View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Queensland eSpace

Accepted Manuscript

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

Jamin J. Day, Matthew R. Sanders

PII: S0005-7894(18)30033-9
DOI: doi:10.1016/j.beth.2018.03.002
Reference: BETH 793

To appear in:

Received date: 19 June 2017
Accepted date: 5 March 2018

Please cite this article as: Jamin J. Day, Matthew R. Sanders , 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. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Beth(2018), doi:10.1016/j.beth.2018.03.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

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

Jamin J. Day and Matthew R. Sanders

Parenting and Family Support Centre

School of Psychology

The University of Queensland

Corresponding author:
Dr. Jamin J. Day
Parenting and Family Support Centre
School of Psychology,
University of Queensland,
St Lucia 4072
Australia
Ph: +61 7 3365 7290
Fax: +61 7 3365 6724
Email: j.day@psy.uq.edu.au

ACCEPTED MANUSCRIPT

Author Note

Acknowledgements

We gratefully acknowledge the families that participated in this trial, Parenting WA (Department of Local Government and Communities, Western Australia) for their generous support and involvement, and the two anonymous reviewers for their insightful and constructive feedback.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Research funding support was provided through the Parenting and Family Support Centre and the University of Queensland.

ACCEPTED MANUSCRIPT

Abstract

In response to recent increases in the dissemination of web-based parenting supports, an important consideration is whether the core benefits of self-directed participation in online parenting interventions are counterbalanced by issues such as high dropout and non-completion rates commonly reported within the Internet intervention literature. This study outlines a randomized controlled trial of Triple P Online, a web-based variant of the Triple P—Positive Parenting Program, delivered with varied levels of support scaffolding. Participants were 183 parents of children aged between 1 and 8 with concerns about their child's behavior and at least one area of disadvantage or family difficulty. Participants were randomized to self-directed Triple P Online, telephone-supported Triple P Online, or a wait-list control (WL). Primary outcomes measured at baseline, post-intervention, and five-month follow-up were negative parenting styles and child behavior problems. Secondary outcomes included: parent confidence, anger, and adjustment; relationship quality; program engagement; and parent satisfaction. Self-directed participants showed short-term treatment effects including reductions in overall negative parenting and frequency of child behavior problems, while practitioner-support led to greater improvements in negative parenting and intensity of difficult child behaviors. Participants in the supported condition were also more likely to complete modules and reported greater program satisfaction. At follow-up, 50% of outcomes for the self-directed condition were significantly better than control, while 94% of outcomes were significantly better than control in the practitioner-supported condition. Although self-directed online approaches to parenting intervention are promising, this research highlights how minimal support can improve effective engagement and enhance outcomes for families.

Keywords: behavioral family intervention, online parenting program, child behavior problems, Triple P—Positive Parenting Program, Triple P Online

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Evidence-based parenting interventions play an important role in the prevention of childhood emotional, behavioral and social difficulties associated with later serious mental health and conduct difficulties (Biglan, Flay, Embry, & Sandler, 2012; Sandler, Schoenfelder, Wolchik, & MacKinnon, 2011). Spurred by recognition within the broader mental health field of the significant global unmet need for services, and corresponding calls for innovative approaches to the development of preventative and treatment-based models that enable broad reach and facilitate greater access to services (Comer & Barlow, 2014; Kazdin & Blase, 2011), there has been growing advocacy by parenting researchers over the last decade for a public health approach to parenting (e.g. Morris et al., 2017). In response, web-based approaches have received increasing attention from parenting researchers in recent years due to the considerable potential of Internet-delivered programs to reach many more recipients than classical forms of parenting interventions (for example intensive individual sessions or group programs).

An emerging body of research supports the efficacy of web-based parenting support for the reduction of disruptive child behaviors in a self-directed context (e.g. Baker, Sanders, Turner, & Morawska, 2017; Sanders, Baker, & Turner, 2012a; Sanders, Dittman, Farruggia, & Keown, 2014a), and with varying levels of clinical support or human feedback embedded (e.g. Enebrink, Högström, Forster, & Ghaderi, 2012; Sourander et al., 2016). The potential benefits of guided delivery versus self-directed online treatments have been well-established within the broader field of web-based psychological and health interventions, including improved treatment adherence and outcomes (Baumeister, Reichler, Munzinger, & Lin, 2014). However, to our knowledge no studies have yet directly compared treatment outcomes, engagement, and consumer satisfaction between self-directed and practitioner-supported variants of a web-based parenting program that seeks to reduce disruptive child behaviors through enhancing the knowledge, skills, and confidence of parents. We believe such an evaluation is timely from both a clinical and policy perspective. Web-based programs are likely to see ongoing and increasing adoption within public health for the dissemination of parenting support, with policy-makers drawn to the notion of cost-effective, large

ACCEPTED MANUSCRIPT

scale dissemination models such as light-touch, self-directed online programs. Likewise, parents are often time poor and may prefer the increased flexibility that comes with accessing information and support through low intensity, self-directed formats such as the Internet (e.g. Metzler, Sanders, Rusby, & Crowley, 2012). However, these benefits of self-directed delivery need to be considered within a context of 'effective engagement' with online programs, such that intended outcomes (in this case, reductions in difficult child behaviors and negative parenting styles, and improvements in parenting skills and confidence) are maximized (Yardley et al., 2016).

To this end, the current study aims to build on prior research by addressing some fundamental questions: what is the direct impact of providing regular practitioner support as an adjunct to a primarily self-directed online parenting program? How can we maximize program engagement (putatively associated with better clinical outcomes), while at the same time providing the minimally sufficient amount of support necessary to see treatment benefits? In line with a public health approach to parenting, so-called 'light-touch', self-directed web-based interventions are more amenable to large-scale dissemination, whereas introducing therapist support both increases costs and reduces scalability (Andersson & Titov, 2014). Hence, it is important to establish whether additional clinical support provides sufficient benefits to parents to justify the associated drawback in terms of reduced (potential) reach. By examining clinical outcomes, program engagement, and satisfaction with treatment across multiple conditions with contrasting levels of intervention and support, we hope to extend the existing evidence base by improving our understanding of the role and importance of the practitioner in this context.

This study focuses on the Triple P Online (TPOL) parenting intervention (Turner & Sanders, 2011), a web-based self-directed adaptation of the empirically-supported Triple P – Positive Parenting Program (Triple P). Triple P is a tiered, multilevel system of behavioral family interventions that draws on social learning theory and cognitive and behavioral principles, targeting parents' knowledge, skills and confidence in effort to reduce or prevent child social, behavioral and emotional problems (Sanders, 2012). Level 4 Triple P interventions are commonly delivered in

ACCEPTED MANUSCRIPT

face-to-face or group contexts over about eight sessions. TPOL is an adaptation of these sessions into eight interactive, online modules that cover topics such as encouraging positive child behaviors, managing difficult child behaviors, and developing a more positive parent-child relationship.

A number of prior TPOL trials have demonstrated treatment effects for various child and parent outcomes when delivered in a self-directed context, including reductions in child behavioral and emotional difficulties, less reported parental anger and use of negative parenting styles, lower levels of conflict with a partner, and decreases in maternal adjustment difficulties (Baker et al., 2017; Sanders et al., 2012a, 2014a). The program has also been shown to reduce ADHD symptoms in children such as restlessness/impulsivity and hyperactivity/inattention (Franke, Keown, & Sanders, 2016), and has shown promise with extremely vulnerable US parents when enhanced with a moderated social-networking component (Love et al., 2016).

In this study we utilized a randomized design to compare clinical outcomes between three groups: a practitioner-supported Triple P Online condition, a self-directed Triple P Online condition, and a wait-list control condition. The primary outcomes of interest were behavior changes directly targeted by the program; namely, reductions in child behavior problems and decreases in parents' use of dysfunctional parenting styles. Secondary outcomes were identified based on findings of prior TPOL and Level 4 Triple P trials (Sanders, Kirby, Tellegen, & Day, 2014b), and included increases in parent confidence and parental relationship quality, along with reductions in parental adjustment difficulties, reported levels of partner conflict, and parents' anger responses towards their child. Finally, we were also interested in consumer reports of satisfaction and indicators of program engagement (e.g. module completion rates, uptake of clinical support calls in practitioner-support condition).

For the practitioner-supported condition, we employed a weekly clinical telephone consultation model based on: (a) the ubiquity of telephone access for most parents; (b) the success of the model for improving web-based treatment outcomes in other eHealth contexts (e.g. Carlbring

ACCEPTED MANUSCRIPT

et al., 2006); and (c) prior evidence that telephone consultations improved outcomes for participants using the "Every Parent's Self-Help Workbook", a self-directed workbook version of Triple P (Morawska & Sanders, 2006). Following from prior efficacy trials, it was hypothesized that parents in both treatment groups (self-directed and practitioner-supported) would show significant improvement relative to a waitlist control condition on both primary outcomes (reductions in the frequency and severity of child behavior problems and dysfunctional parenting practices), and secondary outcomes, including increases in parenting confidence, reductions in parental adjustment, reductions in parental anger towards their child, improvements in the perceived quality of partner relationships, and less parenting-related conflict with a partner.

Given that a clinician's involvement allows for better tailoring of support to the specific needs of the family, along with added motivating factors such as increased accountability (Mohr, Cuijpers, & Lehman, 2011), we expected that parents in the practitioner-supported condition would demonstrate more program engagement (defined as completing more online modules) and significantly greater improvement on primary outcomes (reductions in child behavior problems and dysfunctional parenting styles) than self-directed parents. However, as secondary outcomes are not direct targets of the intervention, we expected the added benefits of clinical support to be smaller, and as such did not hypothesize differences between treatment groups on these measures.

With regards to consumer satisfaction, prior research suggests light touch, self-directed programs are appealing to parents (e.g. Metzler et al., 2012). Given that practitioner support imposes more demands on parents' time and limits capacity for flexibility with regards to participation, it may not be unreasonable to expect self-directed participants to have higher reported satisfaction. However, given we previously hypothesized that parents in the practitioner-supported condition would achieve better primary outcomes, we expected this would outweigh any perceived drawbacks around flexibility, and lead to greater reported satisfaction.

Method

This trial was retrospectively registered with the Australian New Zealand Clinical Trials

ACCEPTED MANUSCRIPT

Registry, ACTRN12614000672651. Approval to conduct the research was granted by the Behavioral and Social Sciences Ethical Review Committee at the University of Queensland, #2012000186.

Participants

Participants were recruited between April 2012 and April 2014 through Facebook, e-newsletters sent through subscription-based parenting networks, pamphlets sent to local childcare and community centers, and referrals and waiting-list mailouts from child health services. Parents registered their interest in participating through an open access website, and were then screened via telephone. For eligibility to participate, families were required to have at least one child aged between two and eight, and report concerns about their child's behavior. Areas of concern were based on items drawn from DSM-IV criteria for conduct disorder, oppositional defiant disorder, and attention-deficit hyperactivity disorder (e.g. "Does your child often say 'no' or refuse to cooperate when asked to do something?"), however formal diagnosis was not a requirement. To narrow our reach to families most likely in need of the type of clinical parenting support provided by a Level 4 Triple P intervention, parents were also required to meet at least one additional socioeconomic or family risk factor that may be associated with higher levels of child social, emotional or behavioral problems (e.g. Lawrence et al., 2015), including: (a) single parent status (separated, divorced or widowed); (b) unemployed (unless supported financially by a partner); (c) in a low education bracket (i.e. one or both parents had not progressed further than high-school); (d) difficulties meeting essential expenses over the last six months; or (e) a score of five or more on the Parent Problem Checklist (PPC; Dadds & Powell, 1991), indicative of clinical levels of conflict with a partner around parenting.

Parents were excluded during screening if they: (a) were currently accessing parenting support elsewhere; (b) had a child with a diagnosed intellectual disability, developmental delay, or Autism Spectrum Disorder (as reported by the parent); (c) reported difficulties reading English without assistance; or (d) reported their Internet connection speed was insufficient to watch a

ACCEPTED MANUSCRIPT

YouTube video. No assessment of child comorbidity was conducted, and parents were not screened on the basis of their own mental health status. Parents not eligible to participate were referred to other local services. Informed consent was obtained through an online portal prior to pre-intervention questionnaires.

Sample characteristics

Parent ages ranged between 22 and 51 (M = 34.9; SD = 5.3), with child ages ranging between 1 and 8 (M = 3.5; SD = 1.5), with 1 one-year-old child enrolled due to a reporting error. About half (53.5%) of the target children were female. Parents were mostly mothers, with only seven fathers recorded as the primary participant (3.8%). The majority of participants lived in Western Australia (79.2%), were married or in a de facto relationship (89%), had children living with their original family (84.7%), and identified as white (93.4%). Regarding education and employment, 60.6% of the participating parents and 38.9% of their reported partners had a university degree, while 54.6% of the participating parents and 96.2% of their partners were in either full- or part-time employment at T1. When asked whether they had difficulties meeting essential expenses within the last six months, 14% responded ‘yes’, while 16.9% reported not having enough money left over for other purchases after paying for essential expenses.

Based on published clinical cut-offs, 64.4% and 60.0% of parents were in the clinical range on the ECBI Problem and Intensity subscales respectively, and 60.4% had a Total score in the clinical range for the Parenting Scale at baseline. Additionally, 78.6% of parents with a partner reported clinical levels of conflict around parenting on the PPC Problem subscale, which may reflect our inclusion of partner conflict as a family risk category during screening. Few parents reported clinical levels of adjustment problems (3.3%, 3.8%, and 2.7% in the clinical range for the DASS Depression, Anxiety and Stress subscales respectively). Chi-square testing showed no differences between groups in terms of base rates of clinical problems.

Design and procedure

A 3 (*condition*: self-directed Triple P Online [TPOL] vs. Triple P Online enhanced with

ACCEPTED MANUSCRIPT

practitioner support [TPOLe] vs. wait-list control [WL]) x 3 (*time*: pre-intervention [T1], post-intervention [T2], five-month follow-up [T3]) randomized design was used for this study.

All questionnaires were completed online using the Qualtrics platform. Random assignment occurred following T1 assessment using a computer-generated sequence of numbers stored in a secure online database and obscured from the research team. Parents and researchers were not blinded to condition following group assignment. Parents assigned to the treatment conditions received their online access code via email, and parents assigned to the TPOLe condition were also contacted by their allocated practitioner to schedule their first telephone consultation.

Program access was provided for four months from the date of first login, at which point access to the program automatically expired. Email reminders were sent to parents that had not logged in to the program one week after their access code was sent, with further follow-up phone calls if still not logged in after two weeks. Automated notification emails were also sent two weeks and again one week prior to expiry to remind participants to finish any remaining modules and download any resources they wished to keep. Parents in the TPOL condition received no further contact from the research team during the four-month intervention period, unless technical assistance was specifically requested. Post-intervention (T2) data collection occurred immediately following program expiry, with follow-up (T3) data collection due five months later. Parents in the WL condition completed assessments four-months after randomization, and again five months later, but otherwise received no further contact during this time. Following completion of T3 assessment, parents in the WL condition were provided with access to the online program.

Intervention

Triple P Online is an eight-module online behavioral family intervention, based on existing Level 4 Triple P interventions (Sanders et al., 2012a). The program covers key concepts of positive parenting, including strategies for managing misbehavior. Multimedia videos are utilized extensively to teach skills and demonstrate strategies, in combination with interactive activities, downloadable resources, and a dynamically-generated workbook for tracking progress through the

ACCEPTED MANUSCRIPT

program. Optional technology-assisted communication tools are embedded, such as SMS prompts during the week reminding parents to try a strategy, or to send module summaries to a partner via email. The program actively encourages the user to set and review goals throughout, while later modules encourage users to identify high-risk situations (e.g. shopping trips) and combine strategies and principles discussed earlier in the program into a cohesive prevention plan.

Up to eight practitioner support sessions were provided to parents randomized to the TPOLe condition. Practitioners in this study were (a) 11 postgraduate psychology students undergoing advanced clinical training at the host institution, and (b) 12 community workers from a state government-funded parenting organization, which regularly delivered Triple P interventions to parents via telemethods. For the community organization, data on years of experience was not provided by practitioners. Anecdotally, many had more than five years' experience implementing Triple P within a community setting. All practitioners were accredited in at least one variant of Level 4 Triple P interventions prior to their involvement and were required to have ongoing clinical supervision and peer support through their organization. Practitioners were also required to attend a half-day training workshop to orient them to the specifics of the project.

Guidelines for telephone consultations were modelled from similar clinical support approaches used in other Level 4 Triple P programs, including: (a) checking that the parent had successfully logged in to the program and/or completed the next module; (b) asking the parent to set an agenda for the session, (c) reviewing module content, (d) reviewing goals or practice tasks from the prior session, (e) discussing agenda items, and (f) discussing an adherence plan if the parent had not engaged with the program since the last consultation. Parents were encouraged to complete one online module and one telephone consultation each week for eight weeks; however, the process was flexible and parents could reschedule calls when legitimate scheduling conflicts arose. Telephone consultations were considered 'missed' if the parent made no attempt to reschedule or notify their practitioner when unable to make a scheduled call.

Measures

ACCEPTED MANUSCRIPT

A modified version of the *Family Background Questionnaire* (Sanders & Morawska, 2010) was used at baseline to collect key demographic information and indicators of socioeconomic status.

Child behavior

Child behavior problems were measured using the *Intensity* and *Problem* subscales of the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999), a 36-item measure of disruptive child behaviors suitable for parents of children aged between 2 and 16. The *Intensity* scale measures the frequency of disruptive behaviors on a scale of 1 (*never*) to 7 (*always*). The *Problem* scale measures how many disruptive child behaviors the parent considers to be a problem using a Yes/No format. The ECBI has shown good test-retest reliability (r = 0.86). Both subscales showed good internal consistency in this sample, with α = .91 and α = .89 respectively.

Parenting practices

Coercive parenting was assessed using the 30-item Parenting Scale (PS; Arnold, O'Leary, Wolff, & Acker, 1993). The PS provides a Total score and three subscale scores (*Laxness*, *Over-reactivity*, and *Hostility*). Parents indicate using a 7-point Likert scale how they would typically handle various disruptive behaviors, with options ranging between more and less effective responses to the behavior. The Total score has good test-retest reliability (r = 0.84). Internal consistency was adequate for the Hostility subscale (α = .66), and good for the Laxness and Over-reactivity subscales (α = .74 for both) and the Total score (α = .86).

Parental adjustment

Parental adjustment was measured using the 21-item Depression, Anxiety and Stress Scale (DASS-21; Lovibond & Lovibond, 1995). The DASS-21 assesses symptoms of depression, anxiety and stress, demonstrates good discriminant and convergent validity, and has good test-retest reliability (r = .71 – .81). Participants indicate to what extent each item applies to them on a scale of 0 to 3. Internal consistency was good for all subscales (Depression, α = .91; Anxiety, α = .79; Stress, α = .89).

ACCEPTED MANUSCRIPT

Parenting confidence

The Parenting Tasks Checklist (PTC; Sanders & Woolley, 2005) assesses task-specific self-efficacy as a measure of parenting confidence. The PTC provides scores on two subscales: *Setting self-efficacy* (e.g. "Going to the doctor") and *Behavioral self-efficacy* (e.g. "Refuses to eat food"). Item responses are given on a scale of 0 (*Certain I can't do it*) to 100 (*Certain I can do it*). Both subscales had good internal consistency ($\alpha = .92$ and $\alpha = .97$ respectively).

Relationship quality and adjustment

Indicators of relationship adjustment included the Parent Problem Checklist (PPC; Dadds & Powell, 1991) and the Relationship Quality Index (RQI), an adaptation of the Quality of Marriage Index (Norton, 1983).

The PPC measures conflict between partners around parenting on a 16-item scale, and has good test-retest reliability ($r = .90$). The *Problem* subscale indicates how often conflict around parenting arises using a Yes/No format, and the *Extent* subscale indicates the perceived severity of these issues on a scale of 1 (*Not at all*) to 7 (*Very much*). Internal consistency for both subscales was good in this sample, with $\alpha = .84$ and $\alpha = .92$ respectively.

The RQI measures relationship satisfaction through five general items rated on a 7-point Likert scale (1 = *Very strongly disagree*, 7 = *Very strongly agree*) and one global item rating overall relationship happiness between 1 (*Unhappy)* and 10 (*Perfectly happy*). The measure demonstrated good internal consistency in the sample ($\alpha = .94$).

Parental anger

Parents' anger responses to their child's problematic behaviors was assessed using the Parental Anger Inventory (PAI; Hansen & Sedlar, 1998). The PAI presents parents with 50 items describing difficult child behaviors and asks them to rate (1) whether the situation has been a problem for them in the past month using a Yes/No format, and (2) how angry the situation makes the parent feel (1 = *Not at all* to 5 = *Extremely*), providing two subscales: *Problem* and *Intensity*. Both subscales have demonstrated good test re-test reliability ($r = .84$ and $r = .91$ respectively) and

ACCEPTED MANUSCRIPT

had good internal consistency in this sample (α = .89 and α = .96).

Parent interviews

Parenting researchers have long recognized the importance of incorporating multimodal or multi-informant assessments into empirical studies, with past efforts typically involving either home or clinic-based observations. Given that the web-based nature of the current intervention voids the requirement that participants are located within close proximity to the research team, there is a subsequent need for new approaches to multimodal assessment that can be readily incorporated into such studies. Harnessing modern technologies such as web-cams is promising for this purpose, but introduces a minimum level of technical proficiency as well as additional equipment costs. Given the combined Internet and telehealth focus of the current study, we instead decided to pilot a telephone interview approach using a modified version of the Parent Daily Report interview (PDR; Chamberlain & Reid, 1987). The aim of these interviews was to capture a brief 'snapshot' of the child's behavior patterns over the preceding 24-hour period, as a low intensity substitute for observational recordings. To obtain an average sampling of daily experiences and behaviors at the time, scores were aggregated over three interviews at each assessment time point, scheduled to take place over a span of one to two weeks (nine interviews per participant in total). While the parent still acts as the main informant in this approach, we were interested in whether the interview format would provide a novel perspective on the child's behavior that is complementary to data collected through online questionnaires.

Due to resource limitations, interviews were conducted with a random subset of participants by the first author and a small team of three trained research assistants (psychology undergraduate students). Interviewed parents primarily provided item ratings for three outcomes: frequency of positive child behaviors, number of difficult child behaviors, and the parent's level of frustration or irritation due to difficult child behaviors. Seven items measured the *frequency of prosocial child behaviors* (e.g. "Being cheerful, showing contentment and self-confidence") using a four-point Likert scale (0 = *Not at all*, 1 = *A bit*, 2 = *Some*, 3 = *A lot*). Nineteen items measured the *number of*

ACCEPTED MANUSCRIPT

occurrences of difficult child behaviors (e.g. "Being aggressive, fighting, hitting, biting, kicking others") using a seven-point Likert scale ranging from *0 times* to *6 or more times*. Finally, for any difficult behaviors that had occurred in the 24-hour period, parents rated their *overall level of frustration* due to that behavior (i.e. "How much did it irritate you?") using the same four-point Likert scale as for prosocial behaviors. The amount of time the parent had spent with the child during the last 24 hours was also recorded to use as a covariate in analyses.

Participant satisfaction

Client satisfaction was assessed using a Client Satisfaction Questionnaire (Sanders, Markie-Dadds, & Turner, 2012b) at post-intervention for the treatment groups only. The CSQ is a 13-item measure commonly used within Triple P research to assess consumer satisfaction across a range of indicators, such as the quality of the service received, whether the program has met the needs of the family, and whether they felt the program had equipped them to deal with problems effectively. Items are phrased as questions, e.g. "To what extent has the program met your needs?", rated on a scale of 1 to 7 with higher scores indicating greater satisfaction. Items are summed to attain a total score ranging between 13 and 91.

Program engagement

Our main indicator of participant engagement and adherence to treatment included website metrics used to track time spent in online modules and number of modules completed. Practitioners also recorded the date and duration of telephone consultations for each participant. We were unable to monitor treatment fidelity within clinical telephone consultations due to privacy concerns and lack of access to the necessary recording equipment within our partner organization.

Statistical approach

Missing data

Review of available data revealed missingness at both the item level and the participant level (i.e. dropout over the course of the intervention period). The assumption of missing completely at random (MCAR) was not supported ($p > .05$), though visual analysis suggested a

ACCEPTED MANUSCRIPT

primarily monotonic mechanism. To capitalize on all available data, multiple imputation of missing data by fully-conditional specification was used to replace missing item-level responses (van Buuren, 2007)[1]. This item-level imputation strategy under the assumption of data missing at random (MAR) was considered appropriate given the large number of informative demographic and descriptive variables available as auxiliary variables at T1 (Collins, Schafer, & Kam, 2001). Fifteen imputations were generated from 30 iterations each, using predictive mean matching as the imputation method and with treatment condition entered as a level-2 cluster variable. Model-based imputation through maximum likelihood estimation of mixed-effects models was used to handle missingness from dropout.

Main analyses

For questionnaire data, longitudinal mixed-effects models were used with the intent-to-treat sample, with assessment time (level 1) nested within participants (level 2). A minimum sample size of 150 was deemed adequate for detecting cross-level interactions with power at the 0.80 level, based on Kreft and Leeuw's (1998) review of power and sample size within multilevel models. All analyses were conducted using *R* (R Core Team, 2016).

For each outcome, a series of hierarchical models were estimated using maximum likelihood (ML). Models were fit within each imputed dataset, and fixed effects estimates pooled using Rubin's (1987) combining rules and the Barnard-Rubin adjustment for degrees of freedom. For each outcome, a baseline random intercept model was first estimated (Model 1). *Time* was added as a fixed effect in Model 2, *time* and *condition* as fixed effects in Model 3; and the *time* x *condition* interaction term entered in Model 4[2]. As the interaction effects were of primary theoretical interest, interaction models were re-estimated using restricted maximum likelihood (REML) and pairwise

[1] Item-level imputation can provide greater statistical accuracy for parameter estimates than scale-level imputation (Enders, 2010). As this approach involves imputing a prohibitively large number of variables, we followed the three-step imputation strategy recommended by Enders (2010), which involves computing temporary subscale scores for use as auxiliary variables that are then iteratively replaced with subscale scores from imputed item-level data.

[2] Random slopes for time were included in Models 2 – 4 to account for variation between participants in rates of change over time, using an unstructured covariance matrix which allows intercept and slope variances and covariances to be freely estimated within the model. Residual within-person variances were estimated using an identity covariance matrix, which assumes error independence and homogeneity. Autoregressive residual errors were tested with a random selection of outcomes but did not improve model fit and so were not explored further.

ACCEPTED MANUSCRIPT

group comparisons conducted to investigate hypothesized short and long-term effects with significance set at $\alpha = .05$. The proportion of variance explained by each model was calculated using the R^2_{lmm} approach recommended by Nakagawa and Schielzeth (2013) for linear mixed-effects models, which partitions variance explained into that which is attributable to fixed effects only (marginal R^2), or both fixed and random effects (conditional R^2). Pseudo-R^2 values were averaged across imputations for reporting.

To analyze interview data, mixed effects models were replaced with hierarchical linear models and imputation of missing data was conducted using the Expectation Maximization algorithm (EM), given these were secondary analyses on a smaller subset of participants.

Effect sizes

Cohen's *d* effect sizes were calculated for each outcome as the mean difference in change scores between groups divided by the pooled pre-treatment standard deviation, with small-sample bias correction applied (Morris, 2008). Effect sizes were pooled across imputations and interpreted using conventions of small (0.2), medium (0.5), and large (0.8).

Reliable and clinical change

Clinical significance and reliable change was calculated at post-intervention for the following primary outcome measures: ECBI Intensity, ECBI Problem, and Total score on the Parenting Scale. Clinical significance is defined as the proportion of participants moving out of the clinical range based on published clinical cut-offs. Reliable change refers to the proportion of individuals in each condition showing more change than would be attributable to measurement error alone. Reliable change was computed using Jacobson and Truax's (1991) reliable change index (RCI) formula based on internal consistency estimates and standard deviations of pre-intervention baseline scores. The proportions showing reliable or clinical change were averaged across imputed datasets and rounded to the nearest whole number for ease of interpretation. Finally, the proportion of participants in each group that showed both reliable change and clinical improvement (i.e. clinically reliable change) was computed (Rajwan, Chacko, Wymbs, & Wymbs, 2014).

ACCEPTED MANUSCRIPT

Results

Preliminary analyses and missing data

Table 1 displays demographic characteristics across the sample according to condition. Chi-square tests for independence and univariate ANOVAs revealed only parent ethnicity differed significantly between groups, however given the very small percentage of parents not identifying as white in the sample overall, this difference was not expected to introduce any interpretive difficulties.

The flow of participants through the study is outlined in Figure 1. In total 373 parents registered an expression of interest in participating. Nine could not be contacted to determine eligibility. The remaining 364 parents were screened for eligibility, with 205 (56%) meeting inclusion criteria. One parent declined to provide consent and 21 parents did not complete T1 assessment, thus 183 parents in total were randomized to condition.

Fifty-seven parents were allocated to the self-directed Triple P Online condition (TPOL), 66 to receive Triple P Online enhanced with practitioner support (TPOLe), and 60 were assigned to the wait-list control condition (WL). Seventeen parents (29.8%) from the TPOL condition did not complete T2 questionnaires and 19 (33.3%) did not complete T3 questionnaires. For TPOLe, 16 parents (24.2%) and 23 parents (34.8%) did not complete T2 and T3 questionnaires respectively, while 7 parents (11.7%) from the WL control condition did not complete T2 or T3 questionnaires. The primary reasons given for withdrawal were enrolment in an alternative parenting program or lack of time. Chi-square analysis indicated a significant difference between the three groups in terms of proportion of parents lost to follow-up, $\chi^2_{(2,\, n = 183)} = 10.43$, $p < .01$. Post hoc comparisons indicated the WL condition had a significantly lower rate of attrition from questionnaires than both the TPOL ($p < .05$) and TPOLe conditions ($p < .01$). We expect the incentive of gaining access to the intervention following the return of T3 data was sufficiently motivating to retain more WL parents in the study.

For parent self-report scales, 2.95% of item responses were missing overall. Seventy-six

ACCEPTED MANUSCRIPT

(41.5%) participants had no missing item-level data across all three assessment time points, and 68.9% of all subscale scores were able to be computed from complete item-level data.

Intervention engagement

Overall, parents completed 4.52 modules on average (*SD* = 3.12; range 0 to 8). Participants in the TPOLe condition completed significantly more modules than TPOL (*M* = 5.62 and 3.25 respectively, median = 7 and 2 respectively), Mann-Whitney *U* = 1090, $p < .001$, $r = .37$. More parents in the self-directed condition did not complete the first module (including not logging in; TPOL: 28.1%, TPOLe: 6.1%), $\chi^2_{(1,\ n = 123)} = 9.32$, $p < .01$ (Yates' continuity correction), and similarly more parents in the practitioner-supported condition completed all eight modules (TPOL: 22.8%, TPOLe: 47.0%), $\chi^2_{(1,\ n = 123)} = 6.76$, $p < .01$ (Yates' continuity correction). Recent data on engagement in behavioral parent training programs found an average pre-treatment attrition rate of 13% (SD = 15%), and a total attrition rate (i.e. any dropout from intervention) of 26% (SD = 18%) (Chacko et al., 2016). Relative to these findings, pre-treatment attrition in our study (i.e. not completing any modules) was lower in the TPOLe condition and higher in the TPOL condition, while overall attrition rates were higher than the average in both conditions (53% and 77.2% for TPOLe and TPOL respectively). Overall, mean module completion time was 62.95 minutes, using a 10% trimmed mean to remove outliers indicative of non-typical program use.

For the practitioner-supported condition, parents participated in 4.36 clinical telephone support sessions on average (*SD* = 2.53; median = 4). Call duration (based on available data) ranged between 5 and 60 minutes, with an average duration of 23.69 minutes (*SD* = 8.26). There was a significant correlation between the number of telephone consultations and number of online modules completed, $\tau = .47$, $p_{(\text{one-tailed})} < .001$ (Kendall's Tau for non-parametric data).

Intervention effects

Means, standard deviations and Cronbach's alphas are shown for each outcome in Table 2, averaged across imputed datasets. Nested model comparisons for hierarchical mixed-effects models are available in the online supplementary material (Table S1). Wald tests indicated that Model 4

ACCEPTED MANUSCRIPT

containing the *time* x *condition* interaction was the best fit for all outcomes except RQI and PAI Intensity. For PS Hostility and DASS Anxiety, Model 4 was significant at the $\alpha = 0.10$ level, but retained for further analyses given the interaction effects were of theoretical interest. For PAI Intensity, Model 2 was the best fit suggesting general improvement over time, $F(2, 63.52) = 5.43$, $p < .01$. The RQI was omitted from further analyses as there was no evidence of change. Pairwise group comparisons in terms of change from baseline at post-assessment (T2) and follow-up (T3) are reported for primary outcomes in Table 3. Secondary outcomes are available in the online supplementary material (Table S2). The main coefficient statistic B_{diff} represents unstandardized differences in change scores between groups, pooled across imputations.

Short-term intervention effects

Short-term intervention effects were explored by examining group differences in rates of change from T1 (baseline) to T2 (post-assessment). At T2, the TPOLe condition showed significant improvement relative to TPOL on the ECBI Intensity subscale, $B_{diff} = -14.18$, $t(254.55) = -3.06$, $p = .002$, $d = 0.50$. There were no other significant differences between the two active treatment conditions at T2. Short-term effects for each treatment condition relative to control are outlined below.

Child behavior

For ECBI Intensity, there was no difference in pre-post change between TPOL and WL, however there was a significant medium to large effect of intervention for the TPOLe condition relative to WL ($d = 0.76$). Both intervention conditions showed significant short-term improvements relative to control on the ECBI Problem scale, with a medium effect of intervention for TPOL ($d = 0.66$) and a large effect for TPOLe ($d = 0.93$).

Parenting practices

Relative to control, both intervention conditions showed significant pre-post change in terms of parenting practices based on PS Total scores, with a small treatment effect for TPOL ($d = 0.39$) and medium effect for TPOLe ($d = 0.73$). The TPOLe condition also showed significant short-term

ACCEPTED MANUSCRIPT

improvements on all PS subscales, while TPOL significant improvement on the PS Laxness subscale, with effect sizes ranging between $d = 0.26 – 0.61$.

Secondary outcomes

Relative to control, parents in both active conditions showed significant pre-post improvements in terms of stress as measured by the DASS-21, with TPOLe also showing short-term improvements in depression and anxiety. Significant short-term improvements in parent confidence were found for both treatment conditions based on PTC Setting and Behavior subscales. Additionally, conflict over parenting as measured by the PPC Problem and Extent subscales was significantly reduced at post-intervention for both active groups relative to control, and short-term reductions in the frequency of anger responses (PAI Problem scale) were found for both conditions. Comparable short-term improvements were found for treatment conditions on the PAI Intensity scale, however the main effect of *time* was a better fitting model, suggesting the most parsimonious interpretation of the data was that there was a general improvement over time for all three conditions. Treatment effects for significant outcomes ranged between $d = 0.32 – 0.57$ for TPOL and $d = 0.34 – 0.57$ for TPOLe.

Long-term intervention effects

Comparisons between the two treatment conditions revealed a significant difference in slopes from baseline to T3 on the ECBI Intensity and Problem subscales, with $B_{diff} = -13.94$, $t(267.95) = -2.49$, $p = .013$, $d = 0.50$ and $B_{diff} = -4.83$, $t(263.43) = -2.79$, $p < .006$, $d = 0.75$ respectively showing greater improvement for TPOLe relative to TPOL. Participants in the TPOLe condition also showed significantly greater reduction in negative parenting practices at T3 relative to TPOL based on PS Total scores, $B_{diff} = -0.30$, $t(243.99) = -2.27$, $p < .024$, $d = 0.70$. There were no further significant differences between treatment conditions in terms of change from pre-treatment to follow-up. Long-term effects for TPOL and TPOLe relative to WL are outlined below.

Child behavior

Relative to WL control, TPOLe participants showed significant improvement at T3 for

ACCEPTED MANUSCRIPT

ECBI Intensity and Problem scales (d = 0.70 and 1.28 respectively). TPOL showed a marginally significant medium effect on the ECBI Problem scale (d = 0.52; p = .050), while no long-term treatment effects were found for ECBI Intensity.

Parenting practices

Both active conditions showed significant long-term treatment effects on parenting practices as measured by the PS Total scale, with d = 0.40 for TPOL and d = 1.06 for TPOLe. At the individual subscale level, the TPOLe condition showed significant long-term improvement on all PS subscales (Laxness, Over-reactivity and Hostility) with effect sizes ranging between d = 0.48 and 0.82. No long-term effects were found for PS Hostility in the TPOL condition and significant short-term effects on PS Laxness were not maintained at follow-up; however long-term treatment effects were shown for PS Over-reactivity (d = 0.44).

Secondary outcomes

Parents in the practitioner-supported TPOLe condition showed significant long-term intervention effects on secondary outcomes including adjustment (DASS Depression, Anxiety and Stress); parent confidence (PTC Setting and Behavior subscales); conflict with a partner (PPC Problem and Extent subscales); and frequency of anger responses towards the child (PAI Problem subscale), demonstrating maintenance of all significant post-intervention gains. Treatment effects for secondary outcomes ranged between d = 0.36 and 0.81. For the TPOL condition, significant long-term effects were found for depression (but not anxiety or stress); parenting confidence; and conflict with a partner. Significant short-term effects found at T2 for parental anger and stress were not maintained at T3. Effect sizes for significant outcomes ranged between d = 0.34 and 0.76.

Parent interviews

Forty-nine parents (26.8%) were randomly assigned to complete PDR interviews during screening. Two parents withdrew prior to completing the first round of interviews and thus were not randomized to condition. Out of nine scheduled interviews (three at each assessment time point), parents completed an average of 7.42 interviews overall (SD = 2.28). Interview attrition was high,

ACCEPTED MANUSCRIPT

likely due to the added burden to participants, with 10/47 (21.3%) at T2 and 13/47 parents (27.7%) at T3 not completing any PDR interviews. Chi-square analysis using Fisher's exact test indicated no difference at T3 between groups in terms of PDR completion, $N = 47$, $p = 0.16$.

Aggregate measures representing frequency of positive child behaviors, number of negative child behaviors, and parents' level of frustration at difficult child behaviors were obtained by averaging sets of interview scores at each time point. Following aggregation, missing data due to participant attrition was imputed using the EM algorithm to attain balanced data for intent-to-treat analyses. Multivariate repeated-measures analysis of variance (MANCOVA) was used to explore the *condition* by *time* interaction for each of the three interview outcomes (child positive, child negative, and parent irritation), with *average time spent with child* (in minutes) entered as a covariate to control for differences in parents' opportunity to interact with their child over the 24 hours being discussed. Multivariate analysis revealed no significant interaction effect for PDR outcomes, $V = 0.06$, $F(12, 384) = 0.63$, $p = .820$, although there was a significant multivariate main effect of time, $V = 0.11$, $F(6, 254) = 2.44$, $p < .05$. Univariate analyses revealed a significant main effect of time for frequency of negative child behaviors, $F(2, 128) = 4.35$, $p < .05$, and parental frustration, $F(2, 128) = 4.48$, $p < .05$, reflecting a general decrease in negative outcomes over the course of the study, while no change was found in terms of frequency of reported positive child behaviors, $F(2, 128) = 1.94$, $p = .15$. Means, standard deviations, and univariate effects for parent interview data are available in the supplementary materials (Table S3).

Clinical and reliable change

Table 4 outlines the proportion of participants in the clinical range at pre-intervention that showed clinically significant change at post-intervention on primary outcomes. Published clinical cutoffs were used, with raw scores of 131 and 15 for ECBI Intensity and Problem scales respectively (Eyberg & Pincus, 1999). For PS Total, a raw score of 3.2 was used, calculated as 1SD above the mean of the non-clinic sample from the validation study (Arnold et al., 1993). All omnibus chi-square tests were significant. Pairwise comparisons showed a significantly higher

ACCEPTED MANUSCRIPT

proportion of participants moving out of the clinical range following treatment for both conditions relative to WL on the ECBI subscales, while only TPOLe showed a significantly higher rate of clinical change on PS Total scores.

Both treatment conditions had significantly greater proportions of participants showing reliable change across the three main indicators. When clinical improvement and reliable change were examined together (i.e. clinically reliable change), in the TPOLe condition, higher proportions of the group showed clinically reliable change for the ECBI Intensity (43.9%), ECBI Problem (39.4%) and PS Total scales (42.4%) than in the TPOL condition (19.3%, 35.1%, and 26.3% respectively) and WL condition (16.7%, 18.3%, and 13.3% respectively).

Participant satisfaction

As hypothesized, participants in the TPOLe condition had significantly higher overall satisfaction (M = 72.19, SD = 12.48) than the TPOL condition (M = 65.13, SD = 12.28), $t(80.46) = -2.65$, $p < .01$. Most parents across both conditions rated the quality of the service received at least 5/7 ("good") or better (TPOL: 89.5%; TPOLe: 88.0%), $N = 88$, $p = 1.00$ (Fisher's exact test). When asked about satisfaction with the *amount* of help received, 90.0% of parents in the TPOLe condition and 68.4% of parents in the TPOL condition reported feeling at least "satisfied" (a score of 5/7 or higher), $\chi^2_{(1, N = 88)} = 5.14$, $p < .05$ (Yates' continuity correction). Similarly, when asked whether they received the *type* of help they wanted (i.e. "Did you receive the type of help you wanted from the program?"), 88% of parents in the TPOLe condition and 68% of parents in the TPOL condition responded with a score of 5/7 ("yes, generally") or higher, $\chi^2_{(1, N = 88)} = 3.95$, $p < .05$ (Yates' continuity correction).

Discussion

This trial provides further empirical support for the efficacy of web-based parenting support using Triple P Online and highlights the additive benefits made possible through supported delivery using regular clinical contact. These findings are in line with evidence from prior eHealth studies demonstrating even small changes in support orientation may significantly impact program

ACCEPTED MANUSCRIPT

adherence and outcomes (e.g. Kleiboer et al., 2015). As hypothesized, relative to WL the practitioner-supported TPOLe condition showed significant reductions in intensity/frequency and number of difficult child behaviors; reductions in negative parenting practices and parental anger; improved adjustment and confidence; and less conflict with a partner at post-intervention, with all short-term changes maintained at follow-up. Relationship quality (measured using the RQI) was the only self-report outcome not showing significant improvement over time, though this was not a direct target of the intervention.

Participants in the self-directed TPOL condition demonstrated initial reductions in the overall number of reported child behavior problems as well as reductions in negative parenting practices, but not intensity/frequency of difficult child behaviors. Secondary outcomes for self-directed participants included lowered stress, improved parental self-efficacy, fewer reported incidents of anger towards their child, and less conflict between partners. Short-term improvements were largely maintained at follow-up, with the exception of parental laxness, anger and stress, while number of child behavior problems were marginally significant at follow-up ($p = .05$), with 95% confidence intervals for effect size ranging between 0.15 – 0.89. In the opposite direction, we saw additional evidence of long-term benefits not seen immediately following treatment in terms of parent over-reactivity and depression.

Direct comparison of outcomes between the two active treatment conditions revealed initial benefits in favor of the practitioner-supported condition for intensity of child behavior problems, with all primary outcomes (frequency and intensity of child behavior problems and reductions in overall negative parenting) significantly better at follow-up. In total, 94% of all self-report outcomes were significantly better than WL for the practitioner-supported condition at follow-up, while 50% of outcomes were significantly better than WL for the self-directed condition. Point estimates of effect sizes were also stronger for the majority of TPOLe outcomes relative to TPOL. Overall, as expected, these findings support stronger treatment effects for both child behavior and parenting when completing the online program with clinical support. For self-directed parents,

ACCEPTED MANUSCRIPT

treatment appeared to mostly impact parenting factors (e.g. negative parenting styles, parenting confidence, and partner conflict), with some evidence of smaller treatment effects on child behavior difficulties.

In terms of clinical change, both treatment conditions showed a significantly higher proportion of children moving out of the clinical range for behavior problems at post-intervention relative to WL control. For negative parenting practices, while significant treatment effects were found for both active conditions, only TPOLe showed more clinical improvement relative to WL. We note that the self-directed condition had the least number of parents in the clinical range at baseline, which may have influenced this finding.

One plausible explanation for the differences in treatment outcomes is the discrepancy between groups regarding intervention engagement (i.e. module completion). As hypothesized, parents completing the program with clinical support showed greater engagement (i.e. were more likely to start and finish the program) than self-directed parents. Because sequential completion of online modules was enforced programmatically, higher module drop-off rates in the self-directed condition means a greater proportion of parents did not progress past early content focused on strategies for preventing challenging behaviors (e.g. descriptive praise, developing positive parent-child relationships), to later program modules that introduce strategies for responding to and managing difficult child behaviors. This pattern of engagement may account for the trend towards reductions in reported frequency of difficult child behaviors for self-directed participants, but not behavioral intensity, in this group.

There are several key questions arising from these findings around the specific mechanisms of action when clinical support is provided. While our results indicated that higher uptake of telephone calls was associated with more module completion, what is not yet known is the primary mechanism responsible for improving outcomes relative to self-directed participation (Kazdin, 2007). For example, as posited by Mohr and colleagues (2011), the addition of human support increases social presence and accountability, factors which in turn are linked with treatment

ACCEPTED MANUSCRIPT

adherence. In a sense, this model positions adherence as the end goal, with mechanisms of action such as increasing accountability and motivation as mechanisms of action driving this.

Alternatively, clinical support allows for more individualized treatment, with sessions tailored to the specific needs of the family. This provides practitioners with an opportunity to focus on improving and addressing parents' core self-regulatory skills such as self-sufficiency, self-efficacy, and self-management (Sanders & Mazzucchelli, 2013), all of which are central theoretical underpinnings of the Triple P framework, theorized to be important mechanisms responsible for promoting lasting parenting change. Subsequently, perhaps greater engagement is a side-effect of these processes, with the core treatment improvements arising from a level of tailored support that is difficult to achieve through automated, self-directed treatment approaches alone. One option for future studies may be audio recording of session consultations to allow for subsequent content analysis, which may provide further insight into the practitioner's role. For example, if discussion primarily focuses on troubleshooting technical difficulties or improving motivation through accountability, these tasks may be candidates for targeting through more sophisticated automated processes. Conversely, if content focuses heavily on the development of self-regulatory skills and tailoring of strategies to the specific family, this would bolster the view that clinical consultations are an essential component for improving program outcomes.

In this vein, it is interesting to note some key comparisons between this trial and the foundational Triple P Online trial reported by Sanders et al (2012a), which saw module completion rates comparable to those observed in our practitioner-supported condition. While Sanders et al (2012a) did not provide clinical support, they did schedule troubleshooting check-ups and reminder prompts via telephone or email for parents that appeared to have disengaged from the online program. Our study provided less scaffolding for self-directed participants, with no reminders or follow-up prompts other than (a) email reminders to start, and (b) notification of program expiry at 3.5 months, but more support for parents in the practitioner-support condition. Comparison of findings from Sanders et al (2012a) with our practitioner-support condition shows similar effect

ACCEPTED MANUSCRIPT

sizes on primary outcomes at post-intervention, but stronger follow-up effects and more significant secondary outcomes in our study. We theorize from this that automated, 'nudge' reminders have some benefit for parents in terms of program engagement, while additional benefits may be achieved through in-person clinical consultations (whether physical or virtual). This view reflects prior findings in the literature (e.g. Baumeister et al., 2014), though further research is needed to extend and replicate in a parenting context, and extricate the specific mechanisms of action.

As hypothesized, parents in the practitioner-supported condition rated their overall satisfaction higher than parents in the self-directed condition, and were more likely to report that the amount and type of help received was what they wanted. On average, parents allocated to support participated in around four out of eight telephone consultations. It seems that despite reducing the flexibility, autonomy, and anonymity often associated with an online program, parents were glad to receive the additional help when offered. That said, our impression was that legitimate scheduling conflicts notwithstanding, some parents had a strong preference for or against allocation to a clinical support condition. Recording parents' preferences prior to randomization may have provided further insight into the uptake and effectiveness of clinical consultations. For example, some parents allocated to clinical support did not participate in any telephone consultations, though this did not necessarily reflect their engagement with the online modules. Other clinical factors such as the practitioner's knowledge, skill and therapeutic alliance were outside the scope of this study to explore, but likely also influenced the effectiveness and uptake of clinical support.

Limitations

There are some limitations to consider when interpreting the results of this trial. First, though we piloted a telephone interview approach to determine its feasibility for use as a multimodal assessment tool alongside a web-based parenting program, we found no treatment effects other than a general decrease in reported negative child and parent behaviors over time for all conditions. As parent interviews were resource intensive and thus only feasible with a small subsample of participants, it is possible that the study lacked power to detect more subtle interaction

ACCEPTED MANUSCRIPT

effects. It is also plausible that the interview process itself affected parents' reporting of negative behaviors. We strongly advocate for multimodal assessment of treatment outcomes, and as such our findings based on parent report alone should be interpreted with care. As dissemination of web-based treatments are not limited by proximity, an important consideration and challenge for researchers going forward will be finding novel ways to conduct multimodal assessment over potentially long distances to obtain rigorous data on program outcomes and from multiple perspectives. Important advances have been made around the use of technology for observation, intervention and assessment within the broader mental health field (e.g. Comer & Barlow, 2014), and we believe this is an area that deserves ongoing attention and innovation in the context of parenting.

Another limitation is the general lack of diversity in the present sample, with participating parents mostly white, mothers, married or living with a partner, and employed with tertiary-level qualifications. We acknowledge the need for ongoing research around effective engagement with fathers, as well as more ethnically diverse populations and those experiencing the highest levels of disadvantage. Promising steps have already been made in this area with an extremely vulnerable population of parents within the United States (Love et al., 2016). A final limitation was the lack of assessment of comorbidity of child and family difficulties. We encourage more research around the issue of comorbidity and the effect of multiple risk factors on the efficacy of online parenting programs. However, in the context of community dissemination of online parenting programs, our use of practitioners from a community agency combined with the absence of rigorous assessment or screening based on comorbidity, provide significant ecological validity for the present study. As such, we expect our findings to be reasonably generalisable to other community settings seeking to implement a similar approach to the dissemination of online parenting programs.

Conclusions

Online parenting programs delivered within a public health framework have significant potential to increase the reach and impact of quality parenting support (Metzler et al., 2012). The

ACCEPTED MANUSCRIPT

current research lends further support for the efficacy of Triple P Online as a self-directed parenting intervention, though lower engagement due to lack of support scaffolding for these parents appears to have reduced treatment effects. We acknowledge that reliance on a practitioner imposes natural limitations on the capacity for widespread reach and availability of an online parenting program, and more work is needed to better establish the minimally sufficient level of clinical support necessary to maximize benefits for families. Further work is also needed to identify how treatment outcomes and support preferences are moderated by parent characteristics such as cultural factors, sociodemographic characteristics, and pre-existing family or adjustment difficulties. As agencies and policy-makers increasingly consider adopting a blended model of web-based parenting support incorporating brief clinical contact, questions around how much practitioner time is required become important for administrative decisions around costing and assigning resources. Based on uptake of telephone consultations in this trial, offering more than four clinical calls (scheduled weekly or fortnightly), may be unnecessary or too imposing for many parents, though moderating factors are currently unclear. Ancillary costs such as unscheduled interactions between practitioners and parents (e.g. emails or SMS) need consideration also. Going forward, clear guidelines for practitioners supporting parents through web-based parenting programs would be beneficial, along with further research around identifying which parents are most likely to benefit from support.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

References

Andersson, G., & Titov, N. (2014). Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *13*(1), 4–11. https://doi.org/10.1002/wps.20083

Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, *5*(2), 137–144. Retrieved from http://psycnet.apa.orgjournals/pas/5/2/137

Baker, S., Sanders, M. R., Turner, K. M. T., & Morawska, A. (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. https://doi.org/10.1016/j.brat.2017.01.016

Baumeister, H., Reichler, L., Munzinger, M., & Lin, J. (2014). The impact of guidance on Internet-based mental health interventions — A systematic review. *Internet Interventions*, *1*(4), 205–215. https://doi.org/https://doi.org/10.1016/j.invent.2014.08.003

Biglan, A., Flay, B. R., Embry, D. D., & Sandler, I. N. (2012). The critical role of nurturing environments for promoting human well-being. *American Psychologist*, *67*(4), 257–271. https://doi.org/10.1037/a0026796

Carlbring, P., Bohman, S., Brunt, S., Buhrman, M., Westling, B. E., Ekselius, L., & Andersson, G. (2006). Remote Treatment of Panic Disorder: A Randomized Trial of Internet-Based Cognitive Behavior Therapy Supplemented With Telephone Calls. *American Journal of Psychiatry*, *163*, 2119–2125.

Chacko, A., Jensen, S. A., Lowry, L. S., Cornwell, M., Chimklis, A., Chan, E., … Pulgarin, B. (2016). Engagement in Behavioral Parent Training: Review of the Literature and Implications for Practice. *Clinical Child and Family Psychology Review*, *19*(3), 204–215. https://doi.org/10.1007/s10567-016-0205-2

ACCEPTED MANUSCRIPT

Chamberlain, P., & Reid, J. B. (1987). Parent observation and report of child symptoms. *Behavioral Assessment*, *9*(1), 97–109.

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, *6*(4), 330–351. https://doi.org/10.1037/1082-989X.6.4.330

Comer, J. S., & Barlow, D. H. (2014). The occasional case against broad dissemination and implementation: Retaining a role for specialty care in the delivery of psychological treatments. *American Psychologist*, *69*(1), 1–18. https://doi.org/10.1037/a0033582

Dadds, M. R., & Powell, M. B. (1991). The relationship of interparental conflict and global marital adjustment to aggression, anxiety, and immaturity in aggressive and nonclinic children. *Journal of Abnormal Child Psychology*, *19*(5), 553–567. https://doi.org/10.1007/BF00925820

Enders, C. K. (2010). *Applied Missing Data Analysis*. Guilford Press.

Enebrink, P., Högström, J., Forster, M., & Ghaderi, A. (2012). Internet-based parent management training: A randomized controlled study. *Behaviour Research and Therapy*, *50*(4), 240–249.

Eyberg, S. M., & Pincus, D. (1999). Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory—Revised: Professional Manual. Odessa, FL: Psychological Assessment Resources.

Franke, N., Keown, L. J., & Sanders, M. R. (2016). An RCT of an Online Parenting Program for Parents of Preschool-Aged Children With ADHD Symptoms. *Journal of Attention Disorders*, 1087054716667598. https://doi.org/10.1177/1087054716667598

Hansen, D. J., & Sedlar, G. (1998). Manual for the PAI: The Parental Anger Inventory. Lincoln, USA: University of Nebraska, Clinical Psychology Training Program.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining

ACCEPTED MANUSCRIPT

meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, *3*, 1–27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432

Kazdin, A. E., & Blase, S. L. (2011). Rebooting Psychotherapy Research and Practice to Reduce the Burden of Mental Illness. *Perspectives on Psychological Science*, *6*(1), 21–37. https://doi.org/Doi 10.1177/1745691610393527

Kleiboer, A., Donker, T., Seekles, W., van Straten, A., Riper, H., & Cuijpers, P. (2015). A randomized controlled trial on the role of support in Internet-based problem solving therapy for depression and anxiety. *Behaviour Research and Therapy*, *72*, 63–71. https://doi.org/10.1016/j.brat.2015.06.013

Kreft, I. G. G., & Leeuw, J. de. (1998). *Introducing Multilevel Modeling*. Sage Publications (CA).

Lawrence, D., Johnson, S., Hafekost, J., Boterhoven de Haan, K., Sawyer, M., Ainley, J., & Zubrick, S. R. (2015). *The Mental Health of Children and Adolescents. Report on the second Australian Child and Adolescent Survey of Mental Health and Wellbeing*. Canberra, ACT.

Love, S. M., Sanders, M. R., Turner, K. M. T., Maurange, M., Knott, T., Prinz, R. J., … Ainsworth, A. T. (2016). Social media and gamification: Engaging vulnerable parents in an online evidence-based parenting program. *Child Abuse & Neglect*. https://doi.org/10.1016/j.chiabu.2015.10.031

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343.

ACCEPTED MANUSCRIPT

https://doi.org/10.1016/0005-7967(94)00075-U

Metzler, C. W., Sanders, M. R., Rusby, J. C., & Crowley, R. N. (2012). Using Consumer Preference Information to Increase the Reach and Impact of Media-Based Parenting Interventions in a Public Health Approach to Parenting Support. *Behavior Therapy*, *43*(2), 257–270.

Mohr, D. C., Cuijpers, P., & Lehman, K. (2011). Supportive Accountability: A Model for Providing Human Support to Enhance Adherence to eHealth Interventions. *Journal of Medical Internet Research*, *13*(1), e30. https://doi.org/10.2196/jmir.1602

Morawska, A., & Sanders, M. R. (2006). Self-administered behavioral family intervention for parents of toddlers: Part I. Efficacy. *Journal of Consulting and Clinical Psychology*, *74*(1), 10–19. https://doi.org/10.1037/0022-006X.74.1.10

Morris, A. S., Robinson, L. R., Hays-Grudo, J., Claussen, A. H., Hartwig, S. A., & Treat, A. E. (2017). Targeting Parenting in Early Childhood: A Public Health Approach to Improve Outcomes for Children Living in Poverty. *Child Development*, *88*(2), 388–397. https://doi.org/10.1111/cdev.12743

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, *11*(2), 364–386. https://doi.org/Doi 10.1177/1094428106291059

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Norton, R. (1983). Measuring Marital Quality: A Critical Look at the Dependent Variable. *Journal of Marriage and Family*, *45*(1), 141–151. https://doi.org/10.2307/351302

R Core Team. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.r-project.org/

ACCEPTED MANUSCRIPT

Rajwan, E., Chacko, A., Wymbs, B. T., & Wymbs, F. A. (2014). Evaluating Clinically Significant Change in Mother and Child Functioning: Comparison of Traditional and Enhanced Behavioral Parent Training. *Journal of Abnormal Child Psychology*, *42*(8), 1407–1412. https://doi.org/10.1007/s10802-014-9877-6

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

Sanders, M. R. (2012). Development, Evaluation, and Multinational Dissemination of the Triple P-Positive Parenting Program. *Annual Review of Clinical Psychology*, *8*(1), 345–379. https://doi.org/doi:10.1146/annurev-clinpsy-032511-143104

Sanders, M. R., Baker, S., & Turner, K. M. T. (2012a). 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*(11), 675–684. https://doi.org/10.1016/j.brat.2012.07.004

Sanders, M. R., Dittman, C. K., Farruggia, S. P., & Keown, L. J. (2014a). A Comparison of Online Versus Workbook Delivery of a Self-Help Positive Parenting Program. *The Journal of Primary Prevention*, *35*(3), 125–133. https://doi.org/10.1007/s10935-014-0339-2

Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014b). The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review*, *34*(4), 337–357. https://doi.org/10.1016/j.cpr.2014.04.003

Sanders, M. R., Markie-Dadds, C., & Turner, K. M. T. (2012b). Practitioner's manual for Standard Triple P (2nd ed.). Brisbane, Australia: Triple P International, Pty. Ltd.

Sanders, M. R., & Mazzucchelli, T. G. (2013). The Promotion of Self-Regulation Through Parenting Interventions. *Clinical Child and Family Psychology Review*, *16*(1), 1–17.

ACCEPTED MANUSCRIPT

https://doi.org/10.1007/s10567-013-0129-z

Sanders, M. R., & Morawska, A. (2010). Family Background Questionnaire. Brisbane, Australia: Parenting and Family Support Centre.

Sanders, M. R., & Woolley, M. L. (2005). The relationship between maternal self-efficacy and parenting practices: implications for parent training. *Child: Care, Health and Development*, *31*(1), 65–73. https://doi.org/10.1111/j.1365-2214.2005.00487.x

Sandler, I. N., Schoenfelder, E. N., Wolchik, S. A., & MacKinnon, D. P. (2011). Long-term impact of prevention programs to promote effective parenting: lasting effects but uncertain processes. *Annual Review of Psychology*, *62*, 299–329. https://doi.org/10.1146/annurev.psych.121208.131619

Sourander, A., McGrath, P. J., Ristkari, T., Cunningham, C., Huttunen, J., Lingley-Pottie, P., … Unruh, A. (2016). Internet-Assisted Parent Training Intervention for Disruptive Behavior in 4-Year-Old Children: A Randomized Clinical Trial. *JAMA Psychiatry*, *73*(4), 378–387. https://doi.org/10.1001/jamapsychiatry.2015.3411

Turner, K. M. T., & Sanders, M. R. (2011). Triple P Online. Brisbane, Australia: Triple P International, Pty. Ltd. Retrieved from http://www.triplep.net

van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, *16*(3), 219–242. https://doi.org/10.1177/0962280206074463

Yardley, L., Spring, B. J., Riper, H., Morrison, L. G., Crane, D. H., Curtis, K., … Blandford, A. (2016). Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. *American Journal of Preventive Medicine*, *51*(5), 833–842. https://doi.org/10.1016/j.amepre.2016.06.015

ACCEPTED MANUSCRIPT

Table 1

Participant sociodemographic characteristics by treatment condition at baseline

Variable		TPOL (*n*=57) M (SD)	TPOLe (*n*=66) M (SD)	WL (*n*=60) M (SD)	*F*
Parent age (years)[a]		34.81 (5.16)	35.45 (5.88)	34.50 (4.81)	0.52
					H[b]
Child age (years)[a]		3.44 (1.45)	3.69 (1.70)	3.43 (1.24)	0.82
		% (*n*)	% (*n*)	% (*n*)	χ^2
Location[c,d]	NSW	10.53 (6)	6.06 (4)	5.00 (3)	6.45
	NT	-	1.52 (1)	-	
	QLD	8.77 (5)	12.12 (8)	16.67 (10)	
	VIC	-	1.52 (1)	-	
	WA	80.70 (46)	78.79 (52)	78.33 (47)	
Child gender	Male	42.10 (24)	42.40 (28)	55.00 (33)	2.63
	Female	57.90 (33)	57.60 (38)	45.00 (27)	
Parent's relationship to child[d]	Mother	94.70 (54)	95.50 (63)	98.30 (59)	1.17
	Father	5.30 (3)	4.60 (3)	1.70 (1)	
Relationship status	Married/de facto	89.50 (51)	86.40 (57)	91.70 (55)	0.92
	Single/separated/divorced	10.50 (6)	13.60 (9)	8.30 (5)	
Household[d]	Original family	84.20 (48)	83.30 (55)	86.70 (52)	1.60
	Step family	3.50 (2)	3.00 (2)	3.30 (2)	
	Sole parent	10.50 (6)	12.10 (8)	6.70 (4)	
	Not specified/Other	1.80 (1)	1.50 (1)	3.30 (2)	
Ethnicity[d]	White	91.20 (52)	89.40 (59)	100.00 (60)	**15.34**[**]
	Asian	-	7.80 (5)	-	
	Other (e.g. Pacific Islander, Arab)	8.80 (5)	3.00 (2)	-	
Parent education	High school	19.30 (11)	24.20 (16)	20.00 (12)	1.32
	Trade/technical college	21.00 (12)	18.20 (12)	15.00 (9)	
	University degree	59.70 (34)	57.60 (38)	65.00 (39)	
Partner education	High school	21.60 (11)	25.00 (14)	25.50 (14)	2.04
	Trade/technical college	35.30 (18)	42.90 (24)	32.70 (18)	
	University degree	43.10 (22)	32.10 (18)	41.80 (23)	
Parent employment status	Full time/part time	63.60 (42)	54.40 (31)	45.00 (27)	4.41
	Not working/job seeking	36.40 (24)	45.60 (26)	55.00 (33)	
Partner employment status[d]	Full time/part time	94.10 (48)	96.40 (54)	98.20 (54)	1.23
	Not working/job seeking	5.90 (3)	3.60 (2)	1.80 (1)	
Able to meet essential expenses[d,e]	No	19.30 (11)	13.60 (9)	10.00 (6)	3.27
	Yes	79.00 (45)	86.40 (57)	88.30 (53)	
	Unsure	1.80 (1)	-	1.70 (1)	
Can afford after expenses[a,f]	Not much	21.10 (12)	16.90 (11)	13.30 (8)	3.40
	Some things	50.90 (29)	49.20 (32)	43.30 (26)	
	Most things	28.10 (16)	33.90 (22)	43.30 (26)	

Note. TPOL = Triple P Online (self-directed condition), TPOLe = Triple P Online Enhanced (practitioner-supported condition), WL = passive control condition. ** = Significant at the .01 level.
[a]Data missing for 1 TPOLe case. [b]Kruskal-Wallis rank sum test used due to non-normal distribution. [c]Location refers to parent's location of residence (Australian State or Territory): NSW = New South Wales; NT = Northern Territory; QLD = Queensland; VIC = Victoria; WA = Western Australia. [d]Fishers' Exact Test used for significance testing as expected frequencies <5 for one or more cells. [e]Able to meet essential expenses in the last 12 months. [f]After essential expenses, how much money is leftover for nonessential purchases.

ACCEPTED MANUSCRIPT

Table 2

Pooled means, standard deviations and internal consistencies for parent self-report measures by treatment condition (TPOL, TPOLe, WL)

		TPOL (*n* = 57)			TPOLe (*n* = 66)			WL (*n* = 60)		
		T1	T2	T3	T1	T2	T3	T1	T2	T3
Measure	α[a]	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Primary outcomes										
ECBI Intensity	.91	136.61 (27.46)	122.97 (28.78)	125.96 (25.49)	139.64 (24.68)	113.19 (24.39)	116.11 (29.12)	145.50 (24.02)	137.45 (26.05)	138.85 (23.08)
ECBI Problem	.89	16.00 (6.63)	10.27 (7.31)	12.34 (7.92)	18.05 (6.40)	10.64 (6.05)	9.56 (6.74)	17.01 (6.72)	15.72 (7.28)	16.96 (7.34)
PS Laxness	.74	2.93 (1.11)	2.34 (0.85)	2.48 (0.90)	3.02 (1.05)	2.26 (0.76)	2.22 (0.73)	3.11 (1.04)	2.80 (1.07)	2.85 (1.07)
PS Over-reactivity	.74	4.00 (1.02)	3.64 (1.19)	3.66 (1.17)	4.10 (1.05)	3.33 (1.02)	3.34 (1.09)	3.88 (1.11)	3.77 (1.09)	4.01 (1.07)
PS Hostility	.66	1.84 (0.81)	1.81 (1.01)	1.79 (0.95)	2.06 (0.99)	1.73 (0.85)	1.70 (0.93)	2.08 (1.18)	2.05 (1.18)	2.24 (1.24)
PS Total	.86	3.29 (0.55)	2.89 (0.67)	2.98 (0.62)	3.40 (0.57)	2.79 (0.58)	2.69 (0.65)	3.33 (0.63)	3.15 (0.72)	3.25 (0.76)
Secondary outcomes										
DASS Depression	.91	3.21 (3.58)	2.35 (3.63)	1.71 (2.56)	4.17 (3.92)	2.56 (3.66)	2.23 (3.38)	3.38 (4.31)	3.76 (5.04)	4.18 (4.95)
DASS Anxiety	.79	2.14 (2.29)	1.78 (2.81)	1.68 (2.61)	2.33 (2.85)	1.37 (1.89)	0.95 (1.80)	2.28 (3.52)	2.57 (3.45)	2.42 (3.33)
DASS Stress	.89	6.42 (3.78)	4.95 (4.63)	5.21 (4.55)	7.58 (4.57)	5.17 (4.09)	4.37 (3.96)	6.47 (4.99)	6.77 (4.77)	7.14 (4.92)
PTC Setting	.92	72.83 (15.21)	84.79 (13.94)	87.26 (10.98)	79.29 (13.33)	88.13 (10.44)	88.79 (13.58)	79.04 (13.63)	82.72 (13.92)	82.54 (12.54)
PTC Behavior	.97	60.83 (19.36)	78.94 (17.25)	79.10 (16.05)	62.14 (18.78)	82.66 (14.97)	83.39 (15.15)	62.08 (22.80)	71.25 (21.79)	69.09 (21.30)
PPC Problem	.84	7.06 (3.84)	4.66 (3.92)	4.62 (4.27)	7.85 (4.00)	5.68 (3.79)	5.64 (3.82)	6.89 (3.72)	6.59 (4.36)	6.72 (4.04)
PPC Extent	.92	40.69 (13.55)	31.57 (15.11)	31.88 (15.80)	45.64 (21.10)	34.50 (18.94)	35.18 (17.94)	40.09 (18.44)	38.40 (18.51)	38.68 (17.58)
RQI	.94	32.47 (8.48)	31.40 (10.77)	32.47 (9.32)	31.00 (9.98)	33.97 (10.41)	32.44 (10.63)	33.07 (8.02)	33.50 (9.28)	31.95 (9.38)
PAI Problem	.89	26.40 (7.28)	21.81 (8.61)	22.59 (10.53)	25.84 (7.96)	22.88 (7.94)	20.51 (9.46)	27.24 (7.44)	26.72 (9.03)	26.26 (8.13)
PAI Intensity	.96	112.67 (34.23)	98.97 (30.63)	97.94 (33.32)	113.12 (28.28)	101.93 (29.87)	99.70 (31.07)	111.72 (33.25)	114.49 (33.59)	112.69 (31.89)

Note. Values averaged across imputed datasets. TPOL = Triple P Online (self-directed condition); TPOLe = Triple P Online Enhanced (practitioner-supported condition); WL = Wait-list control condition. T1 = Time 1 (Pre-assessment); T2 = Time 2 (Post-assessment); T3 = Time 3 (Follow-up assessment). ECBI = Eyberg Child Behavior Inventory; PS = Parenting Scale; DASS = Depression Anxiety Stress Scales; PTC = Parenting Tasks Checklist; PPC = Parent Problem Checklist; RQI = Relationship Quality Inventory; PAI = Parental Anger Inventory. [a]Cronbach's alpha.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 3

Pairwise comparisons showing differences in slopes between groups (change from baseline levels) for primary outcomes

	Pre – Post						Pre – Follow-up					
Outcome	B_{diff}	SE	*t*	df	*p*	*d* [95% CI]	B_{diff}	SE	*t*	df	*p*	*d* [95% CI]
ECBI Intensity												
WL vs TPOL	-7.20	4.52	-1.59	269.0	.113	0.22 [-0.14, 0.58]	-3.29	5.45	-0.60	268.9	.547	0.16 [-0.21, 0.52]
WL vs TPOLe	**-21.38*****	4.37	-4.89	249.9	<.001	0.76 [0.39, 1.13]	**-17.22****	5.24	-3.29	268.7	.001	0.70 [0.34, 1.06]
TPOL vs TPOLe	**-14.18****	4.64	-3.06	254.5	.002	0.50 [0.14, 0.86]	**-13.94***	5.60	-2.49	268.0	.013	0.50 [0.14, 0.86]
ECBI Problem												
WL vs TPOL	**-4.34****	1.49	-2.92	263.9	.004	0.66 [0.29, 1.04]	-3.32$^{\sim}$	1.68	-1.97	266.1	.050	0.52 [0.15, 0.89]
WL vs TPOLe	**-5.98*****	1.42	-4.22	259.4	<.001	0.93 [0.56, 1.31]	**-8.15*****	1.62	-5.02	265.3	<.001	1.28 [0.89, 1.66]
TPOL vs TPOLe	-1.64	1.51	-1.09	249.9	.279	0.26 [0.11, 0.62]	**-4.83****	1.73	-2.79	263.4	.006	0.75 [0.38, 1.12]
PS Laxness												
WL vs TPOL	**-0.34***	0.17	-1.99	257.3	.048	0.26 [-0.10, 0.63]	-0.28	0.20	-1.40	257.4	.164	0.18 [-0.19, 0.54]
WL vs TPOLe	**-0.49****	0.16	-3.06	263.4	.002	0.43 [0.08, 0.78]	**-0.59****	0.19	-3.08	268.5	.002	0.51 [0.16, 0.87]
TPOL vs TPOLe	-0.15	0.17	-0.90	262.3	.369	0.15 [-0.20, 0.51]	-0.31	0.21	-1.50	255.8	.135	0.32 [-0.04, 0.68]
PS Over-reactivity												
WL vs TPOL	-0.25	0.20	-1.25	251.3	.213	0.24 [-0.13, 0.61]	**-0.46***	0.21	-2.24	262.0	.026	0.44 [0.07, 0.81]
WL vs TPOLe	**-0.62****	0.19	-3.22	246.8	.001	0.61 [0.25, 0.97]	**-0.80*****	0.20	-4.10	266.8	<.001	0.82 [0.45, 1.18]
TPOL vs TPOLe	-0.36	0.21	-1.82	262.0	.071	0.39 [0.03, 0.75]	-0.34	0.21	-1.64	262.2	.103	0.40 [0.04, 0.76]
PS Hostility												
WL vs TPOL	-0.08	0.15	-0.53	263.6	.597	0.00 [-0.36, 0.36]	-0.19	0.18	-1.09	264.7	.275	0.20 [-0.16, 0.57]
WL vs TPOLe	**-0.30***	0.15	-2.06	261.7	.040	0.27 [-0.08, 0.62]	**-0.46****	0.17	-2.71	268.7	.007	0.48 [0.13, 0.83]
TPOL vs TPOLe	-0.22	0.15	-1.41	263.8	.159	0.32 [-0.03, 0.67]	-0.26	0.18	-1.45	262.9	.149	0.35 [-0.01, 0.70]
PS Total												
WL vs TPOL	**-0.25***	0.12	-2.08	245.8	.039	0.39 [0.02, 0.76]	**-0.27***	0.13	-2.12	246.5	.035	0.40 [0.02, 0.77]
WL vs TPOLe	**-0.42*****	0.11	-3.71	245.2	<.001	0.73 [0.36, 1.09]	**-0.57*****	0.12	-4.73	265.3	<.001	1.06 [0.69, 1.43]
TPOL vs TPOLe	-0.17	0.12	-1.44	259.9	.151	0.36 [0.00, 0.72]	**-0.30***	0.13	-2.27	244.0	.024	0.70 [0.33, 1.08]

Note. B_{diff} parameter estimates represent differences in slopes (unstandardized beta) between groups. Unless otherwise indicated, negative scores indicate improvement in favor of the second group. Cohen's *d* calculated as difference between groups in pre-post change scores, divided by the pooled baseline standard deviation. Effect sizes pooled across imputed datasets using Rubin's (1987) combining rules and reported as absolute value along with 95% confidence intervals. ECBI = Eyberg Child Behavior Inventory; PS = Parenting Scale. $^{\sim}$ = Borderline significant at $p < .10$; * = Significant at $p < .05$; ** = Significant at $p < .01$; *** = Significant at $p < .001$.

ACCEPTED MANUSCRIPT

Table 4

Pre-post reliable and clinical change

		Clinically improved						Reliably improved					
		Omnibus			Pairwise comparisons (*p*)			Omnibus			Pairwise comparisons (*p*)		
Measure	Group	% (*n*/*n*)	χ^2	*p*[a]	TPOL vs WL	TPOLe vs WL	TPOL vs TPOLe	% (*n*/*n*)	χ^2	*p*[a]	TPOL vs WL	TPOLe vs WL	TPOL vs TPOLe
ECBI Intensity	TPOL	68.2 (15/22)	**23.09*****	<.001	**<.001*****	**<.001*****	1.00	42.5 (17/40)	**27.41*****	<.001	**.024***	**<.001*****	**.016***
	TPOLe	71.4 (20/28)						70.0 (35/50)					
	WL	17.1 (6/35)						18.9 (10/53)					
ECBI Problem	TPOL	65.0 (13/20)	**29.09*****	<.001	**<.001*****	**<.001*****	.920	45.0 (18/40)	**16.24*****	<.001	**.007****	**<.001*****	.525
	TPOLe	59.3 (16/27)						54.0 (27/50)					
	WL	0.0 (0/29)						17.0 (9/53)					
PS Total	TPOL	38.1 (8/21)	**7.53****	.023	.891	**.019***	.093	35.0 (14/40)	**12.97*****	.002	**.047***	**<.001*****	.305
	TPOLe	66.7 (18/27)						48.0 (24/50)					
	WL	32.3 (10/31)						15.1 (8/53)					

Note. Number of participants showing reliable or clinical change calculated from available scale-level data following imputation. Counts averaged across imputations and rounded to the nearest whole number for chi-square analysis; ECBI = Eyberg Child Behaviour Inventory; PS = Parenting Scale; * = Significant at the .05 level; ** = Significant at the .01 level; *** = Significant at the .001 level.

[a]2-tailed *p* value for Fishers' Exact Test reported where expected frequency for any cell is <10.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

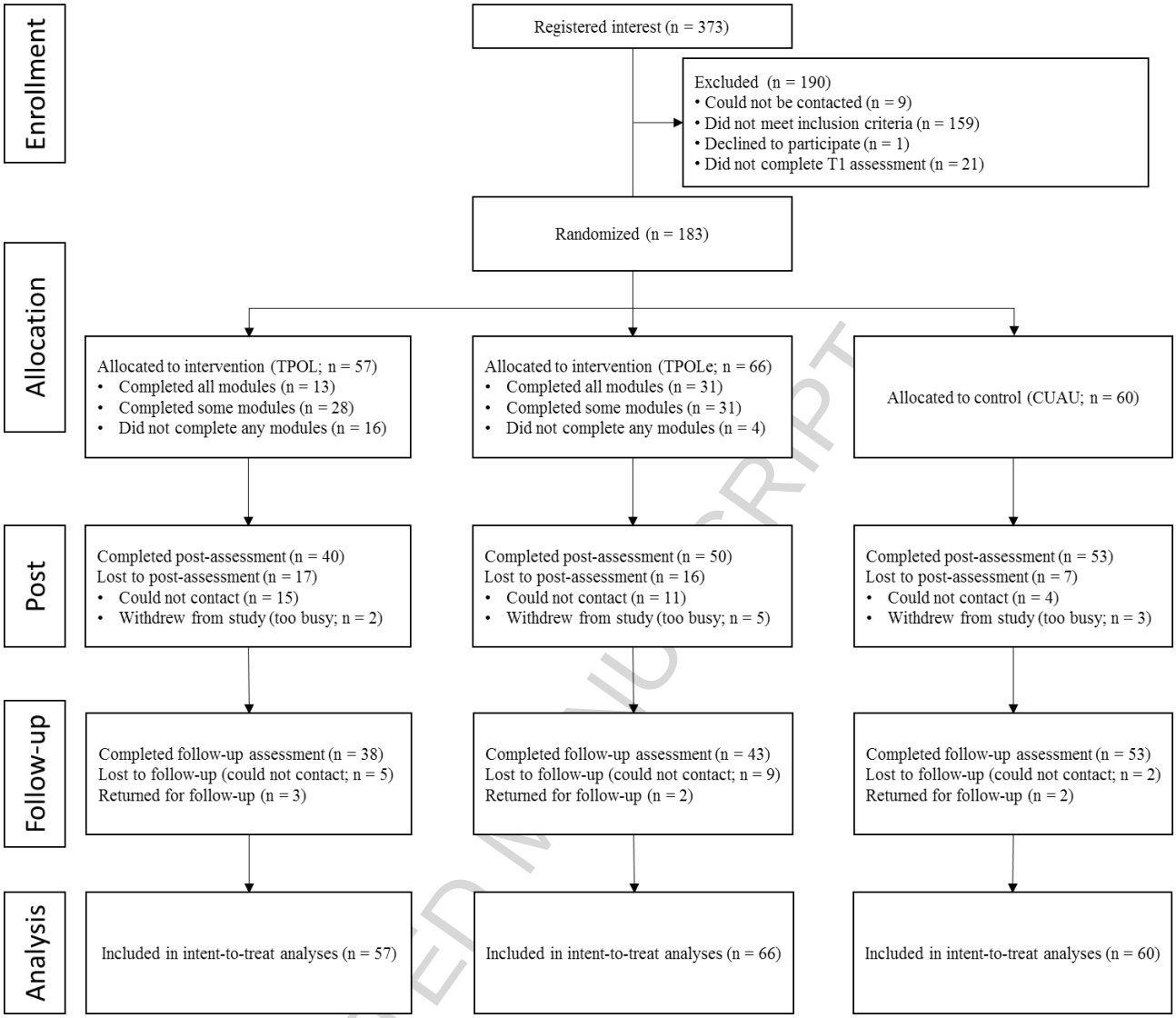

Figure 1. Flow of participants through the study.

ACCEPTED MANUSCRIPT

Highlights

- Randomized controlled trial of an evidence-based online parenting intervention
- Active comparison of self-guided and practitioner-supported program with wait-list control
- Self-directed parents significantly less likely to complete online modules
- Regular, brief clinician contact led to significant gains in parent and child outcomes
- Parents were more engaged and satisfied with treatment when support was provided

ACCEPTED MANUSCRIPT