet parent training accessed from community technology centers. *Prevention Science* 16: 597–608. [CrossRef]
Jones, Deborah J., Rex Forehand, Jessica Cuellar, Justin Parent, Amanda Honeycutt, Olga Khavjou, Michelle Gonzalez, Margaret Anton, and Greg A. Newey. 2014. Technology-enhanced program for child disruptive behavior disorders: Development and pilot randomized control trial. *Journal of Clinical Child and Adolescent Psychology* 43: 88–101. [CrossRef]
Kazak, Anne E., Kimberly Hoagwood, John R. Weisz, Korey Hood, Thomas R. Kratochwill, Luis A. Vargas, and Gerard A. Banez. 2010. A meta-systems approach to evidence-based practice for children and adolescents. *American Psychologist* 65: 85–97. [CrossRef] [PubMed]
Kling, Åsa, Martin Forster, Knut Sundell, and Lennart Melin. 2010. A randomized controlled effectiveness trial of parent management training with varying degrees of therapist support. *Behavior Therapy* 41: 530–42. [CrossRef] [PubMed]
Lee, Pei-chin, Wern-ing Niew, Hao-jan Yang, Vincent Chin-hung Chen, and Keh-chung Lin. 2012. A meta-analysis of behavioral parent training for children with attention deficit hyperactivity disorder. *Research in Developmental Disabilities* 33: 2040–49. [CrossRef]
Leijten, Patty, Maartje A. J. Raaijmakers, Bram Orobio de Castro, and Walter Matthys. 2013. Does socioeconomic status matter? A meta-analysis on parent training effectiveness for disruptive child behavior. *Journal of Clinical Child and Adolescent Psychology* 42: 384–92. [CrossRef]
Liu, Jianghong. 2004. Childhood externalizing behavior: Theory and implications. *Journal of Child and Adolescent Psychiatric Nursing* 17: 93–103. [CrossRef]
Lobo, Michele A., Mariola Moeyaert, Andrea Baraldi Cunha, and Iryna Babik. 2017. Single-case design, analysis, and quality assessment for intervention research. *Journal of Neurologic Physical Therapy* 41: 187. [CrossRef]
Lundahl, Brad, Heather J. Risser, and M. Christine Lovejoy. 2006. A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review* 26: 86–104. [CrossRef]
Maughan, Denita R., Elizabeth Christiansen, William R. Jenson, Daniel Olympia, and Elaine Clark. 2005. Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review* 34: 267. [CrossRef]

Merikangas, Kathleen Ries, Jian-ping He, Marcy Burstein, Sonja A. Swanson, Shelli Avenevoli, Lihong Cui, Corina Benjet, Katholiki Georgiades, and Joel Swendsen. 2010. Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication—Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry* 49: 980–89. [CrossRef]

Mingebach, Tanja, Inge Kamp-Becker, Hanna Christiansen, and Linda Weber. 2018. Meta-meta-analysis on the effectiveness of parent-based interventions for the treatment of child externalizing behavior problems. *PLoS ONE* 13: e0202855. [CrossRef]

Mohr, David C., Michelle Nicole Burns, Stephen M. Schueller, Gregory Clarke, and Michael Klinkman. 2013. Behavioral intervention technologies: Evidence review and recommendations for future research in mental health. *General Hospital Psychiatry* 35: 332–338. [CrossRef]

Morawska, Alina, Helen Tometzki, and Matthew R. Sanders. 2014. An evaluation of the efficacy of a triple P-positive parenting program podcast series. *Journal of Developmental & Behavioral Pediatrics* 35: 128–137.

Morris, Scott B. 2008. Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods* 11: 364–86. [CrossRef]

Muñoz, Ricardo F. 2017. The efficiency model of support and the creation of digital apothecaries. *Clinical Psychology: Science and Practice* 24: 46–49. [CrossRef]

Musitu Ochoa, Gonzalo, Clara Maria Isabel Martinez Sanchez, Santiago Yubero, and Jose Fernando GarcÍa Pérez. 2012. Family socialization practices: Factor confirmation of the Portuguese version of a scale for their measurement. *Revista de Psicodidactica* 17: 159–178.

Mytton, Julie, Jenny Ingram, Sarah Manns, and James Thomas. 2014. Facilitators and barriers to engagement in parenting programs: A qualitative systematic review. *Health Education and Behavior* 41: 127–37. [CrossRef] [PubMed]

Nieuwboer, Christa C., Ruben G. Fukkink, and Jo M. A. Hermanns. 2013. Online programs as tools to improve parenting: A meta-analytic review. *Children and Youth Services Review* 35: 1823–29. [CrossRef]

Nixon, Reginald D. V., Lynne Sweeney, Deborah B. Erickson, and Stephen W. Touyz. 2003. Parent-child interaction therapy: A comparison of standard and abbreviated treatments for oppositional defiant preschoolers. *Journal of Consulting and Clinical Psychology* 71: 251. [CrossRef]

Nock, Matthew K., and Caitlin Ferriter. 2005. Parent management of attendance and adherence in child and adolescent therapy: A conceptual and empirical review. *Clinical Child and Family Psychology Review* 8: 149–66. [CrossRef] [PubMed]

Pelham, William E., E. Michael Foster, and Jessica A. Robb. 2007. The economic impact of attention-deficit/hyperactivity disorder in children and adolescents. *Journal of Pediatric Psychology* 32: 711–27. [CrossRef] [PubMed]

Porzig-Drummond, Renata, Richard J. Stevenson, and Caroline Stevenson. 2015. Preliminary evaluation of a self-directed video-based 1–2-3 Magic parenting program: A randomized controlled trial. *Behaviour Research and Therapy* 66: 32–42. [CrossRef]

Rabbitt, Sarah M., Erin Carrubba, Bernadette Lecza, Emily McWhinney, Jennifer Pope, and Alan E. Kazdin. 2016. Reducing therapist contact in parenting programs: Evaluation of internet-based treatments for child conduct problems. *Journal of Child and Family Studies* 25: 2001–20. [CrossRef] [PubMed]

Robinson, Elizabeth A., Sheila M. Eyberg, and A. William Ross. 1983. Conduct problem behavior: Standardization of a behavior rating scale with adolescents. *Journal of Clinical Child Psychology* 12: 347–54.

Rosenthal, Robert. 1979. The file drawer problem and tolerance for null results. *Psychological Bulletin* 86: 638. [CrossRef]

Sanders, Matthew R. 1999. Triple p-positive parenting program: Towards an empirically validated multilevel parenting and family support strategy for the prevention of behavioral and emotional problems in children. *Clinical Child and Family Psychology Review* 2: 71–90. [CrossRef]

Sanders, Matthew R., Danielle T. Montgomery, and Margaret L. Brechman-Toussaint. 2000. The mass media and the prevention of child behavior problems: The evaluation of a television series to promote positive outcomes for parents and their children. *The Journal of Child Psychology and Psychiatry and Allied Disciplines* 41: 939–48. [CrossRef]

Sanders, Matthew, Rachel Calam, Marianne Durand, Tom Liversidge, and Sue Ann Carmont. 2008. Does self-directed and web-based support for parents enhance the effects of viewing a reality television series based on the Triple P–Positive Parenting Programme? *Journal of Child Psychology and Psychiatry* 49: 924–32. [CrossRef]

Sanders, Matthew R., Sabine Baker, and Karen M.T. Turner. 2012. A randomized controlled trial evaluating the efficacy of Triple P Online with parents of children with early-onset conduct problems. *Behaviour Research and Therapy* 50: 675–84. [CrossRef]

Sanders, Matthew R., Cassandra K. Dittman, Susan P. Farruggia, and Louise J. Keown. 2014. A comparison of online versus workbook delivery of a self-help positive parenting program. *The Journal of Primary Prevention* 35: 125–33. [CrossRef]

Schueller, Stephen M., Kathryn Noth Tomasino, and David C. Mohr. 2017. Integrating human support into behavioral intervention technologies: The efficiency model of support. *Clinical Psychology: Science and Practice* 24: 27–45. [CrossRef]

Serketich, Wendy J., and Jean E. Dumas. 1996. The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy* 27: 171–86. [CrossRef]

Sourander, Andre, Patrick J. McGrath, Terja Ristkari, Charles Cunningham, Jukka Huttunen, Patricia Lingley-Pottie, Susanna Hinkka-Yli-Salomäki, Malin Kinnunen, Jenni Vuorio, Atte Sinokki, and et al. 2016. Internet-assisted parent training intervention for disruptive behavior in 4-year-old children: A randomized clinical trial. *JAMA Psychiatry* 73: 378–87. [CrossRef] [PubMed]

Spencer, Chelsea M., Glade L. Topham, and Erika L. King. 2019. Do online parenting programs create change? A meta-analysis. *Journal of Family Psychology* 34: 346. [CrossRef] [PubMed]

Steinberg, Laurence. 2007. Risk taking in adolescence: New perspectives from brain and behavioral science. *Current Directions in Psychological Science* 16: 55–59. [CrossRef]

Sterne, Jonathan A. C., Alex J. Sutton, John P. A. Ioannidis, Norma Terrin, David R. Jones, Joseph Lau, James Carpenter, Gerta Rücker, Roger M. Harbord, Christopher H. Schmid, and et al. 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 343: d4002. [CrossRef]

Stormshak, Elizabeth A., John R. Seeley, Allison S. Caruthers, Lucia Cardenas, Kevin J. Moore, Milagra S. Tyler, Christopher M. Fleming, Jeff Gau, and Brian Danaher. 2019. Evaluating the efficacy of the Family Check-Up Online: A school-based, eHealth model for the prevention of problem behavior during the middle school years. *Development and Psychopathology* 31: 1873–86. [CrossRef]

Tate, Robyn L., Michael Perdices, Ulrike Rosenkoetter, William Shadish, Sunita Vohra, David H. Barlow, Robert Horner, Alan Kazdin, Thomas Kratochwill, Skye McDonald, and et al. 2016. The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016: Explanation and elaboration. *Archives of Scientific Psychology* 4: 10. [CrossRef]

Thongseiratch, Therdpong, Patty Leijten, and G. J. Melendez-Torres. 2020. Online parent programs for children's behavioral problems: A meta-analytic review. *European Child and Adolescent* 29: 1555–68. [CrossRef] [PubMed]

Webster-Stratton, Carolyn, and M. Jamila Reid. 2003. The incredible years parents, teachers and children training series: A multifaceted treatment approach for young children with conduct problems. In *Evidence-Based Psychotherapies for Children and Adolescents*. Edited by Alan E. Kazdin and John R. Weisz. New York: The Guilford Press, pp. 224–40.

Weisenmuller, Chantel, and Dane Hilton. 2020. Barriers to Access, Implementation, and Utilization of Parenting Interventions: Considerations for Research and Clinical Applications. *American Psychologist* 76: 104. [CrossRef]

Wetterborg, Dan, Pia Enebrink, Kajsa Lönn Rhodin, Martin Forster, Ellen Risto, Johanna Dahlström, Kajsa Forsberg, and Ata Ghaderi. 2019. A pilot randomized controlled trial of Internet-delivered parent training for parents of teenagers. *Journal of Family Psychology* 33: 764–74. [CrossRef] [PubMed]

Wolraich, Mark L. 2002. *NICHQ Vanderbilt Assessment Scale*. Boston: National Institute for Children's Health Quality; American Academy of Pediatrics and National Initiative for Children's Healthcare Quality.

Xie, Yuhuan, J. Faye Dixon, Ong Min Yee, Junshun Zhang, Y. Ann Chen, Sascha DeAngelo, Peter Yellowlees, Robert Hendren, and Julie B. Schweitzer. 2013. A study on the effectiveness of videoconferencing on teaching parent training skills to parents of children with ADHD. *Telemedicine and e-Health* 19: 192–99. [CrossRef] [PubMed]

Zisser, Alison, and Sheila M. Eyberg. 2010. Parent-child interaction therapy and the treatment of disruptive behavior disorders. In *Evidence-Based Psychotherapies for Children and Adolescents*. Edited by John R. Weisz and Alan E. Kazdin. New York: Guilford Press, pp. 179–93.