

A Meta-Analytic Examination of Behavioural Parenting Interventions in the Treatment
of Children's Behaviour Problems

Sarah Bellefontaine

Thesis submitted to the Faculty of Graduate and Post-Doctoral Studies in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in Clinical Psychology

School of Psychology
Faculty of Social Sciences
University of Ottawa

© Sarah Bellefontaine, Ottawa, Canada, 2013

Abstract

The movement towards the provision of evidence-based services has become increasingly important to psychological practice (CPA Task Force on Evidence-Based Practice of Psychological Treatments, 2012; Levant, 2005). One of the key components of this movement is the summarizing and dissemination of important findings to key stakeholders for interventions using techniques such as meta-analysis. Meta-analyses are typically based on reviews of the published literature on a particular topic, but increasingly researchers are acknowledging the importance of examining and including unpublished, grey literature in their reviews in presenting an accurate picture of a body of literature (Hopewell, McDonald, Clarke, & Egger, 2008; McAuley, Pham, Tugwell, & Moher, 2000). However, the inclusion of grey literature in meta-analyses and reviews remains controversial (Coad, Hardicre, & Devitt, 2006; Martin, Perez, Sacristan, & Alvarez, 2005; Seymour, 2010).

A number of reviews of behavioural parenting interventions have been conducted. The body of literature in this area continues to grow quickly, and recent meta-analyses in this area have examined only subsets of the literature, thus necessitating an overall examination of all available literature in this area: published and unpublished. The present meta-analysis synthesized the results of 42 published and 6 unpublished behavioural parenting intervention studies conducted between 1966 and 2011. In the first manuscript, treatment efficacy was examined with respect to three different outcomes (parent behaviour, child behaviour, and parent adjustment). I also examined the differential effects of six moderator variables in two categories: child characteristics (child formal diagnostic status and child age) and methodological characteristics (overall quality, comparison group, experimenter allegiance, and publication status). In the second manuscript, I examined both the advantages and challenges of including

grey literature in meta-analyses, and explored the use of an innovative search program. I also explored the influence of the inclusion of grey literature on outcomes, including whether effect sizes, sample homogeneity, and publication bias indices are affected by the inclusion of grey literature. Results of the present study confirm the efficacy of behavioural parenting interventions on parent behaviour, child behaviour, and parent adjustment outcomes. Further, I extend the knowledge base through the inclusion of recent reports, grey literature, searches of basic search engines such as Google and Yahoo, and the examination of previously unexplored moderator variables. My results underline the importance of consistent reporting practices in the field and their impact on the availability of evidence for policy-makers, administrators, researchers, clinicians and consumers. I also discuss the implications of these findings for the provision of evidence-based services for parents, knowledge translation, and grey literature.

Acknowledgements

This thesis represents the culmination of support from a number of individuals without whom this would not have been possible. First and foremost, I would like to thank my thesis supervisor, Dr. Catherine Lee, for her encouragement and guidance over the past seven years. You have been an invaluable mentor, and have taught me to be a strong woman, both in my roles as a clinician and as a researcher. You always supported me 100%, and that means more to me than you will ever know.

To my committee members, Drs. John Hunsley, Cary Kogan, and Barry Schneider, thank you for your guidance and feedback in both the development of this research project and the writing of this document. I have enjoyed working with each of you collaboratively, and have learned a lot from each of you.

A heartfelt thanks to my labmates Krista Luedemann, Veronica Asgary, Catherine Horvath, Kathy Chan, and Jeffrey Perron for letting me run ideas by them, reading early versions of my work, and supporting me in all my work. Your feedback and support has been invaluable.

To my friends, especially Lindsay Rosval, Michèle Bélanger, Melody Matte, and Stephanie Turner, I cannot express my gratitude for having you there for me throughout this process and ensuring that I remained focused on self-care no matter how busy my schedule became. Your friendships mean the world to me, and each of you has kept me going during the more difficult moments.

Thank you to my parents Ron and Janet Bellefontaine for raising me to be a determined woman. You were both wonderful role models, and have always reminded me that I can do whatever I put my mind to. To my brother Jason Bellefontaine, thank you for believing in me throughout the years, and for keeping in touch with me even when that wasn't easy.

I want to express my deepest gratitude to my husband Brandon Posner for his dedication to me throughout this process and for ensuring that I had all the support I needed to get it done. I am so grateful for your contribution to this dissertation in developing the ReSearch program. Thank you for being the world's best dad to our daughter when I was too busy.

To my daughter Audrey, thank you for reminding me of the simple pleasures in life and to stop now and then to smell the roses, read you a book, or play dollies. I love you.

Statement of Co-Authorship

This thesis contains two manuscripts, one which has been submitted to *Journal of Abnormal Child Psychology*, and the second which has been accepted for publication at *Journal of Child and Family Studies*. Both manuscripts were prepared in collaboration with Dr. Catherine M. Lee, my research supervisor. For each, I am the first author and Dr. Lee is the second author. I was responsible for conceptualizing and planning the meta-analysis, conducting searches, coding studies, training coders, conducting statistical analyses, preparing the manuscripts, and corresponding with journal editors. Dr. Lee was involved in coding and providing support and guidance with respect to the study, search strategy, statistical analyses, as well as manuscript preparation and editing.

Table of Contents

Abstract	ii
Acknowledgements	iv
Statement of Co-Authorship	v
Table of Contents	vi
General Introduction	1
Evidence-Based Practice	1
Grey Literature	8
Disruptive Behavior Disorders	15
Behavioral Parenting Interventions	18
Need For Another Meta-analysis	26
The Current Study	29
Article One	33
Abstract	34
A Meta-Analytic Examination of Behavioral Parent Training in the Treatment of Children's Behavior Problems	35
Method	42
Results	49
Discussion	54
References	60
Table 1	72
Figure 1	75
Table 2	76
Table 3	80
Article Two	81
Abstract	82
Between Black and White: Examining Grey Literature in Meta-Analyses of Psychological Research	83
Exemplar	89
Results	95
Discussion	98
References	104
Figure 1	113
Discussion	114

References	134
Table 1	164
Table 2	167
Appendix A	168
Appendix B	172
Appendix C	174
Appendix D	177
Appendix E	179
Appendix F	181
Appendix G	183
Appendix H	185
Appendix I	187
Appendix J	189

General Introduction

The present dissertation was designed to examine the efficacy of behavioural parenting interventions using meta-analysis, as well as to study the differential effects of variables that may influence outcomes of behavioural parenting interventions. In addition, I examined both the benefits and costs of including unpublished, grey literature in this meta-analysis, as well as the challenges associated with finding such grey literature. Finally I have expanded beyond the usual search methods for identifying grey literature by using the search engines Google and Yahoo.

An overview of the literature on evidence-based practice is presented in order to orient the reader to the importance of research syntheses in general and, in particular, meta-analysis. I then explore the debate in the literature with respect to grey literature inclusion. Finally, an overview of childhood externalizing disorders and behavioural parenting interventions is provided in order to orient the reader to the literature which was synthesized in the current meta-analysis.

Following this introductory literature review, the key research questions and hypotheses are outlined, and then two manuscripts are presented. The first (Bellefontaine & Lee, 2013a) presents the overall findings of the meta-analysis, including moderator analyses; the second (Bellefontaine & Lee, 2013b) provides an examination and discussion of the challenges and benefits of including grey literature, using the previously described meta-analysis as an exemplar. Finally, the implications of this research are discussed in a general discussion section.

Evidence-Based Practice

Over the past two decades, there has been a trend in health care towards evidence-based practice (Hamilton, 2005; Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). Originating in medicine (Evidence-Based Medicine Working Group, 1992), this trend has expanded and been

adopted by other health care professions, including nursing (Craig & Smyth, 2002), social work (Gibbs, 2002) and psychology (CPA Task Force on Evidence-Based Practice of Psychological Treatments, 2012; Levant, 2005). *Evidence-based practice* in psychology was defined by the American Psychological Association's (APA) task force on the topic as integrating clinical expertise with the best research available, taking into account the characteristics and preferences of the people being treated (Levant, 2006). One of the main goals of evidence-based practice is to decrease the use of ineffective interventions and increase the use of ones for which evidence supports its efficacy with particular clinical problems (Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). In 2009, the APA's Presidential task force on the future of psychology practice recommended a focus on ensuring the delivery of evidence-based services by psychologists (APA Presidential Task Force on the Future of Psychology, 2009). One way in which this is achieved is through the development of evidence-based guidelines (Spring, 2007).

Evidence-based practice guidelines.

Guidelines delineating a standard of care have been developed by a number of professional bodies around the world for the treatment of many adult and child psychological disorders. Evidence-based guidelines are based on syntheses of research rather than on expert consensus (Harbour & Miller, 2001; Manchikanti, 2008) and are used to guide the practice of evidence-based services by specifying those interventions that research supports for a particular condition or problem (Spring, 2007). These guidelines may be prepared by national, provincial/state, or regional professional bodies.

For example, evidence-based guidelines from the National Institute for Health and Care Excellence (NICE) in England and Wales are developed through the collaboration of several different professions, require consultation, and are regularly updated to ensure they are based on

the most recent research. NICE has published professional practice guidelines for the assessment and treatment of various mental disorders of childhood. For example, based on a review of the scientific literature, NICE (2006) recommended structured, group-based behavioural parenting interventions as a tool in the management of conduct disorders. According to their recommendations, such training should be 8-12 sessions in length and based on social-learning theory. Professional practice guidelines for the treatment of Attention-Deficit/Hyperactivity Disorder (NICE, 2008) also recommend offering behavioural parenting interventions to caregivers in conjunction with pharmacological interventions for children. Another group which has established interprofessional guidelines in the treatment of children's mental health is Children's Mental Health Ontario. This association consists of individuals and mental health care providers in the province of Ontario whose goal is to ensure that all services provided to children, youth, and their families are evidence-based (Children's Mental Health Ontario, 2008).

As the knowledge base on which practice guidelines are based is constantly expanding, it is recognized that practice guidelines should not be considered to be cast in stone (NICE, 2008). They have a limited time in which they are applicable and relevant, and must be periodically reviewed to determine whether they require revision based on new evidence (Clark, Donovan, & Schoettker, 2006). An essential part of this review process is to consult syntheses of recent research findings.

Research Syntheses

The earliest attempts to synthesize research were done narratively, as opposed to numerically. Narrative synthesis allows researchers to look for patterns in qualitative information collected in research and identify patterns within this information (Baumeister & Leary, 1997). It

has thus been noted to be a good technique to use to build theories or generate alternate research hypotheses (Baumeister & Leary, 1997).

However, narrative reviews do have limitations. For example, criteria for inclusion are rarely explicitly stated, and there are rarely decision rules for weighing the significance of findings (Manchikanti, 2008). Finally, there is no systematic coding and analysis of the findings of included studies; instead, findings are tallied and conclusions are presented based on such a qualitative summary (Manchikanti, 2008). Furthermore, significant bias can be introduced through the omission of portions of the literature due to selection bias (Stanley, 2001), leading to inconsistencies between different reviewers who appear to be examining the same body of literature (Schmidt & Gotzsche, 2005).

More recently, with the amount of available literature increasing exponentially, and a focus on objectivity in research syntheses, researchers have devised procedures to identify, and thus synthesize the best available evidence on a given topic. According to *the Cochrane Handbook for Systematic Reviews of Interventions* (Higgins & Green, 2011), systematic reviews are defined by their presentation of clear, pre-defined inclusion criteria, a specified reproducible methodology, the use of systematic search methods to find all studies that would meet criteria, an assessment of the validity of the findings of a study, as well as the presentation of key characteristics and results of the included studies. Practice guidelines are typically based on conclusions drawn from systematic reviews of the research evidence (Cook, Greengold, Ellrodt, & Weingarten, 1998).

With a greater reliance on systematic reviews to inform evidence-based practice, several bodies have been established to prepare, standardize, and catalog systematic reviews on different topics. These include the Cochrane Collaboration (<http://www.cochrane.org>), the Campbell

Collaboration (<http://www.campbellcollaboration.org>), and the NHS Centre for Review and Dissemination (<http://www.crd.york.ac.uk>).

Cochrane Collaboration. Founded in 1993, the Cochrane Collaboration was developed to help healthcare providers to be able to make well-informed decisions using systematic reviews (Higgins & Green, 2011). To this end, the Cochrane Collaboration has created a standard for how reviews should be presented and structured, a registry which tracks the status of reviews and of randomized controlled trials, offers workshops for reviewers, as well as prepares protocols and software in order to help potential reviewers to facilitate the preparation of systematic reviews (Chalmers, 1993).

Campbell Collaboration. The Campbell Collaboration was inspired by the Cochrane Collaboration and was developed to provide reviews of social and behavioural interventions in education, crime, justice, and social welfare (Davies & Boruch, 2001). Founded in 1999, the Campbell Collaboration's mission was to provide scientifically sound, high quality evidence that could be used and implemented by policy makers, practitioners, and the public (Davies & Boruch, 2001).

NHS Centre for Reviews and Dissemination. Part of the United Kingdom's National Institute for Health Research, the NHS Centre for Reviews and Dissemination maintains a database of research and evidence for health and social care (Centre for Reviews and Dissemination [CRD], 2008). In addition, they are involved in the dissemination of high-quality systematic reviews, as well as the promotion of research synthesis to help others to make informed decisions with regards to health and social care (CRD, 2008).

Meta-analysis.

Systematic reviews offer a standardized way to summarize research findings. Meta-analysis provides a quantitative synthesis of research results. It is a technique used to describe and combine the results of studies numerically, using effect sizes (Hedges & Pigott, 2004). Traditional studies collect and analyze data from individual participants. Similarly, meta-analyses collect and analyze data from different studies. Study effects are generally converted to a standardized measure, the effect size, which allows comparisons across studies (Rosenthal & DiMatteo, 2001). In this way, meta-analyses can be used to summarize the conclusions of a body of research (Harris, 2003) and have thus become the most frequently cited studies (Patsopoulos, Analatos, & Ioannidis, 2005).

Effect sizes for experimental conditions represent the magnitude of the difference between a treatment and no treatment condition on an outcome variable, and can be combined across different studies into overall effect sizes which summarize findings from a body of research (Hedges & Pigott, 2004). Effect sizes range from negative infinity to positive infinity, with a score of zero indicating that there is no difference in the effect between the control and treatment groups (Rosenberg, Adams, & Gurevitch, 2000).

The cumulative or aggregated effect size can be calculated in a number of ways. Typically, a standardized mean difference (d or g) is used so as to ensure that studies are comparable despite their use of different measures of the same outcomes (Borenstein, Hedges, Higgins, & Rothstein, 2009). Using the standardized mean difference, the effect size is calculated by assigning different weights to the effect sizes and variances of different studies within the sample of studies used (Borenstein et al., 2009). A weighted effect size is used so that studies with larger sample sizes, and thus lower variance, are weighted more heavily than are studies with smaller samples and thus greater variance because they more accurately represent

the actual population effect size (Hedges & Olkin, 1985). Because Cohen's d has shown some evidence of bias with small sample sizes, an unbiased, corrected estimate, Hedge's g , is considered suitable (Borenstein et al., 2009).

Meta-analyses can be conducted using either a fixed-effects or random-effects model. In a fixed-effects model, there is assumed to be one true effect size that all studies in the meta-analysis are estimating (Borenstein et al., 2009). However, when it is assumed that there are multiple effect sizes and that the studies included are a random sample of these effects and may not generalize to the population being examined, a random-effects model is employed (Borenstein et al., 2009). The fixed-effects model has been shown to inflate type I error and create confidence intervals that are too narrow when examining a heterogeneous population (Hunter & Schmidt, 2000). When a random-effects model is used, type I error rates and confidence intervals are appropriate whether the studies being examined are homogeneous or heterogeneous (Hunter & Schmidt, 2000). Use of a random-effects model incorporates weights for different studies so as to minimize both within- and between-study variance (Borenstein et al., 2009). A random-effects model was employed in previous meta-analyses of psychotherapy for children's problems (e.g., Weisz, McCarty, & Valeri, 2006) so as to more accurately assess diverse research. Nowadays, it is considered best practice to use a random-effects model, particularly as the random-effects model becomes identical to the fixed-effects model as samples become more homogeneous (Quintana & Minami, 2006). Therefore, I used the random-effects model in the current study to calculate effect sizes.

When there is significant variance between studies, meta-analysts seek to identify the variables that account for this (Borenstein et al., 2009). However, the variation between study effect sizes reflects both of random error as well as the true between-study variance (Borenstein

et al., 2009). To examine whether effect sizes in the sample of studies are homogeneous or heterogeneous, the total heterogeneity of sample (Q ; Hedges & Olkin, 1985) is calculated using the cumulative effect size and confidence interval. This statistic is then tested against a Chi-square distribution with $n-1$ degrees of freedom. A nonsignificant finding indicates that the studies are homogeneous, and thus that no one common effect size exists (Hedges & Olkin, 1985). Heterogeneity in effect sizes indicates the need for moderator analyses (Quintana & Minami, 2006).

In this thesis I conducted a meta-analysis of the outcome of behavioural parenting interventions, examining overall efficacy, the homogeneity of effects, and moderators of treatment effects.

Grey Literature

As with systematic reviews, in meta-analyses the investigator provides explicit criteria for the inclusion of studies (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008). One common criterion is that the study must have been published in a peer-reviewed journal (e.g., Lipsey & Wilson, 2001). Publication in scholarly journals can involve peer-review by several experts in the area in which one has conducted research. By the time of publication, the author is expected to have addressed all ambiguities and have reasonably interpreted the data which he or she is presenting (Solomon, 2007). Thus, it is assumed that the peer-review process serves to enhance the quality of a manuscript by ensuring that it has met scientific standards (Pappas & Williams, 2011).

The majority of meta-analyses only examine the published scholarly literature (Cook et al., 1993; McAuley et al., 2000). This has also been found to be the case with regards to meta-analyses of psychotherapy (Niemeyer, Musch, & Pietrowsky, 2012). Researchers have begun to

question whether reviews that include unpublished literature may offer a more accurate portrait of a body of literature of, for example, registered trials than those that rely on published data (Hopewell et al., 2008; McAuley et al., 2000). Meta-analysts have acknowledged that including unpublished, so-called grey literature in reviews itself causes some challenges in addition to solving some problems (e.g., Coad et al., 2006; Martin et al., 2005; Seymour, 2010). These challenges include the questionable quality of the grey literature, whether its inclusion adds meaningful information, as well as its ability to be located and accessed (McAuley et al., 2000). According to the commonly cited Luxembourg Definition from the Third International Conference on Grey Literature (Gokhale, 1997), *grey literature* refers to print or electronic literature that is not controlled by traditional publishers and is produced by a number of industries (Auger, 1998). Common examples of grey literature include unpublished dissertations, manuscripts, reports, and conference proceedings.

Publication bias.

Researchers (e.g., Clarke & Oxman, 2003; Dickersin & Rennie, 2003) have argued that it is often easier to publish positive results than negative results in peer-reviewed journals. This phenomenon, known as publication bias, refers to the direction and nature of findings determining whether or not they are published in traditional media (Higgins & Green, 2011). A study by Fanelli (2012) which examined the presence of positive-outcome (publication) bias across all disciplines found a strong trend of such bias in psychological/psychiatric literature than in many other disciplines. A landmark meta-analysis of 92 meta-analyses of psychological interventions found that those effect size estimates based only on the published literature were 33% higher than those estimates based on the unpublished literature (Lipsey & Wilson, 1993). Reliance on published literature is thought to lead to an overestimation of the magnitude of

effects (McAuley et al., 2000). Inclusion of the grey literature is commonly thought to minimize publication bias by serving as a balance to the published literature (McAuley et al., 2000; Pappas & Williams, 2011).

Grey literature quality.

Defining the parameters of published literature in terms of date of publication, language, and search strategy is simple, but doing so for grey literature is more difficult and controversial (Auger, 1998; Frater, Myohanen, Taylor, & Keith, 2007; Hopewell et al., 2008). This difficulty in defining the parameters of inclusion, as well as uncertainty regarding the quality of grey literature, has led many meta-analysts to forgo its inclusion. An early review of meta-analyses by Cook et al. (1993) found that 31 to 33% of meta-analyses included grey literature. Echoing this finding, twenty years later, a recent survey of meta-analyses of randomized controlled trials that included individual participant data found that only 29% of those meta-analyses included grey literature (Ahmed, Sutton, & Riley, 2012).

The main challenge that has been consistently identified with regards to grey literature usage is that of the uncertain quality of this literature. Most meta-analysts look for the best evidence sources to include in their reviews so as to ensure the reliability of their findings and produce high-quality reviews (Conn, Valentine, Cooper, & Rantz, 2003). Much variability exists in the quality of grey literature depending on its source. Grey literature studies tend to be smaller than published ones (McAuley et al., 2000). Some research has found that grey literature tend to be of lower quality than published trials (McAuley, 1999). However, the grey literature does include some high quality evidence. For example, dissertations undergo a rigorous review process and have been shown to be of high methodological quality (McLeod & Weisz, 2004). A review of published and unpublished dissertations examining psychosocial interventions for

cancer patients found very few differences in terms of methodological quality (Moyer, Schneider, Knapp-Oliver, & Sohl, 2010). Published dissertations were more likely to have been supported by research funding and were more likely to report significant findings than those that were unpublished (Moyer et al., 2010). Moyer et al. (2010) concluded that unpublished dissertations are of adequate methodological quality to merit inclusion in literature reviews. In contrast, unpublished reports produced by private entities may or may not have gone through such rigorous review processes and are more likely to vary in methodological quality.

The benefits and challenges of including grey literature.

Debate in the literature still abounds regarding whether the inclusion of grey literature really adds meaningful information. A state-of-the-evidence review of 68 studies on the treatment for children and youth with Fetal Alcohol Spectrum Disorder (Benzies, Premji, Hayden, & Serrett, 2006) determined that information found in the 58 documents from the grey literature generally supported the conclusions from the 10 documents from the published literature. Others, however, have found the published literature to overestimate treatment effects. For example, Hopewell et al. (2008) found in a review of meta-analyses of healthcare interventions that greater treatment effects were more likely to be found in the 1402 trials from the published literature than in 346 trials from the grey literature. Research by McAuley et al. (2000) on the literature on grey literature inclusion in meta-analyses determined that the removal of grey literature from 135 meta-analyses of randomized controlled trials on diverse topics led to a 15% increase in the estimated treatment effects, as well as a trend towards greater significance of results.

The inclusion of grey literature may offer additional high quality research to complement that obtained in the published literature. Some researchers have proposed the differential

treatment of different types of grey literature. For example, Martin et al. (2005) argued based on their own research findings that significant and nonsignificant results in the grey literature should be treated differently. They argued that nonsignificant results are harder to publish than are significant results and should therefore be always included in research syntheses (Martin et al., 2005). Significant results, however, may be unpublished due to the delay between the writing of a manuscript and its publication or to the low quality of a manuscript, and therefore advocated for examination of methodological quality prior to its inclusion (Martin et al., 2005).

The major concern remains that such literature is not peer-reviewed and thus its quality has not been vetted in a scientific arena. However, researchers have argued that the peer-review process does not always guarantee the quality of the published literature, asserting that the quality of the published, peer-reviewed literature varies just as much as that of unpublished sources. As previously stated, unpublished dissertations are often of high methodological quality and have been rigorously reviewed (McLeod & Weisz, 2004). One must wonder what high quality information is being missed by excluding the grey literature and how such information may or may not lead to different meta-analytic results. Therefore just as we verify the quality of the published literature that is included in meta-analyses, so too should we expect to have to verify the quality of grey literature sources for inclusion before deeming them of low quality.

Challenges of the peer-review publication process.

Publication of manuscripts in scholarly journals is a long process which prevents a great deal of research from being quickly accessible by research consumers (von Elm, Costanza, Walder, & Tramer, 2003). Furthermore, von Elm et al. (2003), in their examination of publication rates of abstracts submitted to biomedical meetings, found that some research that is conducted is never submitted for publication because of reasons such as obtaining nonsignificant

results or not having enough time. In fact, with journal rejection rates ranging from 31 to 88% (Council of Editors, 2011), some research is never disseminated to those who could benefit from it (Pappas & Williams, 2011). Furthermore, in those cases where research is eventually accepted for publication, significant delays in publication of 2 to 11 months (Council of Editors, 2011) act as further gatekeepers of information. Journals have attempted to address this publication delay through the use of an “On-line First” approach in which an electronic copy of the article, as well as its Digital Object Identifier (DOI) number are available long before the article is printed in the journal (Davidson & Douglas, 1998). However, only works that are accepted for publication are assigned a DOI number, leaving a great number of studies that are not accepted for publication difficult to access. Grey literature has been touted as a solution to such significant barriers to information dissemination (Pappas & Williams, 2011) in that it can be quickly produced, making it more likely to be up-to-date and current (Coad et al., 2006). However, just as the peer-review process is designed to maintain the quality of the published literature, measures must be taken to ensure that similar standards are applied to the grey literature. The exploration of the contribution of grey literature in each field is required to determine its contribution to that body of research.

Knowledge translation.

Publication of research in scholarly journals has been found to be neither sufficient nor ideal for ensuring that research findings are disseminated to those that benefit most from them (Mays & Hogg, 2012). Administrators and practitioners have reported a need to access brief reports and research summaries in an effort to base their policies and practices on the best available evidence (Barwick et al., 2008; Scullion, 2002). Although the grey literature makes research available more quickly than it would be otherwise, its quality is often uncertain, and

searching for grey literature presents challenges for researchers. The advent of the Internet has simplified the task of searching for grey literature. Nonetheless, the use of the internet in searching for grey literature has not eliminated all barriers. Searching for grey literature is often time-consuming and reports often disappear quickly (Coad et al., 2006). Further, search engines such as Google often yield unmanageably high numbers of results (Mahood, van Eerd, & Irvin, 2010) which at times can be adapted to facilitate the search for grey literature (Banks, 2010).

In conducting the present meta-analysis, searches of the grey literature for interventions pertaining to the treatment of disruptive behaviour disorders (DBDs) using behavioural parenting interventions were included in addition to searches of reports published in scholarly journals. The grey literature was accessed through traditional, scholarly databases as well as through documents found through basic search engines such as Google and Yahoo.

Grey literature research questions.

In this thesis, I sought to address several questions about the inclusion of grey literature in a meta-analysis of the outcomes of behavioural parenting interventions. In particular, I examined how much grey literature is available that met my inclusion criteria for my meta-analysis. I then looked at how much of the unpublished literature that was located and included had already been identified in the published literature. I examined whether the effect sizes for grey and published literature were comparable, or whether they differed significantly from one another. In order to address the questions raised by previous researchers, I explored whether a significant difference existed between the methodological quality of studies in the grey literature and those that have been published in peer-reviewed journals. Finally, I determined the overall efficacy of treatment, with and without the inclusion of the grey literature with respect to treatment of DBDs in childhood, as well as the moderators of treatment effects.

Disruptive Behaviour Disorders

Child psychopathology has been examined dimensionally and categorically (Kamphaus & Campbell, 2006). Categorical approaches have predominated in the classification of adult psychopathology and have also been adopted with respect to child psychopathology (Kamphaus & Campbell, 2006). Reflecting a categorical approach, the World Health Organization estimates that 20 percent of children and adolescents have at least one mental disorder, with 4 to 6 percent requiring some degree of clinical intervention (World Health Organization, 2001).

Some experts suggest that a dimensional approach may be more useful in the examination of psychopathology in childhood (Hudziak, Achenbach, Althoff, & Pine, 2007). Factor analytic studies have demonstrated that childhood problems occur on two related, but orthogonal dimensions: internalizing and externalizing problems (Rescorla et al., 2007). Externalizing problems refer to behaviours that are out-of-control and deviant, and they are typically described as discipline problems, behaviour problems, emotional disturbance, acting out, conduct problems, delinquency, and hard-to-manage behaviours (Maughan, Christiansen, Jenson, Olympia, & Clark, 2005). Externalizing behaviour can be difficult on the child, as well as on friends, family, and school, even at subclinical levels (Kazdin, 1987). Behaviour problems are the most common reason for referral of children for psychological services (Chamberlain & Smith, 2003). Given the widespread use of both categorical and dimensional approaches, this meta-analysis was designed to examine children's behaviour problems from both perspectives in order ensure an accurate picture of the efficacy of behavioural parenting interventions with respect to both behaviour problems and related diagnostic conditions is portrayed.

Estimates of the prevalence of behaviour problems in children ages 0 to 5 range from 9.5 to 14.2 percent (Brauner & Stephens, 2006). However, accurate estimates are difficult to obtain,

and estimates often vary from study to study, due to the use of different samples as well as different definitions of what constitutes a behaviour problem. Despite these differences in prevalence estimates, researchers nonetheless agree that such problems are common, and represent a major concern.

Behaviour problems often lead to many difficulties for affected children when left untreated, such as problems at school, difficulties with peers, and trouble with parents and teachers (Frick & McMahon, 2008; Johnston & Mah, 2008). Children who do not eventually grow out of their difficulties with peers, parents, and teachers typically experience poor academic achievement, low socioeconomic status, difficulty with employment, high rates of automobile accidents, family trouble, antisocial behaviour, and mood problems (Frick & McMahon, 2008; Johnston & Mah, 2008). Three diagnoses which fall under the category of disruptive behaviour disorders (DBD) are Attention Deficit/Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD), and Conduct Disorder (CD).

Attention Deficit/Hyperactivity Disorder.

Affecting between 3 and 7 percent of school-age children, Attention Deficit/Hyperactivity Disorder (ADHD) is characterized by symptoms of hyperactivity, impulsivity, and inattention that are inappropriate for the age of the child (American Psychiatric Association [APA], 2000). ADHD is more often diagnosed in males than in females and is more commonly found in children with children who have first-degree biological relatives with ADHD than children in the general population (APA, 2000). Three subtypes of ADHD are diagnosable, based on the symptoms that are primarily experienced (APA, 2000). The Predominantly Inattentive Type is diagnosed when six or more symptoms of inattention are experienced, with few symptoms of hyperactivity or impulsivity. When symptoms of hyperactivity or impulsivity are

the predominant features, the Predominantly Hyperactive-Impulsive Type is diagnosed. Finally, when symptoms of inattention and hyperactivity or impulsivity are both equally predominant, the Combined Type is diagnosed.

Although some children do outgrow this diagnosis, symptoms of ADHD often persist into adulthood, placing people with the diagnosis at an increased risk for psychosocial problems throughout the lifespan (Anastopoulos, Shelton, & Barkley, 2005). Thus, early treatment to address symptoms which cause impairment, as well as the effects on family functioning, are integral. In particular, pharmacological interventions (e.g., stimulant medication) are often used in conjunction with psychosocial interventions (e.g., summer treatment program, behavioural parenting interventions) in order to maximize benefits by targeting both the biological symptoms and social problems associated with ADHD (Pelham & Fabiano, 2008).

Oppositional Defiant Disorder.

Characterized by a pattern of defiant, hostile and negative behaviour towards persons in positions of authority which goes beyond normal childhood behaviour, Oppositional Defiant Disorder (ODD) is a diagnosis which is often found to co-occur with ADHD (APA, 2000). It often refers to behaviour that is characterized as stubborn and disobedient (Kazdin, 2010). Affecting 2-16% of the population, signs of this disorder are normally noticed before age 8, with typical onset occurring no later than early adolescence (APA, 2000). ODD is commonly found in families where one or both parents have a history of ADHD, ODD, Conduct Disorder, Antisocial Personality Disorder, Substance- Related Disorder, or Mood Disorder (APA, 2000). While many children with ODD eventually grow out of it and experience no further difficulties, a significant proportion of children with ODD experience it as a precursor to a more severe behaviour problem, Conduct Disorder (APA, 2000).

Conduct Disorder.

Conduct Disorder (CD) refers to a pattern of behaviour characterized by defiance and hostility towards authority figures, violations of social norms and the basic rights of others, as well as displays of aggressive and often hostile behaviour towards both people and property (APA, 2000). Such acts include firestarting, bullying, cruelty to animals, lying, truancy, forcing others into sexual activity, attacking someone with a weapon, and starting fights with others (Kazdin, 2010). One of the most frequently diagnosed disorders in mental health facilities, CD affects between 1 and 10 percent of the general population, and is found more commonly in males than females (APA, 2000). CD accounts for the highest rate of clinical referrals (Kazdin, 2005). Childhood onset occurs when symptoms of the disorder are present before age 10, whereas the less chronic and severe form, adolescent onset, occurs when criteria are present at or later than age 10 (APA, 2000). CD is more common in children with parental histories of Alcohol Dependence, Mood Disorders, Schizophrenia, ADHD, CD, or with a first-degree relative with Antisocial Personality Disorder (APA, 2000).

Approximately 80% of children who meet criteria for CD will meet criteria for a future psychiatric disorder (Kazdin, 2005). Long term problems which persist into adulthood are common, as are consequences for others in such children's social environments, such as siblings, parents, and teachers (Kazdin, 2005).

Behavioural Parenting Interventions

Treatment guidelines based on scientific findings have been developed for DBDs, such as ADHD (American Academy of Child and Adolescent Psychiatry, AACAP, 2007a; NICE, 2008), ODD (AACAP, 2007b), and CD (AACAP, 1997; NICE, 2006; Children's Mental Health Ontario, 2001; United Nations Department of Crime, UNDOC, 2012). These guidelines have all

identified behavioural parenting interventions as a recommended treatment, either on its own or in combination with other treatment modalities. Behavioural parenting interventions refer to a type of intervention in which parents are taught to become agents of change in their children's behaviour through the use of techniques based on the principles of social learning theory (Bandura, 1971). They are designed to change the parents' behaviour by increasing their use of monitoring, reinforcement, and extinction techniques, and decreasing ineffective parenting behaviours (Briesmeister & Schaefer, 2007). This in turn is presumed to influence the child's behaviour as acceptable behaviours are promoted and maintained and unacceptable behaviours are ignored or mildly punished. Improvements in parents' repertoire of skills to deal with child behaviour problems and reductions in child difficulties are assumed to lead to reduced parental stress and improve parental well-being. Behavioural parenting intervention formats differ between programs, and also sometimes within different programs, and are often times easily adapted to the needs of different groups and different families (Briesmeister & Schaefer, 2007).

Behavioural parenting interventions for working with children with disruptive behaviour began in the 1960's (e.g., Wahler, Winkel, Peterson, & Morrison, 1965). As early as the mid-1970s, researchers such as Nay (1975) had already begun to examine the best strategies (i.e., written, didactic, videotapes, role-playing) for communicating such skills to parents and found no significant differences between these formats.

A variety of behavioural parenting interventions have now been developed for the treatment of DBDs. These include programs developed by Charles Cunningham (Community Parent Education Program; Cunningham, Bremner, & Secord-Gilbert, 1998), Russell Barkley (Defiant Children; Barkley, 1987), Sheila Eyberg (Parent-Child Interaction Therapy; Eyberg & Calzada, 1998), Carolyn Webster-Stratton (Incredible Years; Webster-Stratton & Herbert, 1994),

Matthew Sanders (Triple P – Positive Parenting Program; Sanders, 1999), Gerald Patterson (Oregon Model, Parent Management Training; Patterson, 1982), and Rex Forehand (Helping the Noncompliant Child; Forehand & McMahon, 1981). These programs share a common theoretical foundation that is based in social learning theory (Bandura, 1971). They are all designed to enhance parent-child relationships by teaching parents to increase the number of positive interactions they have with their children, provide rewards for appropriate behaviour, set clear limits with their children, reduce their use of commands, and become more consistent in their discipline (Taylor & Biglan, 1998). These programs are typically delivered by trained and supervised group facilitators who use treatment manuals and materials specified by the program being used. Such behavioural approaches allow parents to set their own parenting objectives and identify personalized goals use role-playing and homework to allow for generalization.

Although these programs all have slightly different focuses and intervention delivery strategies for their participants, they all share a core set of intervention foci. Each of these programs teaches parents to deal with children's behavioural problems by differentially reinforcing any positive behaviour exhibited by the child and ignoring or mildly punishing misbehaviour by the child. Commonly taught strategies include positive attention, delivering clear requests, planned ignoring, response consistency, use of natural consequences, effective use of mild punishment and time out (Maughan et al., 2005). These programs were designed both to treat diagnosed behavioural problems, as well as subclinical levels of problems.

Eyberg, Nelson, and Boggs (2008) provided support for the efficacy of behavioural parenting interventions in the treatment of disruptive behaviour, aggression, delinquency and disruptive behaviour disorders. Their best practice recommendations suggested using behavioural parenting interventions as a first-line intervention in addressing children's disruptive

behaviour. Similarly, Pelham and Fabiano's (2008) review of interventions for ADHD showed support for behavioural parenting interventions.

A number of meta-analyses of behavioural parenting interventions have already been conducted. Table 1 provides a summary. In general, these meta-analyses have found positive effects of behavioural parenting interventions for parent behaviour (Furlong et al., 2012; Kaminski, Valle, Filene, & Boyle, 2008; Lundahl, Risser, & Lovejoy, 2006; Lundahl, Tollefson, Risser, & Lovejoy, 2008; Nowak & Heinrichs, 2008), child externalizing behaviour (Furlong et al., 2012; Lundahl et al., 2006; Lundahl et al., 2008; Maughan et al., 2005; McCart, Priester, Davies, & Azen, 2006; Nowak & Heinrichs, 2008; Piquero, Farrington, Welsh, Tremblay, & Jennings, 2009), and parent adjustment (Furlong et al., 2012; McCart et al., 2006; Nowak & Heinrichs, 2008) outcomes. Effect sizes for parent behaviour ranged from .38 to .54 whereas those for child behaviour were between .30 and .46 when examining weighted effect sizes. Parent adjustment outcomes yielded smaller effect sizes of .17 to .33 in the two studies which examined this outcome.

Given the number of meta-analyses that have examined behavioural parenting interventions, one may wonder whether there is a need for another. However, these previous meta-analyses have notable limitations. For example, several of these examined only particular behavioural parenting interventions (de Graaf, Speetjens, Smit, de Wolff, & Tavecchio, 2008; Fletcher, Freeman, & Matthey, 2011; Nowak & Heinrichs, 2008; Thomas & Zimmer-Gembeck, 2007), thus limiting the generalizability of these findings to only the programs which have been examined. Other meta-analyses have been limited in terms of the age ranges of children (Furlong et al., 2012; Kaminski et al., 2008; Lundahl et al., 2006; Piquero et al., 2009; Thomas & Zimmer-Gembeck, 2007), which makes it difficult to determine the effects of behavioural

parenting interventions at different child ages. Still others have examined studies that looked at children with clinically significant as well as subclinical levels of disruptive behaviour (e.g., de Graaf et al., 2008; Dretzke et al., 2009; Nowak & Heinrichs, 2008), which makes it difficult to ascertain whether effect sizes may be diluted with nondisordered children. Finally, several previous meta-analyses have applied very stringent inclusion criteria, such as requiring them to be randomized controlled trials (e.g., Dretzke et al., 2009; Fletcher et al., 2011; Furlong et al., 2012; Thomas & Zimmer-Gembeck, 2007), requiring inclusion of fathers (Fletcher et al., 2011), or requiring the use of specific outcome measures (Dretzke et al., 2009).

Researchers have used various approaches to calculate effect sizes in previous meta-analyses of behavioural parenting interventions. Although a number of authors of meta-analyses reported having used random-effects models (i.e., de Graaf et al., 2008; Dretzke et al., 2009; Furlong et al., 2012; Kaminski et al., 2008; Lundahl et al., 2006; Maughan et al., 2005; Piquero et al., 2009; Reyno & McGrath, 2006; Thomas & Zimmer-Gembeck, 2007), one (McCart et al., 2006) reported using a fixed-effects model, and one reported using a mixed-effects model (Nowak & Heinrichs, 2008). The authors of three previous meta-analyses did not specify the type of model used in calculating effect sizes (Fletcher et al., 2011; Lundahl et al., 2008; Serketich & Dumas, 1996).

Three outcome variables have been examined in previous meta-analyses: parent behaviour, child behaviour, and parent adjustment. Consistent with the goals of behavioural parenting interventions, there is evidence that such programs produce changes in parent behaviours such as increased use of differential attention, praise, and time-out, and reduced use of spanking (Furlong et al., 2012; Kaminski et al., 2008; Lundahl et al., 2006). In turn, such interventions are often associated with a reduction in child behaviour problems, defined as

externalizing, aggressive, or antisocial behavior (Dretzke et al., 2009; Furlong et al., 2012; Kaminski et al., 2008; Lundahl et al., 2006; Lundahl, Tollefson, Risser, & Lovejoy, 2008; Maughan et al., 2005; McCart et al., 2006; Nowak & Heinrichs, 2008; Piquero et al., 2009; Serketich & Dumas, 1996). Behavioural parenting interventions also provide evidence of improvement in parent adjustment by lowering symptoms of depression, stress, anxiety, and increasing marital satisfaction (Furlong et al., 2012; McCart et al., 2006; Nowak & Heinrichs, 2008).

In general, previous meta-analyses have yielded heterogeneous study effects, indicating the presence of variables that may moderate these outcomes. A number of variables have been suggested in the past as possible moderators of behavioural parenting intervention outcomes, which can be broadly grouped into two categories: characteristics of the child, and characteristics of the study. In turn, I will examine the moderating variables that represent characteristics of the child, and then will examine those moderators that represent study characteristics.

Child moderator variables.

Various characteristics of the child may affect the outcome of behavioural parenting interventions, including characteristics of the child's problems (including the specific diagnosis, the presence of comorbidity, whether the problem was in the clinical range, or the severity of the problem) and the child's age. Previous results regarding child formal diagnostic status have varied depending on the type of outcome assessed. Parent behaviour outcomes have been found to be moderated by child formal diagnostic status when positive and negative parenting behaviours are examined separately (Furlong et al., 2012) but not when they are examined together (de Graaf et al., 2008; Lundahl et al., 2006). With respect to child behaviour outcomes, the findings differ depending on whether children with subclinical problems were included or

not. In one meta-analysis (Lundahl et al., 2006) that looked at clinical and subclinical conduct problems, children with clinical range symptoms evidenced greater effects than did those with nonclinical or mixed symptom levels. However, a meta-analysis that looked only at studies of children with clinically significant problems (Furlong et al., 2012) found no differences between children with more and less severe behaviour problems. Finally, only one meta-analysis (Furlong et al., 2012) examined the effects of child formal diagnostic status on parent adjustment outcomes and found evidence of higher effect sizes for those parents of children with more severe behaviour problems than those with less severe problems. No previous meta-analysis has examined the effects of the presence of different formal diagnostic categories (e.g., ADHD, ODD, CD) as well as comorbid problems on these outcomes for their differential effects existed for children with and without different diagnoses and clinical symptom levels.

Early parenting interventions were designed for parents of young children. I hypothesized that these interventions would be most potent when delivered to families of younger children before dysfunctional patterns of interaction became too entrenched. This leads to the hypothesis that child age will moderate outcomes with larger effects evident in families with younger children. However, in recent years many interventions have been adapted to meet the needs of families with older children. It is difficult to draw firm conclusions on the effects of child age on behavioural parenting intervention outcomes as child age has been examined in different ways, both in terms of the age ranges examined and also in terms of the size of age categories they used. Furthermore, it appears that analyses using age as a continuous variable yielded different results than did analyses using age categories. One meta-analysis which examined child age on a continuum found significant effects with respect to parent behaviour, child behaviour, and parent adjustment outcomes (Nowak & Heinrichs, 2008) whereas when child age was examined in

discrete categories, nonsignificant effects were found (de Graaf et al., 2008; Lundahl et al., 2006). It is my assertion that examining child age in categories may hide differential age effects as age groupings used may not be appropriate nor sensitive enough to capture effects that may be present. Therefore, I proposed to examine age both continuously as well as categorically (preschool age, school age, and high school age).

Methodological moderator variables.

Methodological characteristics, such as methodological quality, comparison group, and experimenter allegiance may also affect the outcomes of behavioural parenting interventions. In examining the size of treatment outcomes, it is important to take into account the methodological quality of the study. One previous meta-analysis (Nowak & Heinrichs, 2008) found greater treatment effects for parent and child behaviour outcomes in those studies with lower methodological quality. This may suggest that studies with lower methodological quality may be overestimating treatment effects (Hempel et al., 2011). To date, methodological quality has not been examined with respect to parent adjustment outcomes.

The nature of the study design also influences the size of the effect. Not surprisingly, psychotherapy yields lower effect sizes when compared to an active control group (i.e., treatment as usual, attention control) versus a passive control group (i.e., waitlist control, no treatment group; Baskin, Tierney, Minami, & Wampold, 2003; Jensen, 2003; Weisz et al., 2006). This has not been examined in previous meta-analyses of behavioural parenting interventions. However, it appears that combining these research designs might not provide an accurate portrait of the intervention's efficacy and may mask true treatment effects.

Allegiance bias refers to a difference in the findings based on the researcher's association with one of the interventions that is being tested (Leykin & Rubeis, 2009; Wampold 2001). In

particular, allegiance bias may be present if supporters of only one of two or more interventions being studied are involved in research, and significant differences (usually favouring the supported intervention) are found (Staines & Cleland, 2007). However, allegiance bias has also been found to be a concern in research in which an intervention is compared to a control group, as it is likely that the favoured condition in such studies would be the implemented intervention (Shadish & Baldwin, 2005). In both cases, this may therefore lead to an inflation of the effect size. Involvement of a program's originator in evaluations of the program has been associated with positively skewed results (Eisner, 2009). Recent research in psychotherapy has demonstrated some evidence of research allegiance being associated with better treatment outcomes (Munder, Gerger, Trelle, & Barth, 2011). However, the presence of allegiance bias has not been examined as of yet in a meta-analysis of behavioural parenting interventions, although a recent meta-analysis of child and youth interventions (Miller, Wampold, & Varhely, 2008), which included behavioural parenting interventions as one of many treatments, did report evidence of allegiance bias.

Need For Another Meta-analysis

Despite the number of previous meta-analyses available on this subject, there are several important reasons why a new meta-analysis of this area is necessary. First, the body of literature in this area has continued to grow since the search for last comprehensive meta-analysis of behavioural parenting interventions completed in 2008 (Piquero et al., 2009). It is notable that two meta-analyses have been completed since but one of these required father involvement and examined only randomized controlled trials (RCTs) of Triple P (Fletcher et al., 2011) and the other was restricted to RCTs and quasi-randomized trials of group-based programs (Furlong et al., 2012).

Second, this meta-analysis was designed to examine the moderating effects of child characteristics such as child age and child formal diagnostic status, including diagnoses of ADHD, ODD, CD, and comorbid problems. It was also designed to examine those moderating effects of study characteristics such as methodological quality, publication status, comparison group, and researcher allegiance in behavioural parenting intervention outcomes. Although several of these variables (i.e., child age, child formal diagnostic status, methodological quality, publication status) have been examined as possible moderators in at least one previous meta-analysis of behavioural parenting interventions (e.g., Lundahl et al., 2006; Maughan et al., 2005; McCart et al., 2006; Nowak & Heinrichs, 2008; Piquero et al., 2009; Serketich & Dumas, 1996), one variable (i.e., experimenter allegiance) has not been previously examined as a moderator. It is important to note that although each of these variables has been previously examined, not all have been examined for all three outcomes that were studied in the present meta-analysis.

Furthermore, despite an increasing trend towards the use of a random effects model to conduct meta-analyses, and greater acceptance of this model compared to fixed effects, this procedure has not been systematically used. In fact, there has been a great deal of variability in the procedures used in previous meta-analyses of behavioural parenting interventions. A number of meta-analyses have been conducted using a random effects model, however a number are unclear regarding the type of model which was used (Fletcher et al., 2011; Lundahl et al., 2008; Serketich & Dumas, 1996). At least one previous meta-analysis used a mixed-effects model (Nowak & Heinrichs, 2008) and yet another reported having used a fixed-effects model (McCart et al., 2006). In this research I used a random effects model.

In addition, as debate continues regarding the possible merits of including grey literature in recent years, it is important to conduct a meta-analysis which thoroughly examines that

literature. There has been limited inclusion of grey literature located using basic search engines, despite its identification by researchers as an easier and quicker way to identify grey literature (e.g., Coad et al., 2006). Two notable exceptions are Nowak and Heinrichs' (2008) meta-analysis, in which GoogleScholar and search engines were searched, as well as Furlong et al. (2012) who used Google to search parenting websites. Notably, both these meta-analyses were quite limited in their scope, and neither addressed the impact of including grey literature on their results, with Furlong et al. (2012) only briefly noting the possibility that some unpublished studies may have not been located.

Many previous meta-analyses (e.g., de Graaf et al., 2008; Dretzke et al., 2009; Lundahl et al., 2006; McCart et al., 2006; Nowak & Heinrichs, 2008; Thomas & Zimmer-Gembeck, 2007) included both clinically significant as well as less severe behavioural problems. Surprisingly, several did not report whether children in their samples had clinically significant or diagnosed behavioural problems (i.e., Fletcher et al., 2011; Lundahl et al., 2008; Maughan et al., 2005; Piquero et al., 2009; Reyno & McGrath, 2006; Serketich & Dumas, 1996).

In addition, the use of search strategies to identify published literature varied greatly in previous meta-analyses. Searches of electronic databases were universally used, although a great deal of variability was found with regards to how many databases were searched (one to 20, with many searching 2 databases) as well as which databases were searched. The PsycINFO database was the only one that was universally searched, with Medline and Educational Resources Information Centre (ERIC) being the second most searched databases. Most meta-analyses reported searching the reference lists of previous reviews or relevant studies (i.e., de Graaf et al., 2008; Furlong et al., 2012; Maughan et al., 2005; McCart et al., 2006; Nowak & Heinrichs, 2008; Piquero et al., 2009; Reyno & McGrath, 2006; Serketich & Dumas, 1996; Thomas &

Zimmer-Gembeck, 2007). Few reported hand searching relevant journals (i.e., Furlong et al., 2012; Maughan et al., 2005; Nowak & Heinrichs, 2008; Piquero et al., 2009) or contacting researchers and/or trial authors (i.e., de Graaf et al., 2008; Furlong et al., 2012; Reyno & McGrath, 2006; Thomas & Zimmer-Gembeck, 2007).

The number of employed search terms to identify research ranged from three to 50 terms. However, it is important to note that the review which included 50 terms (Kaminski et al., 2008) included 28 search terms related to specific program outcomes in which they were interested in examining. Following this, the greatest number of search terms used was 16.

The Current Study

The present study is a meta-analysis of behavioural parenting interventions for the treatment of disruptive behaviour disorders. Whereas many previous meta-analyses focused on varying levels of problems, from not clinically significant to diagnosed, I chose to examine a more homogenous group of children with children whose parents reported behavior problems above a clinical threshold or who had been diagnosed as having a DBD in order to determine the outcome of parent training in families of children with clinically significant problems. It is important to note that prevention programs were not examined as part of this study as the mechanism of preventing behaviour problems in at-risk children may be different than reducing existing diagnosed or clinically significant behaviour problems. The meta-analysis was designed to: 1. include studies reported since the most recent general analyses (i.e., after 2008) in addition to earlier studies, 2. use a great variety of search terms in order to identify the greatest number of relevant articles and unpublished works, and 3. go beyond the traditional search strategies for accessing grey literature by using basic search engines such as Google and Yahoo, which have only been used once previously. In addition, the meta-analysis included studies that expand upon

previously examined age ranges in meta-analyses with a range from toddlerhood to adolescence in order to be able to examine if critical ages exist in which behavioural parenting interventions are most effective.

As a first step I examined the number of studies compared to the numbers reported in previous meta-analyses. Next, I examined effect sizes in terms of three outcome variables: parent behaviour, child behaviour, as well as parent adjustment. These outcomes have all been examined in previous meta-analyses. Consistent with the goals of behavioural parenting interventions, there is evidence that such programs produce changes in parent behaviours such as increased use of differential attention, praise, and time-out, and reduced use of spanking (Kaminski et al., 2008; Lundahl et al., 2006). In turn, such interventions are often associated with a reduction in child behaviour problems (Dretzke et al., 2009; Kaminski et al., 2008; Lundahl et al., 2006; Lundahl, Tollefson, Risser, & Lovejoy, 2008; Maughan et al., 2005; McCart et al., 2006; Nowak & Heinrichs, 2008; Piquero et al., 2009; Serketich & Dumas, 1996). Behavioral parenting interventions also yield improvement in parent adjustment by lowering symptoms of depression, stress, anxiety, and increasing marital satisfaction (McCart et al., 2006; Nowak & Heinrichs, 2008).

Further, I examined the moderating effects of child variables including: presence of a formal diagnosis of ADHD, ODD, CD, or comorbid diagnosis, and child age. Methodological variables examined were overall quality, publication status, comparison group, and experimenter allegiance. The research hypotheses are presented in Table 2. With regards to moderation of outcomes by child variables, I predicted that parents of children with a formally diagnosed DBD would show greater improvement in parent adjustment than would parents of children without formal diagnoses. I also predicted significant age effects when examining child age continuously

such that higher effects would be found for younger children, but not when examining this variable categorically. In examining methodological variables as possible moderators, I predicted an inverse relationship between methodological quality and both parent and child behaviour outcome variables, indicating a possible overestimation of effect sizes in lower quality studies, as has been previously found. I predicted that effect sizes for behavioural parenting interventions would be moderated by the type of control group which was employed with higher effect sizes in studies in which a behavioural parenting intervention was compared to a no-treatment control group than when compared to an active comparison group. Finally, I predicted that studies in which researchers had an allegiance to one intervention being offered over the other would report greater treatment effects than those studies in which the researchers did not have an allegiance to the intervention.

In addition, the meta-analysis was designed to address research questions with respect to the utility of including grey literature in the present meta-analysis. In particular, I examined how many grey studies were obtained that met inclusion criteria for the present meta-analysis. I examined how many of those studies had already been identified in the published literature. I also explored the question of whether effect sizes differed between the grey literature and published literature and also examined whether differences existed in the methodological quality of published studies and unpublished studies.

Roadmap for this thesis.

The two manuscripts that follow examine different foci of the meta-analysis that was conducted for this dissertation. Manuscript one presents the main findings in terms of effect sizes for three types of outcome variables, as well as the results of the subsequent moderator analyses

that were performed. Manuscript two examines the influence of the decision to include grey literature in the present meta-analysis, and discussed the benefits and challenges to its inclusion.

Manuscript One

A Meta-Analytic Examination of Behavioral Parenting Interventions in the Treatment
of Children's Behavior Problems*

Sarah Bellefontaine and Catherine M. Lee

University of Ottawa

*This is a revised manuscript that has been submitted to *Journal of Abnormal Child Psychology*.

Abstract

Behavioral parenting interventions based on social learning theory are designed to enhance relationships between parents and their children through the promotion of positive parenting practices. The present meta-analysis synthesized the results of 42 published and 6 unpublished treatment studies of behavioral parenting interventions conducted between 1966 and 2011. Overall efficacy was examined, as were the effects of six potential moderator variables in two categories: child characteristics and methodological characteristics. Results confirm the findings of previous meta-analyses in demonstrating robust positive effects of behavioral parenting interventions on parent behavior, child behavior, and parent adjustment. Effects were homogeneous for all outcome variables in comparisons between a behavioral parenting intervention and an active comparison group, as well as for child behavior outcomes in comparisons between behavioral parenting interventions and a no-treatment group. Child age was a significant moderator of parent behavior and parent adjustment outcomes in comparisons between a behavioral parenting intervention and no-treatment group; similarly a diagnosis of Attention-Deficit/Hyperactivity Disorder was a moderator of parent adjustment effects. The results underline the importance of consistent reporting practices in the field and their impact on the availability of evidence for policy-makers, administrators, researchers, clinicians and consumers.

Keywords: Parent Training; Meta-analysis; Externalizing Behaviors; Disruptive Behavior Disorders

A Meta-Analytic Examination of Behavioral Parenting Interventions in the Treatment of Children's Behavior Problems

Behavioral parenting interventions are designed to enhance parent-child relationships by teaching parents to increase the number of positive interactions with their children, provide rewards for appropriate behavior, set clear limits with their children, reduce their use of commands, and become more consistent in their discipline (Taylor & Biglan, 1998). These programs are typically delivered by trained facilitators using treatment manuals and materials. Behavioral approaches allow parents to set their own parenting objectives; using role-play and homework to facilitate generalization of skills. Disruptive behavior refers to problems in childhood that are characterized by acting out, out-of-control or hard to manage behaviours, as well as delinquency (Maughan, Christiansen, Jenson, Olympia, & Clark, 2005).

Several meta-analyses have examined the impact of behavioral parenting interventions on either the prevention or treatment of disruptive behavior in children: Table 1 provides a summary of the meta-analyses, their inclusion criteria, findings, and the moderators they examined. Most studies examined children with both clinical and subclinical behavior problems, with only one meta-analysis (Furlong et al., 2012) reporting on studies comprised only of children with clinically significant problems. Given the diverse inclusion criteria and time periods covered, it is not surprising that the numbers of studies included ranges from 13 to 79. The two most recent meta-analyses, which conducted searches in 2011, were more limited : Furlong et al. (2012) focused solely on randomized controlled trials (RCTs) and quasi-controlled trials of group parenting program, and Fletcher et al. (2011) examined only Triple P. The last comprehensive meta-analysis (Piquero et al., 2009) completed a literature search in 2008. Although several meta-analyses focused exclusively on services offered to very young children, some covered a

wide range of ages, from 0 to 18. Furthermore moderator variables were examined in all but three meta-analyses, including child and family characteristics, treatment characteristics, and methodological characteristics. Results of all of these meta-analyses provide evidence that behavioral parenting interventions are efficacious in producing several changes in the families of children with disruptive behavior. First, consistent with their goals, there is evidence they produce changes in parent behaviors such as increased use of differential attention, praise, and time-out, and reduced use of spanking (Furlong et al., 2012; Kaminski, Valle, Filene, & Boyle, 2008; Lundahl, Risser, & Lovejoy, 2006). Second, they show positive effects in terms of reducing child behaviour problems such as aggression, acting out, antisocial behaviour, and non-compliance (Dretzke et al., 2009; Furlong et al., 2012; Kaminski et al., 2008; Lundahl et al., 2006; Lundahl, Tollefson, Risser, & Lovejoy, 2008; Maughan et al., 2005; McCart, Priester, Davies, & Azen, 2006; Nowak & Heinrichs, 2008; Piquero, Farrington, Welsh, Tremblay, & Jennings, 2009; Serketich & Dumas, 1996). Third, they provide evidence that behavioral parenting interventions yield improvement in parent adjustment such as lowering symptoms of depression, stress, and anxiety, and increasing marital satisfaction (Furlong et al., 2012; McCart et al., 2006; Nowak & Heinrichs, 2008). Interventions produced weighted effect sizes of $d = .30$ to $.77$ for each of these outcomes (e.g., Lundahl et al., 2006; Maughan et al., 2005; McCart et al., 2006; Serketich & Dumas, 1996; Piquero et al., 2009).

However, these previous meta-analyses have several limitations. Notably, fixed- or mixed-effects analyses were used in several reports (McCart et al., 2006; Nowak & Heinrichs, 2008). Tests of homogeneity were not conducted prior to examination of moderators (Dretzke et al., 2009; Reyno & McGrath, 2006; Serketich & Dumas, 1996; Thomas & Zimmer-Gembeck, 2007). Most of these previous meta-analysis included children with subclinical problem

behaviors (de Graaf et al., 2008; Dretzke et al., 2009; Lundahl et al., 2006; McCart et al., 2006; Nowak & Heinrichs, 2008; Thomas & Zimmer-Gembeck, 2007) or did not report whether children with subclinical problems were included (Fletcher et al., 2011; Lundahl et al., 2008; Maughan et al., 2005; Piquero et al., 2009; Reyno & McGrath, 2006; Serketich & Dumas, 1996). Some previous studies limited their review to a small number of parenting programs (Furlong et al., 2012; Nowak & Heinrichs, 2008) or design features (Furlong et al., 2012). Therefore, we set out to conduct a comprehensive meta-analysis of behavioral parenting interventions to address these concerns and to get a better picture of what factors influence such interventions for children with disruptive behavior problems.

In addition to updating the body of literature on behavioral parenting interventions to include studies published until December 2011, the current meta-analysis was designed to expand upon current knowledge in several important ways. The present meta-analysis focused on behavioral parenting services for parents of children with disruptive behavior who met diagnostic criteria for a Disruptive Behavior Disorder (DBD) or who scored above a clinical cut-off on a measure such as the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999) or the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001). Studies of children with other disabilities were also included provided that samples also met the problem severity criterion. In addition, many of the most recent meta-analyses examined subsets of this body of literature, with restrictions on either child age, programs examined, or year of publication of included studies. This meta-analysis was designed to present an overall picture of all behavioral parenting interventions from toddlerhood to adolescence, with fewer restrictions regarding year of publication and no restrictions as to the types of programs or specific measures required for inclusion. The only inclusion criterion with respect to design was that the study had to include a

no-treatment, waitlist, or alternative treatment comparison group. Instead, we opted to examine the methodological quality of all studies and determine whether this variable accounts for significant variability in the literature. One novel moderator variable (experimenter allegiance) was examined in response to concerns in the literature about its effects on research outcomes. We also included unpublished literature which extends beyond dissertations and theses in order to present the most complete picture of this body of literature possible, and directly addressed the impact of including such literature on research outcomes. We used a number of terms to access as many potentially relevant sources as possible. Finally, in contrast to previous meta-analyses which have varied in the procedures they have used, we employed the more conservative random-effects model, conducted a-priori tests of heterogeneity, using moderator analyses when appropriate based on a finding of significant heterogeneity.

Moderator Variables

In previous meta-analyses researchers have examined whether treatment effects on parent behavior, child behavior and parent adjustment are homogeneous, or whether they are moderated by child characteristics or methodological characteristics of the study. These moderator analyses address the question of what intervention works for whom.

Child characteristics.

In examining whether parenting interventions have larger effects for families of children with clinically significant problems than for families of children with subclinical difficulties, the results depend on the type of outcome assessed. Child clinical status has not been found to moderate the effect on parent behaviour outcomes (de Graaf et al., 2008; Lundahl et al., 2006) unless positive and negative parenting behaviour are examined separately (Furlong et al., 2012). However, in studies looking at child behavior outcomes, findings are contradictory with one

meta-analysis reporting that children with clinical range symptoms evidenced greater effect sizes than did those with either nonclinical symptom levels or a mix of clinical and nonclinical symptom levels (Lundahl et al., 2006) and another (Furlong et al., 2012) finding nonsignificant differences between children with more and less severe conduct problems. Notably, Furlong et al. (2012) only examined children with clinically significant problems, whereas Lundahl et al. (2006) looked at children with both clinically significant and subclinical problems, which may account for the discrepancy in these findings. One meta-analysis (Furlong et al., 2012) found that parents of children with more severe problems evidenced greater effects in terms of parent adjustment than did parents of children with less severe problems. However, none of the present outcomes have been examined with regards to differential effects of different diagnoses, in addition to looking at subclinical problems and comorbid problems. We examined the differential effects of formal diagnosis of Attention-Deficit/Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD), Conduct Disorder (CD), or two comorbid Disruptive Behavior Disorders (DBDs) and predicted significantly greater effects on parent adjustment outcomes for children with diagnosed problems of ADHD, CD, or ODD, or comorbid disorders than for those children who did not have formal diagnoses.

Studies of behavioral parenting interventions have examined samples that vary both in terms of child age and in the size of the age range examined, rendering examination of the moderating effects of child age a challenge. In addition, it is possible that the relationship between child age and each of the three outcome variables being examined (i.e., parent behavior, child behavior, and parent adjustment) might be accounted for by the way in which age has been examined. A meta-analysis which examined child age as a continuous variable found significant effects on all three outcomes (Nowak & Heinrichs, 2008) whereas meta-analyses which have

examined child age categorically have generally found nonsignificant effects (de Graaf et al., 2008; Lundahl et al., 2006). Thus it appears as though examining age categorically may mask differential age effects by artificially imposing age groupings which may not be sensitive enough to capture the effects which are present. In order to test this, and examine these discrepancies more directly, age was looked at both categorically (preschool age, school age, and high school age) as well as continuously.

Methodological characteristics.

In examining effect sizes, it is important to take into account the methodological quality of the research. Nowak and Heinrichs (2008), for example, found that studies of poorer methodological quality reported higher effect sizes for parent behavior and child behavior than did studies with greater methodological quality, possibly leading to an overestimation of effects. The possible moderating effect of methodological quality on parent adjustment outcomes has not been examined in this literature. Therefore, we examined the possible moderating effect of methodological quality on each of these outcomes.

In addition we examined the possible moderating effects of two other methodological characteristics: experimenter allegiance and type of comparison group. Within the adult psychotherapy literature, research (e.g., Eisner, 2009; Petrosino & Soydan, 2005) has suggested that when program developers are involved in testing the intervention program they developed, more positive outcomes are found. This, in turn, can lead to an overestimation of effects, and thus a distortion of findings (Leykin & DeRubeis, 2009). Previous studies of psychotherapy research have found evidence of more positive outcomes when the investigator has allegiance to one intervention or condition over another (e.g., Paley & Shapiro, 2002). This experimenter allegiance hypothesis has not been tested in a meta-analysis of parenting programs, although a

recent meta-analysis of child and youth treatment which included behavioral parenting interventions reported evidence of researcher allegiance effects (Miller, Wampold, & Varhely, 2008). We tested the hypothesis that studies in which the researchers have an allegiance to the training program would report greater treatment effects.

Research has found that studies which compare an intervention to a passive comparison group (i.e., waitlist control or no treatment group) yield greater effect sizes than do those in which an intervention is compared to an active comparison group (i.e., treatment as usual, or another treatment; Jensen, 2003; Weisz, McCarty, & Valeri, 2006). Thus, combining these different research designs may not provide an accurate portrait of the efficacy of an intervention, and may dilute or inflate overall effect sizes. With this in mind, we decided to examine effect sizes separately for no-treatment and active comparison groups. In the present study, active comparison groups consisted of social skills training, attention control, monitoring, support, treatment as usual, parent counselling and support, collaborative problem-solving, as well as parental stress and problem solving. We predicted that those studies in which a behavioral parenting intervention was compared to a passive control group would evidence greater effect sizes than those compared to an active comparison group.

Grey literature.

Grey literature refers to print or electronic literature that is produced by government, academia, business and industry, and is not controlled by commercial publishers (Auger, 1998). Grey literature includes unpublished studies, conference abstracts, conference proceedings, book chapters, government and agency reports, and unpublished doctoral dissertations. Reviewers who have not included grey literature in their meta-analyses have generally cited the difficulty in obtaining unpublished studies as the main reason for not doing so (e.g., McCart et al., 2006). In

general, the unpublished studies reported in meta-analyses have consisted largely of dissertations and theses (e.g., Piquero et al., 2009; Serketich & Dumas, 1996).

In the field of adult psychotherapy, some authors have found that meta-analytic findings solely based on published studies may over-estimate treatment effect sizes (Staines & Cleland, 2007). Staines and Cleland (2007), reported that population effect sizes for unpublished psychotherapy studies were only 70 to 90 percent as large as the effect sizes of published studies. To respond to this concern, Tricco et al. (2008) suggested that unpublished studies should be included in meta-analyses so as to minimize bias. However, many meta-analysts remain concerned about the quality of grey literature and choose not to include it despite research demonstrating that certain sources of grey literature such as dissertations are of high quality (McLeod & Weisz, 2004).

The meta-analyses in Table 1 that reported mean effect sizes based solely on published studies (de Graaf et al., 2008; Dretzke et al., 2009; Lundahl et al., 2006; McCart et al., 2006; Reyno & McGrath, 2006; Thomas & Zimmer-Gembeck, 2007) found higher mean effect sizes ($d = .42-1.31$) than did meta-analyses which included unpublished studies ($d = .30-.35$; Maughan et al., 2005; Piquero et al., 2009). The only exception to this pattern was an early meta-analysis (Serketich & Dumas, 1996) in which the unweighted effect size ($d = .86$) may have been inflated by the restrictive criteria used to identify studies, the small number of studies included in the analysis ($N = 26$), as well as the fact that only 4 unpublished studies were included.

The current study was designed to examine the overall impact of grey literature on the meta-analytic outcomes of behavioral parenting intervention studies. In particular, we examined whether the conclusions would be affected by the inclusion of unpublished, grey literature.

Method

Search Strategy

This meta-analysis included studies published in English from 1966 until December 2011, as well as the grey literature from 1966 until December 2011. A comprehensive search of PsycINFO, PubMed, Scopus, and ERIC was used to identify relevant studies published in peer-reviewed journals and book chapters. Hand searches of key journals were undertaken to identify relevant articles which may have been overlooked by other search strategies (see Appendix A for a full list of key journals). In addition, the reference lists of relevant articles, reviews and chapters were examined to find any other studies which were not identified by other means.

Search terms used in identifying relevant articles included terms related to intervention types and program names (*behavio(u)ral family interaction; behavio(u)r modification; behavio(u)ral parent training; behavio(u)r/behavior therapy; contingency management; parent-child interaction; Parent Effectiveness Training; Parent Management Training; parent training; Community Parent Education; COPE; Defiant Children; Incredible Years; Parent-Child Interaction Therapy; PCIT; Positive Parenting Program; Triple P*), as well as disruptive behavior (*ADD; ADHD; aggressive; antisocial; Attention Deficit Disorder; Attention Deficit Hyperactivity Disorder; behavio(u)r disorder; CD; conduct disorder; conduct problem; DBD; defiant; disruptive; Disruptive Behavio(u)r Disorder; externalizing problem; hyperkinetic; ODD; Oppositional Defiant Disorder*) and different combinations and permutations of these search terms (see Appendix B for a list of search terms). They included all those that had been used in previous meta-analyses, as well as several additional terms in order to access the greatest number of relevant articles.

Identifying grey literature.

PsycINFO was used to identify dissertations dating back to 1995 and Proquest Dissertations and Theses was also searched to identify unpublished dissertations which were not available using PsycINFO as well as to obtain full-text versions of dissertations found through other search strategies. Further unpublished studies were identified using Scopus, a database which includes both conference proceedings and trade publications across a number of different disciplines. Requests to researchers who had published in the field asking them to forward unpublished studies or data yielded no further studies.

In addition, online searches were performed using the Google and Yahoo search engines to identify further unpublished studies. Using these search engines, the first 500 search results were examined. This strategy was chosen to ensure the largest possible number of relevant identified resources within a manageable number of search results. To maximize the relevance of searches performed on the internet search engines Google and Yahoo, only combinations of search terms (e.g., *parent training* and *disruptive*) were used. Initial searches were performed in 2009. Updates to searches were performed between August and December 2011.

Inclusion Criteria

To be included in the meta-analysis, studies had to meet a number of criteria (refer to Appendix C for study inclusion coding manual; see Appendix D for coding sheet). First, it had to examine a behavioral parenting intervention in which parents practised skills (e.g., praise, differential reinforcement, time-out, response cost), in groups or individually, using methods such as psychoeducation, role play, using videotapes and telephone or internet consultation. To permit comparison of the results of the present meta-analysis with previous meta-analyses the articles were required to have examined the parents of children between the ages of 2 and 18 years. Eligible studies had to include a comparison group: either a waitlist control group, no

treatment group, or a group in which another treatment was employed. A sample size of at least 15 participants in each group was required to ensure adequate power to detect meaningful differences between groups (Kazdin & Bass, 1989). Studies had to describe services offered to treat clinically significant problems in the child, indicated either by meeting diagnostic criteria established by DSM-III, DSM-IV, or DSM-IV-TR for a Disruptive Behavior Disorder, or by scoring above a clinical cut-off on at least one psychometrically sound behavioral rating scale or clinical interview examining externalizing behavior. Finally, studies were required to have included measures of at least one of the outcome variables being examined: child behavior, parent behavior, or parental adjustment.

Data Extraction and Coding

Identifying multiple reports from the same data-set.

Often a single study yields multiple publications, such as when overall findings are presented, and follow-up data are subsequently published. This also may occur when results of a subset of participants are reported (e.g., data from a study of children with Oppositional Defiant Disorder [ODD] are reported, and then another report describes the data of a subsample of those participants with intellectual deficits). In meta-analysis, it is essential that each sample is counted only once, so therefore it was necessary to determine whether articles were based on the same sample and to identify the most complete report.

The detection heuristic proposed by Wood (2008) was adapted to psychological research by the authors (see Appendix E for coding manual and Appendix F for coding sheet). Studies were examined on nine criteria to establish whether they referred to the same sample: whether they shared one or more authors, referenced another report, described the same treatment, and reported the same recruitment and referral procedures, time period, age range of children,

response rates, assessment tools, and parent ages. If at least seven of these criteria were met, then the study was coded as a multiple, and the report with the most comprehensive data set was included in the meta-analysis.

Calculation of study effect sizes.

Means, standard deviations, and group sample sizes for each group, or between-group effect sizes and sample sizes, were recorded to calculate the effect size for each study. For studies that reported data for more than one measure for each outcome examined, all effect size data were recorded. Effect sizes were weighted according to sample sizes. These were, in turn, averaged to calculate the overall effect size of behavioral parenting interventions for each of the outcomes examined: child externalizing behavior, parent behavior, and parent adjustment.

Methodological quality.

The methodological quality of included studies was examined to determine whether the quality of such studies exerts an influence on the magnitude of effects which are found. Previously used scales apply only to randomized controlled trials (e.g., Jadad et al., 1996) or apply to health care interventions, with a great focus on reporting of information (e.g., Downs & Black, 1998). Two raters determined methodological quality on a scale from 0 to 7 (see Appendix G for coding manual and Appendix H for coding sheet), with higher scores representing greater methodological quality. Two criteria were examined to determine the methodological quality, with the total score obtained on these two criteria resulting in the methodological quality score.

Studies were first examined for their group assignment procedures. Studies that used random group assignment were assigned a score of 2, studies that ensured group equivalency but did not randomly assign were assigned a score of 1. In the case that no information was provided

on assignment procedures and group equivalence, a rating of 0 was given. Nine studies were given a score of 0, two studies were given a score of 1, and 37 studies were given a score of 2.

Secondly, psychometric soundness of measures was determined for each measure examined on a scale of 0 to 5 using a coding scheme based on the criteria proposed by Hunsley and Mash (2008). Measures were identified as being psychometrically sound if they fulfilled 5 criteria: (1) norms were based on large, appropriate clinical samples, (2) an internal consistency of at least $\alpha = .70$, (3) test-retest reliability correlations of .70 or greater, as well as adequate evidence of (4) content and (5) construct validity. As reporting of psychometric data varied significantly by author and publication, if the description of the measure provided in the article itself did not provide adequate information for determination of the adequacy of these measures of psychometric soundness, the manual for the specific measures was consulted. To determine the overall psychometric soundness of the measures in the study, the average psychometric soundness of all relevant measures was calculated. The full scoring manual is available from the first author.

Moderator variable coding.

Studies that met inclusion criteria, were not multiples, and provided enough information to calculate effect sizes were coded with respect to 6 moderator variables in two categories (see Appendix I for coding manual; refer to Appendix J for coding sheet). Child variables included child clinical status and comorbidity, child age. Child formal diagnostic status and comorbidity were coded for the presence or absence of diagnoses of *ADHD*, *ODD*, *CD* and/or *comorbid conditions* in the sample. Child age was coded as the mean age of children involved in the study, and was then also grouped at into age categories of *preschool*, *school-age*, and *high school age* so that it could be examined both categorically and continuously. Methodological characteristics

included type of comparison group, experimenter allegiance, and peer-reviewed status.

Methodological quality was coded on a scale of 0 to 7 based on the criteria described above. The type of comparison group was coded as *no-treatment comparison* for those studies in which a behavioural parenting program was compared to a waitlist or no-treatment group, and was coded as *active comparison* for those studies which compared a behavioural parenting intervention to treatment as usual or another type of intervention. Experimenter allegiance was coded positively if one or more of the authors of the research was involved in the development of the behavioural parenting program which was being tested. For publication status, studies were coded as published or unpublished based on whether they had been previously published in a peer-reviewed journal.

Coding Reliability Procedures

Inclusion/exclusion coding.

Two raters coded studies to determine whether they met inclusion criteria. Raters were undergraduate- and graduate-level research assistants, trained by the first author through joint coding of studies until a criterion of at least $k=.70$ was reached. To avoid coder drift, frequent checks were conducted. In addition, the second author coded 20% of the articles. Any discrepancies were resolved through joint re-examination of the study methodology. Kappa for study inclusion/exclusion coding was .91 between raters and the first author.

Multiple articles on the same data-set coding.

All studies identified for inclusion were screened and flagged by the first author to examine further for evidence of them being multiple articles from a single data-set. All articles were then coded by both authors working together in order to determine if they were indeed

referring to the same data-set. Kappa was therefore not calculated. Disagreements were solved by jointly re-examining the study methodology.

Effect size coding.

All study effect size data for each of the three outcomes (e.g., parent behavior, parent adjustment, and child behavior) were extracted by two raters (first and second authors) coding independently. Values extracted were *ns* for each group as well as the between-groups *d*, or the *M*, *SD*, and *n* for the intervention and control groups. Disagreements were addressed by having the raters jointly re-examine the study methodology and reach consensus.

Moderator variable coding.

Moderator variables were extracted by two raters (the first author and a trained undergraduate research assistant) for each report. Kappas for coded variables ranged from .55 to 1.00. Disagreements were resolved through a joint review of the study methodology. If consensus between the two raters could not be reached regarding a study, the issue was discussed with the second author and an executive decision was made as to the disagreement at hand, in consultation with the first author. Such decisions were made by referring to the research article or report in question.

Results

Our search identified 1746 potentially relevant reports, of which we were able to obtain full-text versions of 1518 for evaluation. Of those, reasons for exclusion included: reviews or theoretical papers ($n = 239$), not examining behavioral parenting interventions ($n = 417$), having no comparison group ($n = 422$), inadequate sample size ($n = 180$), children outside the 2-18 range ($n = 12$), not examining clinically significant child behavior problems ($n = 116$), outcomes other than those being examined ($n = 13$), a behavioral parenting intervention combined with

another intervention ($n = 20$), and other ($n = 4$). Our search yielded 74 articles and 21 unpublished documents that met initial inclusion criteria. We retained 42 published articles and six unpublished documents (5 dissertations and 1 unpublished report) after eliminating multiple reports and articles lacking necessary details for calculating effect size data. Table 2 presents a summary of all studies included in the meta-analysis.

Of those 48 studies that were included in the present meta-analysis, 19 (39.6%) had not been included in previous meta-analyses. Of those 19 new references, 2 had been published after the most recent review (Furlong et al., 2012). The remaining 17 studies were not included in previous meta-analyses due to not meeting inclusion criteria of the meta-analysis (i.e., child age outside the range being examined, comparison group was other than required by the meta-analytic criteria, or reporting nonsignificant results).

The total number of participants was 1762 across all behavioral parenting intervention groups and 1416 for comparison groups. A random-effects model using weighted effect sizes was used to calculate Hedge's g for parent reports of parent behavior, child behavior, and parent adjustment along with 95% Confidence Intervals (CI) for each. As the most commonly reported outcomes were based on parent-report, only those outcomes were used. Within studies, effect sizes of multiple parent-report measures of parent behavior, child behavior, and parent adjustment were aggregated into a single study effect size for each of the aforementioned outcomes by calculating the mean of each individual effect sizes using the Comprehensive Meta-Analysis Program (Borenstein, Hedges, Higgins, & Rothstein, 2005). If data from both parents were available, effect sizes were aggregated into a single parent effect size for each measure. Overall effects were calculated separately for studies that compared behavioral parenting

interventions to a waitlist or no treatment control group, and for those that compared behavioral parenting interventions to an active treatment group.

Parent Behavior

Active treatment comparison. Four studies compared behavioral parenting interventions to an active treatment comparison group. The overall effect size for these studies is $g = .20$ (95% CI = .01 - .39). The calculated fail-safe N indicated that 2 studies with nonsignificant results would have to be added in order to reduce the overall effect size to below a statistically significant level. A test of heterogeneity indicated that the sample was homogeneous ($Q[3] = 2.14, p = .544$) and so no moderators were examined.

No treatment comparison. The 24 studies that compared a behavioral parenting intervention to a no treatment comparison and reported data for parent behavior measures (e.g., GBPS, PS, PPI) yielded an overall effect size of $g = .71$ with a 95% CI of .51 to .92. A calculation of the fail-safe N indicated that 991 studies with nonsignificant results would have to be added in order to reduce the overall effect size to a level below statistical significance. A test of heterogeneity indicated significant variability in the effect sizes ($Q[23] = 85.39, p < .001; I^2 = 73\%$), making it possible to examine for moderating variables.

Moderators were examined using the Comprehensive Meta-Analysis Program to perform mixed effects regression for continuous variables and by conducting t -tests between different levels of categorical moderator variables. Table 3 presents a summary of the moderators examined, the reliability with which they were coded, as well as whether it was found to have a significant effect on each of the outcome variables for which moderators could be examined.

Child and parent characteristics as moderators of parent behavior outcomes.

When behavioral parenting interventions were compared to a no treatment control group, child formal diagnostic status and comorbidity did not moderate parent behavior outcomes. Examined categorically, child age also did not moderate parent behavior outcomes. However, when examined continuously using meta-regression age was found to be a significant moderator, such that as child age increases (up to age 9), effect sizes increase ($\text{Tau}^2 = .20$, $\text{SE} = .09$).

Methodological characteristics as moderators of parent behavior outcomes.

Effect sizes for parent behavior varied significantly between those studies that were unpublished ($g = .19$; 95% CI = $-.12-.49$) and those that were published ($g = .58$; 95% CI = $.49-.68$; $t[26] = 2.64$, $p = .014$). Independent groups t -tests revealed that methodological quality, experimenter allegiance, and publication status were not significant moderators of parent behavior.

Child Behavior

Active treatment comparison. Child behavior outcomes were reported in 13 studies where behavioral parenting interventions were compared to an active treatment comparison. The overall effect size for these studies is $g = .27$ (95% CI = $.14 - .40$). The calculated fail-safe N indicated that 53 studies with nonsignificant results would have to be added in order to reduce the overall effect size to below a statistically significant level. A test of heterogeneity indicated that the sample was homogeneous ($Q[12] = 9.33$, $p = .675$) and so no moderators were examined.

No treatment comparison. Across all studies, published and unpublished, child behavior measures (e.g., ECBI, CBCL) yielded an effect size of $g = .58$ with a 95% CI of $.49$ to $.67$. A fail-safe N was calculated, revealing that 1355 studies with nonsignificant results would have to be added in order to reduce the overall effect size to a level below statistical significance. A test

of heterogeneity was performed on the 13 effect sizes, and revealed significant homogeneity ($Q[32] = 27.75, p = .682$). Therefore, no moderator analyses were conducted.

Parent Adjustment

Active treatment comparison. Nine studies examined parent adjustment outcomes in studies where behavioral parenting interventions were compared to an active treatment comparison. The overall effect size for these studies is $g = .36$ (95% CI = .14 - .58). The calculated fail-safe N indicated that 26 studies with nonsignificant results would have to be added in order to reduce the overall effect size to below a statistically significant level. A test of heterogeneity indicated that the sample was homogeneous ($Q[8] = 11.04, p = .199; I^2 = 28\%$) and so no moderators were examined.

No treatment comparison. An examination of those 25 studies that reported on measures of parent adjustment (e.g., Beck Anxiety Inventory [BAI; Beck & Steer, 1990], Parenting Stress Index [PSI; Abidin, 1995]) comparing behavioral parenting interventions to no treatment comparisons yielded an overall effect size of $g = .49$ with a 95% CI of .33 to .64. A calculation of the fail-safe N indicated that 496 studies with nonsignificant results would have to be added to reduce the overall effect size to a level below statistical significance. Significant variability in the effect sizes was revealed by the test of heterogeneity ($Q[24] = 52.44, p = .001; I^2 = 54\%$), indicating the need to examine possible moderating variables.

Child and parent characteristics as moderators of parent adjustment.

Child age and formal diagnostic status were both significant moderators of parent adjustment. Examined categorically, child age did not moderate parent adjustment outcomes. However, when examined continuously using meta-regression, age was found to be a significant moderator such that as child age increases (up to age 10), effect sizes increase ($\text{Tau}^2 = .07, \text{SE} =$

.05). With regards to child formal diagnostic status, those studies in which children were diagnosed with ADHD had significantly greater effects than those in which children were not diagnosed with ADHD ($t(23) = 2.754, p = .01$).

Methodological characteristics as moderators of parent adjustment.

Independent groups *t*-tests indicated that methodological quality, experimenter allegiance, and publication status were not significant moderators of parent adjustment.

Discussion

Our search and coding yielded 19 new studies that had not been included in previous meta-analyses. Given the different inclusion criteria applied in previous meta-analyses (e.g., RCTs, specific parenting programs, diagnostic restrictions, age range, only published studies), not all of these studies would have been eligible for inclusion in previous meta-analyses. Of the 19 new studies, two were published since the most recent meta-analysis, and five were unpublished dissertations or manuscripts which were completed during the time period covered by previous meta-analyses. The remaining 12 sources were published during the time period covered by previous meta-analyses, but may have not been included due to being in lesser known journals, not examining the outcomes which were the focus of other meta-analyses, and/or looking at children who were outside the age range being examined.

In this meta-analysis, we determined the effects of parenting interventions delivered to parents of children with externalizing problems, examining parent behavior, child behavior, and parent adjustment outcomes. Findings were robust for most outcomes. We examined all studies, published and unpublished, from 1966 to December 2011, used the widely accepted random-effects analysis, examined children from 2 to 18 years of age with clinically significant behavior problems, examined all behavioral parenting programs, and conducted separate analyses for

those studies in which behavioral parenting interventions were compared a no-treatment group and those in which it was compared to an active comparison group. Therefore, to our knowledge, this meta-analysis is the most extensive one that has been conducted to date.

Moderator Variables

Due to the homogeneity of effects when compared to an active comparison group, moderators for parent behavior and parent adjustment were only examined for studies comparing a behavioral parenting intervention to a no-treatment group. The results are presented below.

Child characteristics as moderators of parent behavior outcomes.

Child formal diagnostic status and child age, examined categorically, did not moderate the effects of behavioral parenting interventions on parent behavior outcomes. However, significant effects were found when child age was examined continuously. With respect to child formal diagnostic status, the current results are consistent with the nonsignificant findings of previous meta-analyses of behavioral parenting interventions (de Graaf et al., 2008; Lundahl et al., 2006). In the present sample, children with diagnosed problems were often studied in samples that contained children with undiagnosed problems. Thus, moderating effects of child formal diagnostic status may have been overshadowed by children with undiagnosed problems.

Examined continuously, increases in child age were associated with greater parent behavior effects. This pattern is the reverse of that reported by Nowak and Heinrichs (2008). The discrepancy in findings may be partially accounted for by the difference in age range examined by Nowaks and Heinrichs (0 to 14), and the fact that the majority of children examined in that meta-analysis did not have clinically significant or diagnosed problems.

Methodological characteristics as moderators of parent behavior outcomes.

No methodological variables moderated the effect on parent behavior outcomes. It is possible that the inclusion criteria used to select studies for this meta-analysis restricted the range of methodological quality ratings.

Contrary to findings of other studies (e.g., Eisner, 2009; Miller, Wampold, & Varhely, 2008; Petrosino & Soydan, 2005), there was no effect of experimenter allegiance. Due to the low reliability in the coding this variable, these results should be interpreted with caution.

Publication status did not moderate the effects of behavioral parenting interventions on parent behavior outcomes. This finding is inconsistent with the concept of publication bias (Staines & Cleland, 2007), in which published had significantly larger effect sizes than unpublished studies. Inclusion criteria with respect to methodological quality may have reduced differences between published and unpublished studies, as they yielded similar ratings of methodological quality. This preliminary finding requires replication.

Child and methodological characteristics as moderators of child externalizing behavior.

Although some earlier meta-analyses reported heterogeneity of effects (e.g., Kaminski et al., 2008; Lundahl et al., 2006; Maughan et al., 2005; McCart et al., 2006; Nowak & Heinrichs, 2008; Piquero et al., 2009), in this meta-analysis child behavior effects were homogeneous, consistent with the findings reported by Lundahl et al. (2008) and Furlong et al. (2012). The discrepancy may be accounted for by the use of unclearly defined, fixed effects or mixed effects analyses in earlier reports. Furthermore, most of the meta-analyses that reported heterogeneity in child effects included both clinical and nonclinical populations which may have reduced the homogeneity of their samples.

Child characteristics as moderators of parent adjustment.

Child formal diagnostic status and child age were significant moderators of parent adjustment outcomes. Those studies in which families of children diagnosed with ADHD yielded greater effects on parent adjustment than did studies with children who were not diagnosed with ADHD. Examined continuously, child age was a significant moderator of parent adjustment outcomes such that as child age increased parent adjustment effects also increased. Younger children with clinically significant problems may represent an early onset group (Loeber, 1991) which experiences a number of other stressors present in the environment (McCabe, Hough, Wood, & Yeh, 2001). It appears in this sample that parents of children with early-onset problems did not demonstrate the same gains in terms of their own adjustment as did parents of older children.

Methodological characteristics as moderators of parent adjustment.

None of the methodological characteristics (i.e., comparison group, methodological quality, experimenter allegiance, and publication status) moderated parent adjustment outcomes.

Contribution of the grey literature

The limited number of studies from the grey literature did not have a significant effect on outcomes. Although there was no benefit to including the grey literature in the present meta-analysis, it would be premature to generalize this finding to other, less well-established bodies of research. Researchers in emerging areas may be prudent to determine whether inclusion of the grey literature would contribute to the accuracy and strength of research findings which are being communicated to program developers, group leaders, and parents.

Limitations

As with any meta-analysis, the findings presented in this paper have a limited shelf-life. As new information becomes available in the field, published or unpublished, it will be important

to re-examine the conclusions we have presented to ensure that practitioners, administrators, and consumers of behavioral parenting interventions are operating and making decisions on the best available data.

Due to the inclusion of children in more than one diagnostic category in most studies, it was impossible for us to determine the differential effectiveness of parent training interventions across different diagnostic categories. In this meta-analysis we chose not to use measures of methodological quality that were developed primarily for use with RCTs. We recognize that our tool gave considerable weight to measurement issues. In addition, although it would be helpful to examine other moderator variables, such as the level of treatment integrity reported, the way these variables are reported in the literature does not permit such comparisons. Future meta-analyses may benefit from examining such variables and their impact on outcomes, but for this to become possible, the literature must evolve to report on such variables in a more standardized way. The Consolidated Standards of Reporting Trials [CONSORT] group has published guidelines to help streamline the reporting of trials in order to facilitate understanding and clarity of reports (Moher et al., 2010; Schulz, Altman, & Moher, 2010). Although CONSORT does not refer to treatment integrity, as research syntheses are becoming important vehicles of knowledge translation about a particular field, it is recommended that the CONSORT guidelines be expanded to include guidelines for consistent reporting of other variables that are often examined in syntheses (e.g., participant age, treatment integrity, recruitment strategies). Notably, an extension of CONSORT for public health, psychology, social work, criminology and education (UPSCaLE; Grant, Montgomery, and Mayo-Wilson, 2012) is currently under development and may address these concerns.

The current meta-analysis found behavioral parenting interventions to be efficacious yielding improvement not only on parent and child behavior, but also parent adjustment. The effects are robust across clinical samples. Moderation analyses did identify subgroups that may be more vulnerable. Notably, parents of younger children with clinically significant problems derived fewer benefits in terms of changing parenting behaviors or improved adjustment. The positive effects on parent adjustment were marked for parents of children with ADHD. This suggests an important link between adopting behavioral strategies to manage the child's behavior and reduction in parental stress and distress.

References

Note. Articles preceded by an * were included in the analyses.

Abidin, R. R. (1995). *Parenting Stress Index: Professional Manual* (3rd. ed.). Odessa, FL:

Psychological Assessment Resources.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA school-age forms & profiles*.

Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

*Anastopoulos, A. D., Shelton, T. L., DuPaul, G. J., & Guevremont, D. C. (1993). Parent training for attention-deficit hyperactivity disorder: Its impact on parent functioning. *Journal of Abnormal Child Psychology*, 21, 581-596.

Auger, C. P. (1998). *Information sources in grey literature* (4th ed.). London: Bowker-Saur.

*Barkley, R. A., Guevremont, D. C., Anastopoulos, A. D., & Fletcher, K. E. (1992). A comparison of three family therapy programs for treating family conflicts in adolescents with attention-deficit hyperactivity disorder. *Journal of Consulting and Clinical Psychology*, 60, 450-462.

*Barkley, R. A., Shelton, T. L., Crosswait, C., Moorehouse, M., Fletcher, K., Barrett, S., ... Metevia, L. (2000). Multi-method psycho-educational intervention for preschool children with disruptive behavior: Preliminary results at post-treatment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41, 319-332.

Beck, A. T., & Steer, R. A. (1990). *Manual for the Beck Anxiety Inventory*. San Antonio, TX: Psychological Corporation.

*Bor, W., Sanders, M. R., & Markie-Dadds, C. (2002). The effects of the triple p-positive parenting program on preschool children with co-occurring disruptive behavior and

attentional/hyperactive difficulties. *Journal of Abnormal Child Psychology*, 30, 571-587.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis Version 2*. Englewood, NJ: Biostat.

*Braet, C., Meerschaert, T., Merlevede, E., Bosmans, G., Van Leeuwen, K., & De Mey, W. (2009). Prevention of antisocial behaviour: Evaluation of an early intervention programme. *European Journal of Developmental Psychology*, 6, 223-240.

*Connolly, L., Sharpy, J., & Fitzpatrick, C. (2001). Evaluation of a group treatment programme for parents of children with behavioural disorders. *Child & Adolescent Mental Health*, 6, 159-165.

*Costin, J., Lichte, C., Hill-Smith, A., Vance, A., & Luk, E. (2004). Parent group treatments for children with oppositional defiant disorder. *Australian e-journal for the Advancement of Mental Health*, 3, 1-8.

*Cruz, M. (2009). *Behavior parent training and fathers of young children with ADHD* (Unpublished doctoral dissertation). University at Buffalo, State University of New York, New York.

De Graaf, I., Speetjens, P., Smit, F., de Wolff, M., & Tavecchio, L. (2008). Effectiveness of the Triple P positive parenting program on parenting: A meta-analysis. *Family Relations*, 57, 553-566.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomized studies of health care interventions. *Journal of Epidemiology Community Health*, 52, 377-384.

- Dretzke, J., Davenport, C., Few, E., Barlow, J., Stewart-Brown, S., Bayliss, S., ... Hyde, C. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: A systematic review of randomized controlled trials. *Child and Adolescent Psychiatry and Mental Health*, 3, 1-10.
- Eisner, M. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology*, 5, 163-183.
- Eyberg, J., & Pincus, D. (1999). *The ECBI and SESBI-R: Eyberg Child Behavior and Sutter-Eyberg Student Behavior Inventory – Revised: Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Fletcher, R., Freeman, E., & Matthey, S. (2011). The impact of behavioural parent training on fathers' parenting: A meta-analysis of the triple p – positive parenting program. *Fathering*, 9, 291-312.
- Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S. M., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database for Systematic Reviews*, 2012(2), 1-344.
- *Gardner, F., Burton, J., & Klimes, I. (2006). Randomised controlled trial of a parenting intervention in the voluntary sector for reducing child conduct problems: Outcomes and mechanisms of change. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 47, 1123-1132.
- Grant, S., Montgomery, P., & Mayo-Wilson, E. (2012). Development of a CONSORT extension for interventions in public health and related disciplines. *The Lancet*, 380, S14.

- *Greene, R. W., Ablon, J. S., Monuteaux, M. C., Goring, J. C., Henin, A., Raezer-Blakely, L., . . . Rabbitt, S. (2004). Effectiveness of collaborative problem solving in affectively dysregulated children with oppositional-defiant disorder: Initial findings. *Journal of Consulting & Clinical Psychology, 72*, 1157-1164.
- Hopewell, S., McDonald, S., Clarke, M. J., & Egger, M. (2008). Grey literature in meta-analyses of randomized trials of health care interventions (Review). *The Cochrane Library*, 1-18.
- Hunsley, J., & Mash, E. J. (2008). Developing criteria for evidence-based assessment: An introduction to assessments that work. In J. Hunsley & E. J. Mash (Eds.), *A Guide To Assessments That Work* (pp. 3-14). New York: Oxford University Press.
- *Hutchings, J., Bywater, T., Daley, D., Gardner, F., Daley, D., Whitaker, C., . . . Edwards, R. T. (2007). Parenting intervention in Sure Start services for children at risk of developing conduct disorder: Pragmatic randomised controlled trial. *British Medical Journal, 334*, 678-682.
- Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials, 17*, 1-12.
- Jensen, P. S. (2003). What is the evidence for evidence-based treatments? *Emotional & Behavioral Disorders in Youth, 3*, 37-41.
- Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36*, 567-589.

- Kazdin, A. E., & Bass, D. (1989). Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting and Clinical Psychology, 57*, 138-147.
- *Kazdin, A. E., Siegel, T. C., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. *Journal of Consulting and Clinical Psychology, 60*, 733-747.
- *Kling, A., Forster, M., Sundell, K., & Melin, L. (2010). A randomized controlled effectiveness trial of parent management training with varying degrees of therapist support. *Behavior Therapy, 41*, 530-542.
- *Landy, S., & Menna, R. (2006). An evaluation of a group intervention for parents with aggressive young children: Improvements in child functioning, maternal confidence, parenting knowledge and attitudes. *Early Child Development and Care, 176*, 605-620.
- *Larsson, B., Fossum, S., Clifford, G., Drugli, M. B., Handegard, B. H., & Morch, W. T. (2009). Treatment of oppositional defiant and conduct problems in young Norwegian children: Results of a randomized controlled trial. *European Child & Adolescent Psychiatry, 18*, 42-52.
- *Lehner-Dua, L. L. (2001). *The effectiveness of Russell A. Barkley's parent training program on parents with school-aged children who have ADHD on their perceived severity of ADHD, stress, and sense of competence* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3030505)
- *Leung, C., Sanders, M. R., Leung, S., Mak, R., & Lau, J. (2003). An outcome evaluation of the implementation of the Triple P Positive Parenting Program in Hong Kong. *Family Process, 42*, 531-544.

Leykin, Y., & DeRubeis, R. J. (2009). Allegiance in psychotherapy outcome research :

Separating association from bias. *Clinical Psychology: Science and Practice*, 16, 54-65.

Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training:

Moderators and follow-up effects. *Clinical Psychology Review*, 26, 86-104.

Lundahl, B. W., Tollefson, D., Risser, H., & Lovejoy, M. C. (2008). A meta-analysis of father involvement in parent training. *Research on Social Work Practice*, 18, 97-106.

Maughan, D. R., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review*, 34, 267-286.

McCabe, K. M., Hough, R., Wood, P. A., Yeh, M. (2001). Childhood and adolescent onset conduct disorder: A test of the developmental taxonomy. *Journal of Abnormal Child Psychology*, 29, 305-316.

*McCabe, K. M., & Yeh, M. (2009). Parent-child interaction therapy for Mexican Americans: A randomized clinical trial. *Journal of Clinical Child & Adolescent Psychology*, 38, 753-759.

McCart, M. R., Priester, P. E., Davies, W. H., & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology*, 34, 527-543.

*McGilloway, S., Bywater, T., Mhaille, G. N., Furlong, M., O'Neill, D., Comiskey, C., ...

Donnelly, M. (n.d.). *Proving the power of positive parenting: A randomised control trial to investigate the effectiveness of the Incredible Years Basic Parent training programme in an Irish context (short-term outcomes)*. Retrieved from

http://eprints.nuim.ie/1816/1/Research_Report_-_Incredible_Years_Ireland_Study.pdf

- McLeod, B. D., & Weisz, J. R. (2004). Using dissertations to examine potential bias in child and adolescent clinical trials. *Journal of Consulting and Clinical Psychology, 72*, 235-251.
- Miller, S., Wampold, B., Varhely, K. (2008). Direct comparisons of treatment modalities for youth disorders: A meta-analysis. *Psychotherapy Research, 18*, 5-14.
- *Moharreri, F., Shahrivar, Z., Tehrani-doost, M., & Mahmoudi-Gharaei, J. (2008). Efficacy of the positive parenting program (triple p) for parents of children with attention deficit/hyperactivity disorder. *Iranian Journal of Psychiatry, 3*, 59-63.
- Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., ... Altman, D. G., for the CONSORT Group. (2010). CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trial. *British Medical Journal, 340*, c869-897.
- *Nixon, R. D. V., Sweeney, L., Erickson, D. B., & Touyz, S. W. (2003). Parent-child interaction therapy: A comparison of standard and abbreviated treatments for oppositional defiant preschoolers. *Journal of Consulting & Clinical Psychology, 71*, 251-260.
- Nowak, C., & Heinrichs, N. (2008). A comprehensive meta-analysis of triple p – positive parenting program using hierarchical linear modeling: Effectiveness and moderating variables. *Clinical Child and Family Psychology Review, 11*, 114-144.
- *Ogden, T., & Amlund-Hagen, K. (2008). Treatment effectiveness of parent management training in Norway: A randomized controlled trial of children with conduct problems. *Journal of Consulting & Clinical Psychology, 76*, 607-621.
- Paley, G., & Shapiro, D. A. (2002). Lessons from psychotherapy research for psychological interventions for people with schizophrenia. *Psychology and Psychotherapy, 75*, 5-17.

- Petrosino, A. & Soydan, H. (2005). The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1, 435-450.
- Piquero, A. R., Farrington, D. P., Welsh, B. C., Tremblay, R., & Jennings, W. G. (2009). Effects of early family/parent training programs on antisocial behavior and delinquency. *Journal of Experimental Criminology*, 5, 83-120.
- *Pisterman, S., Firestone, P., McGrath, P., Goodman, J. T., Webster, I., Mallory, R., & Goffin, B. (1992). The role of parent training in treatment of preschoolers with ADDH. *American Journal of Orthopsychiatry*, 62, 397- 408.
- *Pisterman, S., McGrath, P., Firestone, P., & Goodman, J. T. (1989). Outcome of parent-mediated treatment of preschoolers with Attention Deficit Disorder with hyperactivity. *Journal of Consulting & Clinical Psychology*, 57, 628-635.
- *Plant, K. M., & Sanders, M. R. (2007). Reducing problem behavior during care-giving in families of preschool-aged children with developmental disabilities. *Research in Developmental Disabilities*, 28, 362-385.
- Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems – a meta-analytic review. *Journal of Child Psychology and Psychiatry*, 47, 99-111.
- *Ross, C. M. (2002). *Bridging the gaps in services for families of children with attention-deficit/hyperactivity disorder: Examining the effectiveness of parent management training groups in a rural community* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. MQ85908)

- *Sanders, M. R., Markie-Dadds, C., Tully, L. A., & Bor, W. (2000). The Triple P-Positive Parenting Program: A comparison of enhanced, standard, and self-directed behavioral family intervention for parents of children with early onset conduct problems. *Journal of Consulting and Clinical Psychology, 68*, 624-640.
- *Schuhmann, E. M., Foote, R. C., Eyberg, S. M., Boggs, S. R., & Algina, J. (1998). Efficacy of parent-child interaction therapy: Interim report of a randomized trial with short-term maintenance. *Journal of Clinical Child Psychology, 27*, 34-45.
- Schulz, K. F., Altman, D. G., Moher, D., for the CONSORT Group. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *British Medical Journal, 340*, 698-702.
- *Scott, S., Spendor, Q., Doolan, M., Jacobs, B., & Aspland, H. (2001). Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice. *British Medical Journal, 323*, 1-7.
- Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy, 27*, 171-186.
- *Sheeber, L. B., & Johnson, J. H. (1994). Evaluation of a temperament-focused, parent-training program. *Journal of Clinical Child Psychology, 23*, 249-259.
- *Sofronoff, K., Jahnel, D., & Sanders, M. (2011). Stepping Stones Triple P seminars for parents of a child with a disability: A randomized controlled trial. *Research in Developmental Disabilities, 32*, 2253-2262.
- *Sofronoff, K., Leslie, A., & Brown, W. (2004). Parent management training and Asperger syndrome. A randomized controlled trial to evaluate a parent based intervention. *Autism, 8*, 301-317.

- *Sommer, J. L. (2007). *Efficacy of parent training for children with attention-deficit/hyperactivity disorder: A therapeutic component analysis* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 1443241)
- *Sonuga-Barke, E. J. S., Daley, D., Thompson, M., Laver-Bradbury, C., & Weeks, A. (2001). Parent-based therapies for preschool attention-deficit/hyperactivity disorder: A randomized controlled trial with a community sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40, 402-408.
- Staines, G. L., & Cleland, C. M. (2007). Bias in meta-analytic estimates of the absolute efficacy of psychotherapy. *Review of General Psychology*, 11, 329-347.
- Taylor, T. K., & Biglan, A. (1998). Behavioral family interventions for improving child-rearing: A review of the literature for clinicians and policy makers. *Clinical Child and Family Psychology Review*, 1, 41-60.
- *Taylor, T. K., Schmidt, F., Pepler, D., & Hodgins, C. (1998). A comparison of eclectic treatment with Webster-Stratton's parents and children series in a children's mental health center: A randomized controlled trial. *Behavior Therapy*, 29, 221-240.
- Thomas, R., & Zimmer-Gembeck, M. J. (2007). Behavioral outcomes of parent-child interaction therapy and Triple P-positive parenting program: A review and meta-analysis. *Journal of Abnormal Child Psychology*, 35, 475-495.
- *Thomas, R., & Zimmer-Gembeck, M. J. (2011). Accumulating evidence for parent-child interaction therapy in the prevention of child maltreatment. *Child Development*, 82, 177-192.

- *Thompson, R. W., Ruma, P. R., Schuchmann, L. F., & Burke, R. V. (1996). A cost-effectiveness evaluation of parent training. *Journal of Child & Family Studies*, 5, 415-429.
- *Thorell, L. B. (2009). The community parent education program (COPE): Treatment effects in a clinical and community-based sample. *Clinical Child Psychology and Psychiatry*, 14, 373-387.
- Tricco, A. C., Tetzlaff, J., Sampson, M., Fergusson, D., Cogo, E., Horsley, T., & Moher, D. (2008). Few systematic reviews exist documenting the extent of bias: A systematic review. *Journal of Clinical Epidemiology*, 61, 422-434.
- *Turner, K. M., Richards, M., & Sanders, M. R. (2007). Randomised clinical trial of a group parent education programme for Australian Indigenous families. *Journal of Paediatrics and Child Health*, 43, 429-437.
- *van de Wiel, N. M. H., Matthys, W., Cohen-Kettenis, P. T., Maassen, G. H., Lochman, J. E., & van Engeland, H. (2007). The effectiveness of an experimental treatment when compared to care as usual depends on the type of care as usual. *Behavior Modification*, 31, 298-312.
- *Van Den Hoofdakker, B. J., Van Der Veen-Mulders, L., Sytema, S., Emmelkamp, P. M. G., Minderaa, R. B., & Nauta, M. H. (2007). Effectiveness of behavioral parent training for children with ADHD in routine clinical practice: A randomized controlled study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46, 1263-1271.
- *Van Zeijl, J., Mesman, J., Van Ijzendoorn, M. H., Bakermans-Kranenburg, M. J., Juffer, F., Stolk, M. N., ... Alink, L. R. A. (2006). Attachment-based intervention for enhancing sensitive discipline in mothers of 1- to 3-year-old children at risk for externalizing

- behavior problems: A randomized control trial. *Journal of Consulting and Clinical Psychology*, 74, 994-1005.
- *Ward, S. L. (1997). *Glasser's parent training model: Effects on child and parent functioning* (Unpublished doctoral dissertation). University of Arizona, Arizona.
- *Webster-Stratton, C. (1992). Individually administered videotape parent training: "Who benefits?" *Cognitive Therapy and Research*, 16, 31-35.
- *Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct-problem children: Comparison with two cost-effective treatments and a control group. *Journal of Consulting & Clinical Psychology*, 56, 558-566.
- *Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology*, 33, 105-124.
- Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin*, 132, 132-149.
- *Whittingham, K., Sofronoff, K., Sheffield, J. K., & Sanders, M. R. (2009). Stepping stones Triple P: An RCT of a parenting program with parents of a child diagnosed with an autism spectrum disorder. *Journal of Abnormal Child Psychology*, 37, 469-480.
- *Wiggins, T. L., Sofronoff, K., & Sanders, M. R. (2009). Pathways triple P-positive parenting program: Effects on parent-child relationships and child behavior problems. *Family Process*, 48, 517-530.
- Wood, J. (2008). Methodology for dealing with duplicate study effects in meta-analysis. *Organizational Research Methods*, 11, 79-95.

Table 1

Meta-Analyses of Behavioral Parenting Interventions for Children's Disruptive Behavior

Authors	Diagnoses (or status)	# Studies	Inclusion Criteria	Child Ages	Conclusions (ES)	Moderators
de Graaf, Speetjens, Smit, de Wolff, & Tavecchio (2008)	Clinical: Y (26.32%); Dx: Y (5.26%)	19	1975-2006; Triple P level 4 intervention; PS and PSOC	n/a	P Style = .68 parent competence = .65	C & F variables; Tx Characteristics
Dretzke et al. (2009)	Clinical: Y (71.64%); Dx: Y (14.93%)	24	until 2006; RCT; ECBI or CBCL	0 to 18	SMD CB = -.67 (favors Interv. Grp)	n/a
Furlong et al. (2012)	Clinical: Y (100%) Dx: Y (46%)	15	Until January 2011; RCTs and Quasi-randomized studies	3 to 12	SMD CB = -.53 SMD PB = -.77 SMD PA = -.36	C & F variables; Tx Characteristics; Method. Characteristics
Fletcher, Freeman, & Matthey (2011)	Clinical: n/a Dx: n/a	28	1950-2011; RCTs; fathers targeted or eligible; Triple P	n/a	P Style = .77	Method. Characteristics
Kaminski, Valle, Filene, & Boyle (2008)	Clinical: Y (n/a); Dx: Y (n/a)	77	1990-2002	0 to 7	CB = .30 PB = .39	Tx Characteristics; Method. Characteristics

Authors	Diagnoses (or status)	# Studies	Inclusion Criteria	Child Ages	Conclusions (ES)	Moderators
Lundahl, Risser, & Lovejoy (2006)	Clinical: Y (67.16%); Dx: N	63	1974- August 2003	M = 6.79 years	CB = .42 PB = .45	C & F variables; Tx Characteristics
Lundahl, Tollefson, Risser, & Lovejoy (2008)	Clinical: n/a Dx: n/a	32	Until 2005	n/a	CB = .48 PB = .54	n/a
Maughan, Christiansen, Jenson, Olympia, & Clark (2005)	Clinical: n/a Dx: n/a	79	1966 – 2001	3 to 16	CB = .30	C & F variables; Tx Characteristics; Method. Characteristics
McCart, Priester, Davies, & Azen (2006)	Clinical: N; Dx: Y (20%)	30	Until 2004	under 18 years	CB = .38 PA+PS = .33 PB = .30	C & F variables; Tx Characteristics
Nowak & Heinrichs (2008)	Clinical: Y (16.36%); Dx: Y (5.45%)	55	1970-2007; Triple P	0 to 14	PB = .38; CB = .35; PA = .17	C & F variables; Tx Characteristics; Method. Characteristics
Piquero, Farrington, Welsh, Tremblay, & Jennings (2009)	Clinical: n/a Dx: n/a	55	Until 2008	1 to 5 (or M = ~5)	CB = .35	Tx Characteristics; Method. Characteristics

Authors	Diagnoses (or status)	# Studies	Inclusion Criteria	Child Ages	Conclusions (ES)	Moderators
Reyno & McGrath (2006)	Clinical: n/a Dx: n/a	31	1980 to 2004	1 to 16	n/a	C & F variables
Serketich & Dumas (1996)	Clinical: n/a Dx: n/a	26	Until 1995	M = 6.05	CB = 0.86	C & F variables; Tx Characteristics; Method. Characteristics
Thomas & Zimmer-Gembeck (2007)	Clinical: Y (79.17%) Dx: n/a	24	1980-2004; RCTs and single group follow-up; PCIT or Triple P	3 to 12	triple P = -.73 PCIT = -1.31	n/a

Note. ASB=Antisocial Behavior; BP=Behavior Problems; C=Child; CB=Child Behavior; CBCL=Child Behavior Checklist;

CL=Clinical; DBD=Disruptive Behavior Disorders; DX=Diagnosed; ECBI=Eyberg Child Behavior Inventory EP=Externalizing

Problems; ES=Effect Size; F=Family; M=Mean; Method. = Methodological; N = No; n/a=Not reported; NC= Non-clinical;

ODD=Oppositional Defiant Disorder; P=Parent; PA=Parent Adjustment; PB=Parent Behavior; PCIT=Parent-Child Interaction

Therapy; PS=Parenting Scale; PSOC=Parenting Sense of Competence Scale; RCT=Randomized Control Trial; SMD=Standardized

Mean Difference; Tx=Treatment; Y = Yes.

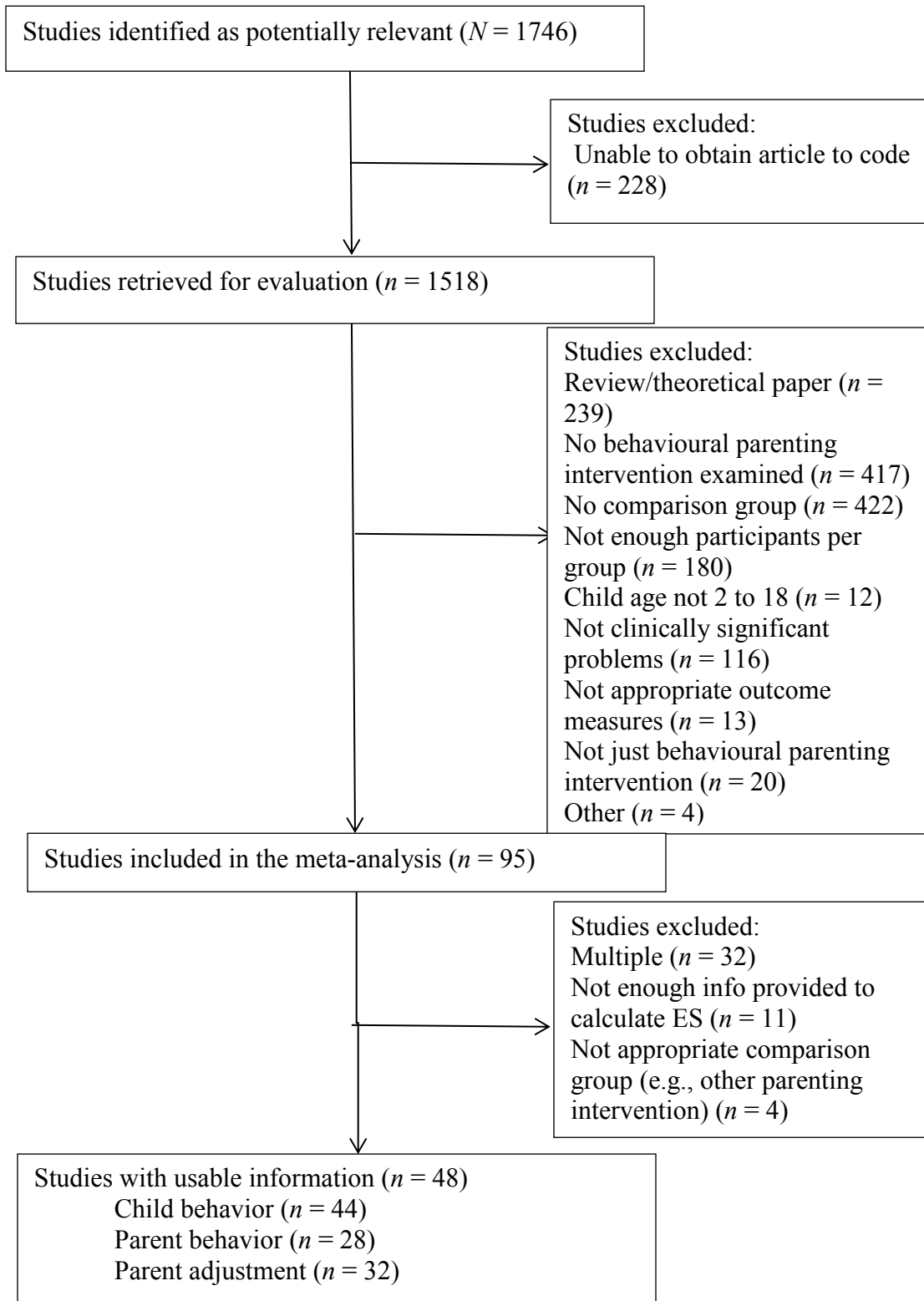


Figure 1. Flowchart of included and excluded studies

Table 2

Studies Included in the Present Meta-Analysis

Study	N (PT, Comparison)	Parent Behavior <i>g</i>	Child Behavior <i>g</i>	Parent Adjustment <i>g</i>
Anastopoulos, Shelton, DuPaul, & Guevremont (1993)	19, 15	-	0.86	0.74
Barkley, Guevremont, Anastopoulos, & Fletcher (1992)	20, 21	-	0.22	0.50
Barkley et al. (2000)	39, 42	0.10	0.26	0.18
Bor, Sanders, & Markie-Dadds (2002)	18, 27	0.92	1.02	0.58
Braet et al. (2009)	30, 19	0.39	0.20	0.04
Connolly, Sharpy, & Fitzpatrick (2001)	27, 18	-	0.51	0.25
Costin, Lichte, Hill-Smith, Vance, & Luk (2004)	22, 18	-	0.34	0.23
Cruz (2009)		0.03	-	-
Gardner, Burton, & Klimes (2006)	34, 26	0.58	0.51	0.37
Greene et al. (2004)	19, 28	-	0.79	-
Hutchings, et al. (2007)	104, 49	0.95	0.51	0.57
Kazdin, Siegel, & Bass (1992)	21, 25	-	0.29	0.09

Study	N (PT, Comparison)	Parent Behavior <i>g</i>	Child Behavior <i>g</i>	Parent Adjustment <i>g</i>
Kling, Forster, Sundell, & Melin (2010)	60, 40	0.62	0.74	-
Landy & Menna (2006)	20, 15	-	0.77	0.51
Larsson, Fossum, Clifford, Drugli, Handegard, & Morch (2009)	35, 25	1.12	0.60	0.75
Lehner-Dua (2001)	23, 25	-	0.25	0.85
Leung, Sanders, Leung, Mak, & Lau (2003)	33, 36	0.79	0.80	0.67
McCabe & Yeh (2009)	20, 18	0.64	0.62	0.92
McGilloway et al. (n.d.)	103, 46	-	0.77	-
Moharreri, Shahrivar, Tehrani-doost, & Mahmoudi-Gharaei (2008)	30, 30	3.74	-	2.03
Nixon, Sweeney, Erickson, & Touyz (2003)	19, 17	1.18	0.86	0.72
Ogden & Amlund-Hagen (2008)	59, 53	0.16	0.23	-
Pisterman et al. (1992)	23, 22	0.77	-	-
Pisterman, McGrath, Firestone, & Goodman (1989)	23, 23	1.25	-	-
Plant & Sanders (2007)	26, 24	0.79	0.35	0.51
Ross (2002)	50, 22	0.22	0.21	0.05

Study	N (PT, Comparison)	Parent Behavior <i>g</i>	Child Behavior <i>g</i>	Parent Adjustment <i>g</i>
Sanders, Markie-Dadds, Tully, & Bor (2000)	61, 71	0.61	0.68	0.32
Schuhmann, Foote, Eyberg, Boggs, & Algina (1998)	22, 20	-	1.30	-
Scott, Spendor, Doolan, Jacobs, & Aspland (2001)	56, 28	-	0.75	-
Sheeber & Johnson (1994)	20, 20	-	0.54	0.62
Sofronoff, Jahnel, & Sanders (2011)	27, 27	0.47	0.35	0.18
Sofronoff, Leslie, & Brown (2004)	18, 15	-	1.13	-
Sommer (2007)	75, 17	0.27	0.25	0.11
Sonuga-Barke, Daley, Thompson, Laver-Bradbury, & Weeks (2001)	30, 22	-	0.61	0.72
Taylor, Schmidt, Pepler, & Hodgins (1998)	15, 17	-	0.60	0.18
Thomas & Zimmer-Gembeck (2011)	41, 34	0.31	0.58	-0.21
Thompson, Ruma, Schuchmann, & Burke (1996)	39, 27	-	0.32	0.86
Thorell (2009)	56, 57	-	0.44	-
Turner, Richards, & Sanders (2007)	20, 18	0.55	0.26	0.20

Study	<i>N</i> (PT, Comparison)	Parent Behavior <i>g</i>	Child Behavior <i>g</i>	Parent Adjustment <i>g</i>
van de Wiel et al. (2007)	38, 26	-	0.10	-
Van Den Hoofdakker et al. (2007)	47, 47	-	0.05	0.06
Van Zeijl et al. (2006)	120, 117	0.14	0.13	-
Ward (1997)	22, 15	-	0.43	0.90
Webster-Stratton (1992)	48, 36	0.51	0.52	0.42
Webster-Stratton, Kolpacoff, & Hollinsworth (1988)	23, 25	0.64	0.74	0.29
Webster-Stratton, Reid, & Hammond (2004)	27, 25	0.54	0.93	-
Whittingham, Sofronoff, Sheffield, & Sanders (2009)	29, 30	0.21	0.21	-
Wiggins, Sofronoff, & Sanders (2009)	27, 22	0.89	0.56	-

Table 3

Summary of Moderator Analyses

Moderator	Reliability of rating	Parent behavior	Parent adjustment
Child			
Child formal diagnostic status and Comorbidity	.78 - .95		x
Child age	.75	X	x
Methodology			
Comparison group	.78		
Overall quality	.90		
Experimenter allegiance	.55		
Publication status	1.00	X	

Manuscript Two

Between Black and White: Examining Grey Literature in Meta-Analyses of Psychological
Research*

Sarah Bellefontaine and Catherine M. Lee

University of Ottawa

*This manuscript is in press at *Journal of Child and Family Studies*.

Abstract

The inclusion of grey literature in meta-analyses and reviews is controversial. We examine both the advantages and challenges of including grey literature in meta-analyses. An exemplar meta-analysis of behavioral parenting interventions on child behavior outcomes is used to demonstrate these issues. It also explores the influence of the inclusion of grey literature on outcomes, including whether effect sizes are affected by the inclusion of grey literature and describes the challenges of searching for grey literature using traditional search engines such as Google and Yahoo. Homogeneity and publication bias are also examined. Based on these results, recommendations are presented for meta-analysts and researchers.

Keywords: meta-analysis, grey literature, unpublished, evidence-based, behavioral parenting interventions

Between Black and White: Examining Grey Literature in Meta-Analyses of Psychological Research

Over the past two decades, there has been a trend in health care towards evidence-based practice (Hamilton, 2005; Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). Originating in medicine, the principles of evidence-based practice have been embraced by other professions including psychology (APA Task Force on the Future of Psychology, 2009), social work (Gilgun, 2005), nursing (Royle & Blythe, 1998), and psychiatry (American Academy of Child and Adolescent Psychiatry [AACAP], 2006). Evidence-based practice in psychology was defined by the American Psychological Association's (APA) task force on the topic as integrating clinical expertise with the best research available, taking into account the characteristics and preferences of the people being treated (APA, 2005; Levant, 2006).

Evidence-based practice relies on the results of research syntheses. Whereas Eysenck's (1952) narrative review of the psychotherapy outcome literature contained only 24 studies, this body of literature has grown exponentially and now includes thousands of studies. For example, a recent review of meta-analyses of Cognitive Behavioral Therapy, representing only a portion of this literature, examined the results of 332 outcomes studies in 16 meta-analyses (Butler, Chapman, Forman, & Beck, 2006).

Meta-analysis

Attempting to sift through the outcomes of the large volume of studies on a given topic presents a daunting challenge for any researcher trying to synthesize a body of literature (Glass, 1977; Light & Smith, 1971). A key tool in the process of reviewing research literature is a meta-analysis of relevant research findings. Until the 1970s, research on psychological treatment was

summarized narratively (e.g., Emrick, 1975; Eysenck, 1952). Meta-analysis is a way to quantitatively describe treatment outcomes in each study by calculating effect sizes, which represent the magnitude of the effect of treatment on an outcome variable (Hedges & Pigott, 2004). The effect sizes of various studies can then be combined into an overall effect size which numerically summarizes the results of research on a given topic (Hedges & Pigott, 2001).

Smith and Glass (1977) conducted the first meta-analysis of outcomes in psychotherapy for adults. A decade later, Weisz, Weiss, Alicke, and Klotz (1987) reported the first meta-analysis of the effectiveness of psychotherapy for children. These landmark publications, and all those that followed in their footsteps, carefully described the inclusion and exclusion criteria used in selecting the studies that comprised their dataset. Such criteria often specify the language of publication, date of publication, and precise methodological features that must be met for a study to be included in the meta-analysis. One of the most common initial criteria to be specified was that research should be published in a scholarly journal.

Both early proponents (e.g., Smith, Glass, & Miller, 1980) and critics (Eysenck, 1978) of meta-analysis have used the adage “garbage in – garbage out.” This refers to the fact that the quality of any meta-analysis is limited by the quality of the studies of which it is comprised. It is of prime concern, therefore, for meta-analysts to define the characteristics of a good quality study (McLeod & Weisz, 2004; Moher et al., 1999; Wilson & Lipsey, 2001). For some authors, publication in a peer-reviewed journal is an essential indicator of quality (e.g., Abby, Massey, Galandiuk, & Polk, 1994; Hopewell, McDonald, Clarke, & Egger, 2008). On the other hand, researchers such as McLeod and Weisz (2004) have raised concerns that publication is not necessarily a good indicator of quality, and that the published literature represents only a subsample of a research body. Consequently, they suggested that reliance on published literature

may yield a distorted view of findings as published literature may include a disproportionate number of positive findings, thus resulting in effect sizes that overestimate the true effects of treatment.

Grey Literature

In this age of information, the results of research studies can be found in many other sources in addition to what is available in published scholarly journals. Unpublished work is also referred to as grey literature. The most widely cited definition of grey literature was adopted at the Third International Conference on Grey Literature (Gokhale, 1997) and is referred to as the 'Luxembourg Definition.' According to this definition, *grey literature* refers to print or electronic literature that is produced by government, academia, business and industry, and is not controlled by commercial publishers (Auger, 1998). This can include materials such as unpublished studies, conference abstracts, conference proceedings, book chapters, government and agency reports, as well as unpublished doctoral dissertations.

Proponents of the inclusion of grey literature in meta-analyses argue that it balances the literature by minimizing the effects of publication bias (McAuley, Pham, Tugwell, & Moher, 2000; Pappas & Williams, 2011). Generally, research which demonstrates positive results tends to be published more easily than those studies that yield null or negative outcomes (Decullier, Lhéritier, & Chapuis, 2005; Dwan et al., 2008; Hopewell et al., 2008; Pappas & Williams, 2011), so it is argued that reliance on published literature may lead to an overestimation of effects (McAuley et al., 2000).

Calls for inclusion of grey literature in meta-analyses and systematic reviews have emphasized the importance of ensuring the quality of such literature (e.g., McAuley et al., 2000).

Although it is relatively easy to define the parameters of published literature by specifying the dates, language, search strategy, and methodology (e.g., a randomized controlled trial), the parameters of grey literature has been a topic of great debate (Auger, 1998; Frater, Myohanen, Taylor, & Keith, 2007; Hopewell et al., 2008; Pappas & Williams, 2011). In addition, critics argue that as there is no peer review process to vet the quality of grey literature, such studies are often of poor, or at the very least uncertain, quality (Sacks, Reitman, Pagan, & Kupelnick, 1996). As well, concerns often abound regarding the difficulty in both locating and accessing grey literature (Hopewell et al., 2008).

On average, across all disciplines, early studies reported that 31 to 33% of meta-analyses utilize grey literature (Cook et al., 1993). Meta-analyses which include grey literature have reported varied results. A review of grey literature in the area of Fetal Alcohol Spectrum Disorder by Benzies Premji, Hayden, and Serrett (2006) found benefits and drawbacks to the inclusion of grey literature, but in general found that information derived from the grey literature supported and reinforced the conclusions of journal articles. A review of meta-analyses of health care interventions reported greater treatment effects in published trials than in those found only in the grey literature (Hopewell et al., 2008). McAuley et al. (2000), in their review of the effects of grey literature inclusion on meta-analytic results, concluded that removal of grey works in meta-analysis resulted in a 15% larger estimate of the treatment effect and a trend towards more significant results.

Most meta-analysts seek to include the best evidence sources to review in their work so as to ensure the reliability of their findings and produce high-quality reviews (Conn, Valentine, Cooper, & Rantz, 2003). Much variability exists in the quality of grey literature depending on its source. For example, dissertations are subject to a rigorous review process and have evidence

demonstrating their high quality (McLeod & Weisz, 2004). By contrast, unpublished reports produced by private entities may or may not have gone through such rigorous review processes and are more likely to vary widely in the quality of their research. However, researchers such as Roth (2010) and Seymour (2010) have noted similar variability the quality of published articles, suggesting that the peer-review process may not be an adequate indicator of quality.

Martin, Perez, Sacristan, and Alvarez (2005) suggested that significant and nonsignificant results in grey literature must be treated differently. They noted that nonsignificant results are harder to publish than are statistically significant results; furthermore, they stated that statistically significant data may be unpublished due to either the delay between writing and publication, or due to low quality of the manuscript. The authors recommended that, with respect to psychiatric research, grey literature with nonsignificant results should be included in meta-analyses, but that grey literature with positive results should be examined for methodological quality before considering inclusion.

The road to publication of research is long and often includes many submissions and revisions. With journal rejection rates ranging from 31% to 88% and lag times between article acceptance and publication ranging from 2 to 11 months (Council of Editors, 2011), the publication process from manuscript preparation to publication can take years, if it is at all successful. Therefore, some research conducted is never disseminated through traditional, published media (Pappas & Williams, 2011).

Grey literature is viewed by some as being a solution to the significant time lag between research and publication (Pappas & Williams, 2011). Advocates argue that the grey literature provides a vehicle for disseminating studies with null or negative results which might otherwise

never be disseminated, or may take much longer to be published than studies with confirmatory results (Decullier et al., 2005; Dwan et al., 2008; Hopewell & Clarke, 2001), and ensures that key stakeholders (i.e., parents, service providers, and administrators) can be made aware of important information so as to make better informed choices. Grey literature can generally be produced quickly, making it more likely to be up-to-date and current (Coad et al., 2006).

The recent emphasis on knowledge transfer may also have contributed to an expansion of the grey literature. There is growing awareness of the need for both administrators and practitioners to have access to brief reports and research summaries (Barwick et al., 2008; Scullion, 2002). Publishing in scientific journals is likely neither sufficient nor the best way to ensure that research findings get disseminated to those that benefit most from them, including policy makers, administrators, and potential consumers (Mays & Hogg, 2012).

Challenges and benefits of grey literature.

Although the advent of the Internet has made it possible to locate many grey literature sources, searching for it can be time-consuming (Coad et al., 2006). Search engines such as Google, Bing, and Yahoo yield many results that can be adapted to more easily find grey literature using particular search techniques (Banks, 2010). However, online resources often change quickly and without notice (Coad et al., 2006), leaving the source of previously found information untraceable and researchers frustrated.

The Present Study

To illustrate the challenges of finding, accessing, and utilizing grey literature in a quantitative literature review, we present an example from a meta-analysis of the outcomes of parenting interventions. Search strategies used for accessing grey literature are detailed,

including the use of traditional search engines (e.g., Google and Yahoo) that have not been widely employed in previous meta-analyses despite their identification by researchers as providing easier and quicker ways to identify grey literature (e.g., Coad et al., 2006).

Exemplar

This exemplar is a meta-analysis of the literature regarding the efficacy of behavioral parenting interventions in the treatment of disruptive behaviour disorders. The majority of meta-analyses of behavioral parenting interventions as a treatment for disruptive behaviour disorders have included only published studies (e.g., de Graaf, Speetjens, Smit, de Wolff, & Tavecchio, 2008; Dretzke, Davenport, Frew, Barlow, Stewart-Brown, Bayliss, et al., 2009; Fletcher, Freeman, & Matthey, 2011; Kaminski, Valle, Filene, & Boyle, 2008; Lundahl, Risser, & Lovejoy, 2006; Lundahl, Tollefson, Risser, & Lovejoy, 2008; McCart, Priester, Davies, & Azen, 2006; Reyno & McGrath, 2006; Thomas & Zimmer-Gembeck, 2007); a much smaller number have included the grey literature (e.g., Furlong et al., 2012; Maughan et al., 2005; Nowak and Heinrichs, 2008; Piquero et al., 2009; Serketich & Dumas, 1996). In those meta-analyses, grey literature was defined as dissertations and theses, and in two cases also as unpublished or in-preparation manuscripts. Three of the meta-analyses (i.e., Furlong et al., 2012; Nowak & Heinrichs, 2008; Serketich & Dumas, 1996) did not examine publication status as a potential moderator. The other two meta-analyses (i.e., Piquero et al., 2009; Maughan et al., 2005) found publication status to be a nonsignificant moderator variable of outcome, however, neither examined the possible differences in the methodological quality of published versus unpublished studies.

This meta-analysis was designed to address several research questions. First, we examined the number of studies in the grey literature that met our inclusion criteria. Second, we determined whether these grey literature reports referred to studies we had already located in the published literature. Third, we examined whether effect sizes for child behavior, parent behavior, and parent adjustment outcomes varied for published versus grey literature. Finally, based on researchers' concerns that unpublished studies represent lower quality studies, we explored whether there was a significant difference in methodological quality between published and unpublished studies that were included in the present meta-analysis.

Traditional search methods employed

This meta-analysis included both published studies and grey literature from 1966 until December 2011. A comprehensive search of PsycINFO, PubMed, Scopus, and ERIC was used to identify relevant studies published in peer-reviewed journals and book chapters. Hand searches of key journals were undertaken to identify relevant articles which may have been overlooked by other search strategies (see Appendix A for a full list of key journals), and the reference lists of relevant articles, review papers and chapters were examined to find any other studies which were not identified by other means.

The 41 search terms used to identify relevant literature included all those that had been used in previous meta-analyses, as well as several additional terms in order to access the greatest number of relevant articles. These included terms related to intervention types and program names (*behavio(u)ral family interaction; behavio(u)r modification; behavio(u)ral parent training; behavio(u)r/behavior therapy; contingency management; parent-child interaction; Parent Effectiveness Training; Parent Management Training; parent training; Community Parent Education; COPE; Defiant Children; Incredible Years; Parent-Child Interaction*

Therapy; PCIT; Positive Parenting Program; Triple P), as well as disruptive behavior (*ADD; ADHD; aggressive; antisocial; Attention Deficit Disorder; Attention Deficit Hyperactivity Disorder; behavio(u)r disorder; CD; conduct disorder; conduct problem; DBD; defiant; disruptive; Disruptive Behavio(u)r Disorder; externalizing problem; hyperkinetic; ODD; Oppositional Defiant Disorder*) and different combinations and permutations of these search terms. Appendix B presents a complete listing of all search terms used.

Identifying Grey Literature

We defined grey literature as any research report that had not been published in scholarly journals. This included dissertations and theses, conference proceedings, book chapters, unpublished manuscripts, unpublished data, and reports.

PsycINFO was used to identify dissertations dating back to 1995. Proquest Dissertations and Theses was also searched to identify unpublished dissertations which were not available using PsycINFO (i.e., those produced before 1995) as well as to obtain full-text versions of dissertation abstracts found through other search strategies. Further unpublished studies were identified using Scopus, a database which includes peer-reviewed publications, conference proceedings, and trade publications across a number of different disciplines. Requests sent to researchers who had published in the field asking them to forward unpublished studies or data yielded no further studies.

In addition, online searches were performed using the Google and Yahoo search engines to identify further unpublished studies. Using these search engines, the first 500 search results were examined. This was decided in order to ensure the largest possible number of relevant identified resources within a manageable number of search results. To maximize the relevance of

searches performed on the internet search engines Google and Yahoo, only combinations of search terms (e.g., *parent training* and *disruptive*) were used. Initial searches were performed in 2009. Updates to searches were performed in August to December 2011.

Google Search Strategy

Google search did not allow for any aggregation of search results, so searches were performed one at a time. At the time of initial search, Google allowed a number of search results to be viewed before blocking out searchers for a period of approximately one hour, decreasing the user-friendliness of the experience for researchers. Attempts to contact Google in order to have this function removed were unsuccessful. By the time that the search updates were performed (August to December 2011), Google had added functions which allowed for easier updating of searches (e.g., ability to narrow down results by date added). In addition, heavy searchers were no longer blocked out of searching for a period of time, but were simply asked, after a certain number of results, to perform a CAPTCHA (i.e., Completely Automated Public Turing test to tell Computers and Humans Apart) in order to distinguish between humans performing searches and automated computer searches. CAPTCHAs involve having a user type a word, words, or pseudo-words presented in a distorted picture into a text box. The reason for such tests is that it is a task which is thought to be easy for humans to perform, but difficult for computers to do. Through the use of CAPTCHAs, Google sought to reduce violations of its terms of services, for example, to lessen the impact of computer viruses which use search engines as their vehicle to vulnerable websites (i.e., search worms; Provos, McClain, & Wang, 2006).

Yahoo Search Strategy

As the Google search strategy yielded a great number of results, many of them duplicates, measures were undertaken to see if there was a way in which Yahoo searches could be sped up by removing duplicate results. Thus, a program called ReSearch (Posner & Bellefontaine, 2011) was developed. The program was designed to work with a publicly available search application programming interface (API) which was specific to Yahoo called Yahoo Search Boss (<http://boss.yahoo.com>). In particular, ReSearch used the API in order to collect the first 500 results of each search and aggregate them. The program was able to run multiple searches and combine the search terms as required. It was also programmed to search for and remove all duplicate results, ordering the results by their frequency over all searches run and providing the user with the highest ranking each result obtained in any search. All results were stored in a database, allowing for future comparisons with the goal of being able to update searches. From the time of the initial searches (in 2009) to the time of the search update (in 2011), the Yahoo API changed from version 1 to version 2, necessitating an update in the ReSearch program. During that time, Yahoo and Microsoft also developed a search alliance whereby Yahoo search results were generated by the newer, more powerful Bing search engine. This led to an exponential increase in the search results yielded through Yahoo.

The full coding procedures are presented elsewhere (Bellefontaine & Lee, 2013). Documents were coded in five separate rounds of coding: 1) for inclusion in the meta-analysis (see Appendix C and D respectively for coding manual and coding sheet); 2) to determine whether articles were multiple reports from the same data-set (see Appendix E for coding manual and Appendix F for coding sheet); 3) to extract information to calculate effect sizes; 4) to determine methodological quality (see Appendix G for coding manual and Appendix H for

coding sheet); and 5) to code moderator variables (see Appendix I for coding manual and Appendix J for coding sheet).

Often a single research project yields multiple publications, for example when overall findings are presented, and follow-up data are subsequently published, or when data are published on the complete sample, and then a subset of the sample. In meta-analysis, it is essential that each sample is counted only once, so it was necessary to determine whether articles were based on the same sample and to identify which report is the most complete. The coding scheme used (Appendix E) was based in part on a detection heuristic proposed by Wood (2008), which was adapted to psychological research. A list of criteria were examined to establish whether studies were multiples: whether they shared one or more authors, referenced another report that was included in the meta-analysis, or described the same treatment, recruitment and referral procedures, time period, age range of children, response rates, assessment tools used, and parent ages as other included articles. If at least seven of these conditions were met, then the study was coded as a multiple, and the report with the most comprehensive data set was included in the meta-analysis. Data from six grey literature studies that were identified as multiples of studies already in the meta-analysis were excluded from the meta-analysis in favor of their published counterparts. Most (83%) of these journal articles were published within two years of the publication of the grey literature report, and all were published within four years.

Methodological quality was determined for both published and unpublished literature on a scale from 0 to 7 (see Appendix G and Appendix H for the coding manual and coding sheet respectively). Up to two points were awarded based on whether a randomization procedure was used, and the average psychometric quality of measures on a 5-point scale was used, based on

criteria proposed by Hunsley and Mash (2008) regarding the norms, internal consistency, test-retest reliability, content validity, and construct validity of measures used.

Results

Question 1. How Many Studies That Met Inclusion Criteria Were Found in Grey Literature?

The main results of this meta-analysis are reported elsewhere (Bellefontaine & Lee, 2013). The search yielded 78 published articles and 21 grey literature documents (eight were unpublished studies, three conference proceedings, and 10 dissertations) that met our initial inclusion criteria. Almost one-third of these grey literature documents were obtained from dissertation abstracts, with the rest coming from PsycINFO, Scopus, Google, or Yahoo.

Question 2. How Many of the Grey Literature Reports Referred to Studies That Had Already Been Identified in the Published Literature?

Figure 1 is a flow chart of the inclusion and exclusion of grey literature. 537 reports were identified as potentially relevant, of which 391 were able to be accessed. The rest were unable to be accessed due to unavailability online, through the institutional library, as well as interlibrary loans. Four of these were located but have yet to be received after more than one year. After coding for inclusion, 18 reports remained. Coding to identify multiple reports on the same study reduced the number to 11 grey literature reports. Of the seven reports that were excluded, one was a grey report of the same dataset as another grey report that was included, and six were eventually published. Five studies reported insufficient data to allow for the calculation of effect sizes. Therefore, the final number of grey sources included in the meta-analysis was six,

comprised of five dissertations and one unpublished report, representing 12.5% of the 48 total sources included in the meta-analysis.

Question 3. Were Effect Sizes Different in the Grey Literature Than in the Published Literature?

Parent behavior outcomes. Across all 24 studies comparing behavioral parenting interventions to a waitlist or no-treatment control group, parent behavior measures yielded an overall effect size of $g = .71$ with a 95% CI of .51 to .92. The test of heterogeneity indicated significant heterogeneity ($Q[23] = 85.39, p < .001; I^2 = 74\%$). The analysis was repeated independently for published and unpublished studies. In the published literature, representing 22 studies, behavioral parenting interventions were found to have an effect of $g = .76$ with a 95% CI of .55 - .98 on parent behavior outcomes. Among the two unpublished studies, the effect size of behavioral parenting interventions on parent behavior outcomes was $g = .15$ (95% CI = -.23 - .52). The difference between the effect sizes of the published and unpublished studies was nonsignificant ($t(22) = 1.664, p = .11$).

Once again, the analyses were repeated for the four studies comparing behavioral parenting interventions to an active control group. The overall effect size was $g = .20$ (95% CI = .01 - .39). A test of heterogeneity indicated that the effect sizes were homogeneous ($Q[3] = 2.14, p = .544; I^2 = 0\%$). Effect sizes remained comparable for the three published ($g = .20$; 95% CI = -.01 - .41) and one unpublished ($g = .27$; 95% CI = -.26 - .80) studies.

Child behavior outcomes. A random-effects model using weighted effect sizes was used to calculate Hedge's g for parent reports of child behavior, along with 95% Confidence Intervals (CI). Heterogeneity was assessed by conducting the Q test, and the magnitude of heterogeneity

was estimated using I^2 . File-drawer analyses were conducted to determine the number of nonsignificant studies needed to reduce the effect below statistical significance. Within studies, effect sizes of different parent-report measures of child behavior were aggregated into a single study effect size by calculating the mean of each individual effect size. Across all 33 studies that compared behavioral parenting interventions to a waitlist or no-treatment control group, published and unpublished, child behavior measures yielded an effect size of $g = .58$ (95% CI = .49 - .67). The test of heterogeneity indicated homogeneity in the effect sizes ($Q[32] = 27.75$, $p = .682$; $I^2 = 0\%$). Effect sizes were then calculated separately for published and unpublished studies. Behavioral parenting interventions were found to produce effect sizes that were comparable for the 30 published ($g = .58$; 95% CI = .49 - .68) and three unpublished ($g = .51$; 95% CI = .15 - .88) studies.

The analyses were then repeated for those 13 studies that compared behavioral parenting interventions to an active control group. This yielded an overall effect size of $g = .27$ with a 95% CI of .14 to .40. A test of heterogeneity indicated that the effect sizes were homogeneous ($Q[12] = 9.33$, $p = .675$; $I^2 = 0\%$). Effect sizes were once again comparable for the 11 published ($g = .27$; 95% CI = .13 - .41) and two unpublished ($g = .25$; 95% CI = -.13 - .63) studies.

Parent adjustment outcomes. Overall, the 25 studies that compared behavioral parenting interventions to a waitlist or no-treatment control group had effects on parent adjustment measures of $g = .49$ with a 95% CI of .33 to .65. The Q test for heterogeneity showed significant variability in the effects ($Q[24] = 52.44$, $p = .001$; $I^2 = 54\%$). Similar effect sizes were found for the 23 published ($g = .49$; 95% CI = .33 - .66) and two unpublished ($g = .45$; 95% CI = -.39 - 1.28) studies.

The analyses were repeated for the nine studies that compared behavioral parenting interventions to an active control group. A test for heterogeneity indicated that the effects were homogeneous ($Q[8] = 11.04$, $p = .199$; $I^2 = 28\%$). The overall effect size was $g = .38$ with a 95% CI of .14 to .62. Effect sizes for the seven published ($g = .32$; 95% CI = .08 - .56) and two unpublished ($g = .48$; 95% CI = -.21 - 1.18) studies were comparable.

Question 4. Does the Methodological Quality of Published Studies Differ From That of Unpublished Studies?

Average methodological quality was calculated separately for the 42 published ($M = 5.65$, $SD = 1.51$) and 6 unpublished ($M = 4.45$, $SD = 1.88$) studies. An independent samples t-test was performed which demonstrated no significant difference between the quality of studies based on publication status ($t(46) = 1.77$; $p = .083$).

Discussion

The substantial investment of effort in searching for grey literature yielded an initial pool of 78 published studies and 21 grey literature reports; coding for multiple reports from the same data set and for the availability of data to code effect sizes yielded a final sample of 42 published studies and 6 grey literature reports. The grey literature therefore represented 12.5% of the studies included in the present meta-analysis, which is within the range of representation of grey literature found in meta-analyses (McAuley et al., 2000). Notably, McAuley et al. (2000) found that most meta-analyses reported only one grey source. In many of the previous meta-analyses of behavioral parenting interventions, no grey sources were included, two meta-analyses included three unpublished sources (de Graaf et al., 2008; Piquero et al., 2009), one included two unpublished sources (Furlong et al., 2012) and another included four dissertations (Serketich &

Dumas, 1996). Notably, two previous of the previous meta-analyses included 18 grey sources. However one of these (Nowak & Heinrichs, 2008) included both German and English sources, and the other (Maughan et al., 2005) was not limited to clinically significant problems and included within-groups and single-subject designs in addition to between-subjects designs. It is noteworthy that the six grey literature reports were primarily made up of theses and dissertations and included only one other report. It would, therefore, have been possible to identify 83.3% of the grey literature that was included by searching only for theses and dissertations. Six grey reports were excluded from analyses as they were later published. Few of the previous meta-analyses of behavioral parenting interventions have examined the attrition of grey literature reports due to publication. Piquero et al. (2009) reported excluding two unpublished studies due to non-independence of data, but it is unclear whether these were due to being published or not.

The findings from this meta-analysis are generally consistent with the findings reported by Maughan et al. (2005) as well as Piquero et al. (2009) of nonsignificant differences in effect sizes between published and unpublished studies. Parent behavior, child behavior, and parent adjustment measures yielded similar effect sizes for published and unpublished literature. This is consistent with findings of other meta-analyses in this area, whether or not they included grey literature. However, it is important to note that in some cases, particularly for unpublished studies, the number of studies included remained quite low and may thus be masking true differences in the effects. This made it difficult to examine homogeneity separately for published and unpublished studies, and therefore study results should be replicated with newer sources included.

Based on concerns raised by previous researchers (e.g., Roth, 2010; Pappas & Williams, 2011; Seymour, 2010) we also examined whether methodological quality differed between

published and unpublished studies. Our results indicated no significant differences between published and unpublished studies in terms of methodological quality. Therefore, it appears, at least in the body of literature pertaining to the effects of behavioral parenting interventions on children's disruptive behavior, that quality of the literature is comparable. However, our ratings of methodological quality focused mainly on group randomization procedures and the psychometric properties of the measures that were included. It may be that differences between published and unpublished studies may have become apparent had we used a more elaborate coding scheme for methodological quality.

Inclusion of Grey Literature

Only a small amount of grey literature that was eventually included in this meta-analysis, with no significant difference found between the effect sizes obtained with and without the inclusion of grey literature. Therefore, it seems that the inclusion of grey literature did not add a great deal of new information to the evaluation of outcomes of parenting interventions, with the exception of parent behavior outcomes. Given the minimal effort required to access dissertations and theses, and the unpredictability of which outcomes may be affected by grey literature exclusion, it is prudent to include them in meta-analyses, especially as a number of theses will go on to be published in scholarly journals in less than five years. Within newer research areas, it still may be prudent to search for and include grey literature. In such areas, grey literature typically provides the most current, up-to-date information, providing a snapshot in time as to what is happening with a body of literature. For this exemplar meta-analysis, it appears as though most grey literature is eventually published. Further research is needed to determine whether, in general, good quality grey literature is eventually published in other established research areas.

Search Strategies for Accessing Grey Literature

The internet has greatly simplified the task of searching for grey literature. Modern search engines such as Yahoo and Google yield numerous search results at the click of a button, making it incredibly simple to access a multitude of grey documents on a particular subject. The caveat is that one must know what particular words (and combinations of words) are the “gatekeeper” to desired results. Furthermore, the volume of results obtained by such methods is time-consuming to sort through and yields a relatively small number of useful sources. Google and Yahoo, from our experience, provided similar useful yields, each with their respective pros and cons. For example, Google removed a number of duplicate results, but restricted searches. Yahoo, on the other hand, provided many more duplicate results, but allowed for an independent program to be developed to search through its results. However, it is important to note that, in error, the sources approximately half of the grey materials obtained was not recorded, and so determining the utility of each of these search strategies is not possible in the present meta-analysis.

Although the present authors did not evaluate the usefulness of MSN Bing search as an alternative due to its small share of the search market at the time that these searches were conducted, it is recommended that it be evaluated as an alternative for those who would like to include grey literature results in systematic reviews and meta-analyses.

Limitation

Unfortunately, 13% of identified resources, including articles, books, theses or conference proceedings, could not be obtained to adequately assess for inclusion into the present meta-analysis. 137 of these were dissertations which were identified primarily (68.6%) using

PsycINFO and Proquest Dissertations and Theses. Many were from unidentified or non-Canadian universities and 129 of them had been defended at least 15 years prior to 2011. Only one previous meta-analysis (Dretzke et al., 2009) examined the amount of inaccessible resources and reported that they were unable to obtain 9% of the articles they had identified. The present findings point to a gap in researchers' ability to access all reports.

Recommendations for Researchers

The decision of whether or not to search for grey literature is a challenging one. Our recommendation is that all meta-analyses should include all relevant and accessible dissertations and theses, as in most cases they represent future published literature. However, those areas of study where limited published literature is available should perform a search of the grey literature through at least one traditional search engine (e.g., Google, Yahoo, or Bing) to ensure that all relevant literature is being included and that conclusions reached are not overestimating the effect size in a given field.

The inclusion of grey literature in this exemplar meta-analysis raised some challenges. Notably, determining whether studies were multiple studies based on the same dataset was a particular challenge due to inconsistent reporting of information. Several studies did not directly reference other manuscripts that were based on the same dataset. Therefore, careful examination of sample characteristics needed to be undertaken, making it a time-consuming process. The Consolidated Standards of Reporting Trials [CONSORT] group has published guidelines to help streamline the reporting of trials in order to facilitate understanding and clarity of reports (Moher et al., 2010; Schulz, Altman, & Moher, 2010). The CONSORT statement is designed to address this problem by ensuring that trials are registered, and having authors provide the name and

number of the registry of their trials. It is therefore recommended that trials be registered and that CONSORT guidelines with respect to reporting trial numbers and registry names be followed. If this cannot be done, it is recommended that studies based on a common dataset systematically refer to one another so as to minimize the number of duplicate studies which are inadvertently included in systematic reviews and meta-analyses. These measures have been recommended in the Publication Manual of the American Psychological Association (2010). By taking small steps to increase the accessibility of grey literature to researchers, we improve the overall quality of the reviews many clients, practitioners, and researchers rely on.

References

- Abby, M., Massey, M. D., Galandiuk, S., & Polk, Jr., H. C. (1994). Peer review is an effective screening process to evaluate medical manuscripts. *Journal of the American Medical Association*, 272, 105-107. DOI:10.1001/jama.1994.03520020031008
- American Academy of Child and Adolescent Psychiatry (AACAP). (2006). Policy statement on Evidence-based Practice. Retrieved September 10, 2012 from http://aacap.org/cs/root/policy_statements/evidence_based_practice.
- American Psychological Association. (APA) (2005). *Policy statement on evidence-based practice in psychology*. Retrieved from <http://www.apa.org/practice/resources/evidence/evidence-based-statement.pdf>
- American Psychological Association. (APA). (2010). *Publication manual of the American Psychological Association*. Washington, DC: APA.
- American Psychological Association, Task Force on the Future of Psychology. (2009). *2009 Presidential Task Force on the Future of Psychology Practice Final Report*. Retrieved from <http://www.apa.org/pubs/info/reports/future-practice.pdf>
- Auger, C. P. (1998). *Information sources in grey literature (guides to information sources)*. 4th edition. London: Bowker Saur.
- Banks, M. A. (2010). Blog posts and tweets: The next frontier for grey literature. In D. J. Farace, & J. Schopfel (Eds.), *Grey literature in library and information studies*. Berlin: De Gruyter Saur.
- Barwick, M., Boydell, K. M., Stasiulis, E., Ferguson, H. B., Blase, K., Fixsen, D. (2008). Research utilization among children's mental health providers. *Implementation Science*, 3, 19. DOI:10.1186/1748-5908-3-19

- Bellefontaine, S., & Lee, C. M. (2013). *A meta-analytic examination of behavioral parent training in the treatment of children's disruptive behavior*. Manuscript submitted for publication.
- Benzies, K. M., Premji, S., Hayden, K. A., & Serrett, K. (2006). State-of-the-evidence reviews: Advantages and challenges of including grey literature. *Worldviews on Evidence-based Nursing*, 3, 55-61. DOI: 10.1111/j.1741-6787.2006.00051.x
- Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review*, 26, 17-31. DOI: 10.1016/j.cpr.2005.07.003
- Coad, J., Hardicare, J., & Devitt, P. (2006). Searching for and using grey literature. *Nursing Times*, 102, 35-36.
- Conn, V. S., Valentine, J. C., Cooper, H. M., & Rantz, M. J. (2003). Grey literature in meta-analyses. *Nursing Research*, 52, 256-261.
- Cook, D.J., Guyatt, G.H., Ryan, G., Clifton, J., Buckingham, L., Willan, A., ... Oxman, A. D. (1993). Should unpublished data be included in meta-analyses? Current convictions and controversies. *The Journal of the American Medical Association*, 269, 2749-2753. DOI: 10.1001/jama.1993.03500210049030
- Council of Editors. (2011). Summary report of journal operations, 2010. *American Psychologist*, 66, 405-406. DOI: 10.1037/a0023978
- De Graaf, I., Speetjens, P., Smit, F., de Wolff, M., & Tavecchio, L. (2008). Effectiveness of the Triple P positive parenting program on parenting: A meta-analysis. *Family Relations*, 57, 553-566. DOI: 10.1177/0145445508317134

- Decullier, E., Lhéritier, V., & Chapuis, F. (2005). Fate of biomedical research protocols and publication bias in France : Retrospective cohort study. *BMJ*, 331, 19. DOI: <http://dx.doi.org/10.1136/bmj.38488.385995.8F>
- Dretzke, J., Davenport, C., Few, E., Barlow, J., Stewart-Brown, S., Bayliss, S., ... Hyde, C. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: A systematic review of randomized controlled trials. *Child and Adolescent Psychiatry and Mental Health*, 3, 1-10. DOI:10.1186/1753-2000-3-7
- Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A-W., Cronin, E., ... Williamson, P. R. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE*, 3, e3081. DOI:10.1371/journal.pone.0003081
- Emrick, C.D. (1975). A review of psychologically oriented treatment of alcoholism. *Journal of Studies on Alcohol*, 36, 88-108.
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting and Clinical Psychology*, 16, 319-324.
- Eysenck, H. J. (1978). An exercise in mega-silliness. *American Psychologist*, 33, 517. DOI: 10.1037/0003-066X.33.5.517.a
- Fletcher, R., Freeman, E., & Matthey, S. (2011). The impact of behavioural parent training on fathers' parenting: A meta-analysis of the triple p – positive parenting program. *Fathering*, 9, 291-312. DOI: 10.3149/fth.0903.291
- Frater, J., Myohanen, L., Taylor, E., & Keith, L. (2007). What would you tell me if I said grey literature? The New York academy of medicine grey literature report. *Journal of Electronic Resources in Medical Libraries*, 4, 145-153.

- Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S. M., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database for Systematic Reviews*, 2012(2), 1-344. DOI: 10.1002/14651858.CD008225.pub2.
- Gilgun, J. F. (2005). The four cornerstones of evidence-based practice in social work. *Research on Social Work Practice*, 15, 52-61. DOI: 10.1177/1049731504269581
- Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research Education*, 5, 351-379. DOI: 10.3102/0091732X005001351
- Gokhale, P. (1997). Grey Literature Varieties - Definitional Problems. In *Third International Conference on Grey Literature: Perspectives on the Design and Transfer of Scientific and Technical Information*, 13-14 November 1997.
- Hamilton, J. (2005). Clinician's guide to evidence-based practice. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44, 494-498.
- Hedges, L.V., & Pigott, T.D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, 6, 203-217.
- Hedges, L.V., & Pigott, T.D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9, 426-445. DOI: 10.1037/1082-989X.9.4.426
- Hopewell, S., & Clarke, M. (2001). Methodologists and their methods: Do methodologists write up their conference presentations or is it just 15 minutes of fame? *International Journal of Technology Assessment in Health Care*, 17, 601-603.
- Hopewell, S., McDonald, S., Clarke, M. J., & Egger, M. (2008). Grey literature in meta-analyses of randomized trials of health care interventions. *The Cochrane Library*, 1-18.
- Hunsley, J., & Mash, E. J. (2008). Developing criteria for evidence-based assessment: An

- introduction to assessments that work. In J. Hunsley & E. J. Mash (Eds.), *A Guide To Assessments That Work* (pp. 3-14). New York: Oxford University Press.
- Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology*, 36, 567-589. DOI: 10.1007/s10802-007-9201-9
- Levant, R. F. (2006). Making psychology a household word. *American Psychologist*, 61, 383-395.
- Light, R. J., & Smith, P. V. (1971). Accumulating evidence: Procedures for resolving contradictions among different research studies. *Harvard Educational Review*, 41, 429-471.
- Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*, 26, 86-104. DOI: 10.1016/j.cpr.2005.07.004
- Lundahl, B. W., Tollefson, D., Risser, H., & Lovejoy, M. C. (2008). A meta-analysis of father involvement in parent training. *Research on Social Work Practice*, 18, 97-106. DOI: 10.1177/1049731507309828
- Martin, J. L. R., Perez, V., Sacristan, M., & Alvarez, E. (2005). Is grey literature essential for a better control of publication bias in psychiatry? An example from three meta-analyses of schizophrenia. *European Psychiatry*, 20, 550-553. DOI: 10.1016/j.eurpsy.2005.03.011
- Maughan, D. R., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review*, 34, 267-286.

- Mays, G. P., & Hogg, R. A. (2012). Expanding delivery system research in public health settings: Lessons from practice-based research networks. *Journal of Public Health Management Practice, 18*, 485-498. DOI: 10.1097/PHH.0b013e31825f75c9
- McAuley, L., Pham, B., Tugwell, P., & Moher, D., (2000). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *The Lancet, 356*, 1228-1231.
- McCart, M. R., Priester, P. E., Davies, W. H., & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology, 34*, 527-543. DOI: 10.1007/s10802-006-9031-1
- McLeod, B. D., & Weisz, J. R. (2004). Using dissertations to examine potential bias in child and adolescent clinical trials. *Journal of Consulting and Clinical Psychology, 72*, 235-251. DOI: 10.1037/0022-006X.72.2.235
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet, 354*, 1896-1900. DOI: 10.1016/S0140-6736(99)04149-5
- Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., ... Altman, D. G., for the CONSORT Group. (2010). CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trial. *British Medical Journal, 340*, c869-897. DOI: 10.1136/bmj.c869
- Nowak, C., & Heinrichs, N. (2008). A comprehensive meta-analysis of triple p – positive parenting program using hierarchical linear modeling: Effectiveness and moderating

- variables. *Clinical Child and Family Psychology Review*, 11, 114-144. DOI: 10.1007/s10567-008-0033-0
- Pappas, C., & Williams, I. (2011). Grey literature: Its emerging importance. *Journal of Hospital Librarianship*, 11, 228-234. DOI: 10.1080/15323269.2011.587100
- Piquero, A. R., Farrington, D. P., Welsh, B. C., Tremblay, R., & Jennings, W. G. (2009). Effects of early family/parent training programs on antisocial behavior and delinquency. *Journal of Experimental Criminology*, 5, 83-120. DOI 10.1007/s11292-009-9072-x
- Posner, B. P., & Bellefontaine, S. (2011). *ReSearch* [Computer Software].
- Provos, N., McClain, J., Wang, K. (2006, November). *Search worms*. Paper presented at the 4th Workshop on Recurring Malcode, Fairfax, VA. Retrieved from http://www.citi.umich.edu/u/provos/papers/search_worms.pdf
- Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems – a meta-analytic review. *Journal of Child Psychology and Psychiatry*, 47, 99-111. DOI: 10.1111/j.1469-7610.2005.01544.x
- Roth, B. J. (2010). An academic perspective on grey literature. *Archaeologies: Journal of the World Archaeological Congress*, 6, 337-345. DOI: 10.1007/s11759-010-9141-9
- Royle, J., & Blythe, J. (1998). Promoting research utilisation in nursing: The role of the individual, organisation, and environment. *Evidence-Based Nursing*, 1, 71-72. DOI: 10.1136/ebn.1.3.71
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence-based medicine: What it is and what it isn't. *British Medical Journal*, 312, 71-72. DOI: 10.1136/bmj.312.7023.71

- Sacks, H. S., Reitman, D., Pagano, D., & Kupelnick, B. (1996). Meta-analysis: An update. *Mount Sinai Journal of Medicine*, 63, 216-224.
- Schulz, K. F., Altman, D. G., Moher, D., for the CONSORT Group. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *British Medical Journal*, 340, 698-702. DOI: 10.1136/bmj.c332
- Scullion, P.A. (2002). Effective dissemination strategies. *Nurse Researcher*, 10, 65–67.
- Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy*, 27, 171-186. DOI: 10.1016/S0005-7894(96)80013-X
- Seymour, D. J. (2010). Sanctioned inequity and accessibility issues in the grey literature in the United States. *Archaeologies: Journal of the World Archaeological Congress*, 6, 233-269. DOI: 10.1007/s11759-010-9144-6
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 36, 753-760.
- Smith, M., Glass, G., & Miller, T. (1980). *The Benefits of Psychotherapy*. Baltimore, MD: John Hopkins University Press.
- Thomas, R., & Zimmer-Gembeck, M. J. (2007). Behavioral outcomes of parent-child interaction therapy and triple p-positive parenting program: A review and meta-analysis. *Journal of Abnormal Child Psychology*, 35, 475-495. DOI 10.1007/s10802-007-9104-9
- Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology*, 55, 542-549. DOI: 0022-006X/87/J00.75

Wilson, D. B., & Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Bulletin*, 6, 413-429.

Wood, J. (2008). Methodology for dealing with duplicate study effects in meta-analysis. *Organizational Research Methods*, 11, 79-95. DOI: 10.1177/1094428106296638

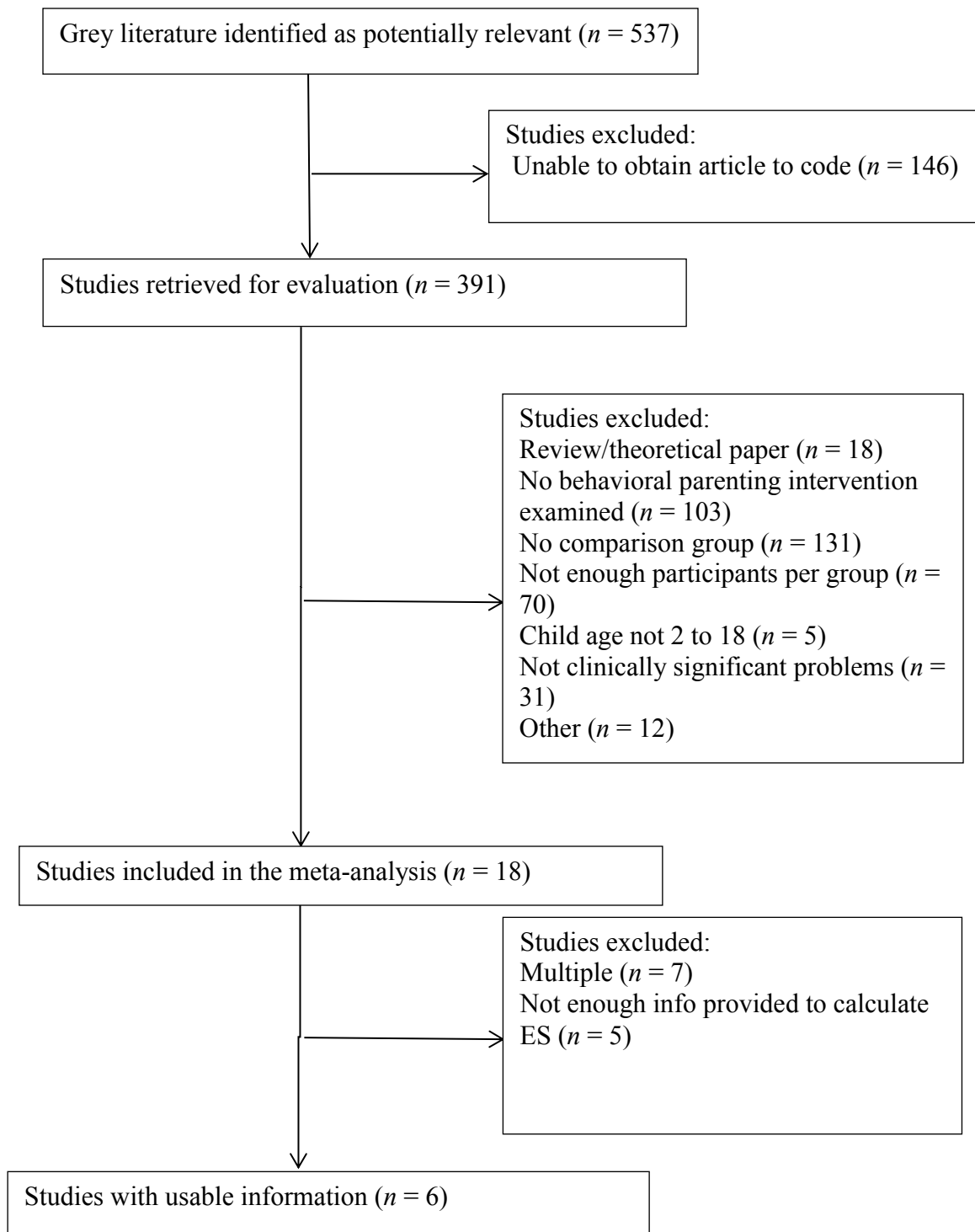


Figure 1. Flowchart of included and excluded grey literature

Discussion

This thesis is a meta-analysis of parenting interventions. My search of the literature identified 1746 potentially relevant reports, of which I obtained full-text versions of 1518 to determine whether they were appropriate for inclusion. 1423 were excluded from the present meta-analysis for a variety of reasons, including: reviews or theoretical papers ($n = 239$), not examining behavioral parenting interventions ($n = 417$), having no comparison group ($n = 422$), inadequate sample size ($n = 180$), children outside the 2 to 18 year range ($n = 12$), not examining clinically significant child behavior problems ($n = 116$), outcomes other than those being examined ($n = 13$), a behavioral parenting intervention combined with another intervention ($n = 20$), and other reasons ($n = 4$). In total, 74 articles and 21 unpublished documents met initial inclusion criteria. After coding for multiple reports on the same sample as well as for the presence of information necessary for calculating effect sizes, we retained 42 published articles and six unpublished documents. Of those, 19 (39.6%) had not been included in previous meta-analyses. Two studies were published since the most recent review (Furlong et al., 2012) and five others were unpublished reports and dissertations which were completed during the timeframe covered by previous meta-analyses. The remaining 12 studies had not been included in previous meta-analyses due to not meeting inclusion criteria of the meta-analysis (i.e., child age outside the range being examined, comparison group was other than required by the meta-analytic criteria, or reporting nonsignificant results).

All 48 included studies examined parents of children with clinically significant behavioural problems, with 54.16% including children with at least one diagnosed problem. To my knowledge, the only other meta-analysis to have such a cohesive group of studies with respect to formal diagnostic status is Furlong et al. (2012), which restricted their studies to

randomized controlled trials (RCTs) and quasi-controlled studies. The present age range of children was 2 to 18 years, whereas all but two of the previous meta-analyses (Dretzke et al., 2009; McCart et al., 2006) had narrower age ranges. It is also important to note that a number of previous studies had limitations regarding the programs that were examined (de Graaf et al., 2008; Fletcher et al., 2011; Furlong et al., 2012; Nowak & Heinrichs, 2008; Thomas & Zimmer-Gembeck, 2007), the measures that were used (de Graaf et al., 2008; Dretzke et al., 2009) or the study design (Dretzke et al., 2009; Fletcher et al., 2011; Furlong et al., 2012; Thomas & Zimmer-Gembeck, 2007). This is in contrast to the present meta-analyses, which purposely transcended each of these.

The preceding two manuscripts examined different foci of the meta-analysis I conducted. In the first manuscript I examined the overall effects of behavioural parenting interventions on parent behaviour, child behaviour, and parent adjustment outcomes. In the second manuscript I focused on the contribution of grey literature to the present meta-analysis. In this section I will review the most notable findings of each manuscript and highlight their contributions, limitations, and implications for future research.

Main Effects and Moderator Analyses

Effects were found for behavioural parenting interventions on all outcomes, confirming and extending the findings of previous meta-analyses which reported effects on the most proximal outcome to behavioural parenting interventions, parent behaviour (Kaminski et al., 2008; Lundahl et al., 2006; Lundahl, et al., 2008; Nowak & Heinrichs, 2008), as well as on child externalizing behaviour (Lundahl et al., 2006; Lundahl et al., 2008; Maughan et al., 2005; McCart et al., 2006; Nowak & Heinrichs, 2008; Piquero et al., 2009), and on parent adjustment outcomes (McCart et al., 2006; Nowak & Heinrichs, 2008).

In previous meta-analyses researchers have generally found significant heterogeneity with respect to parent behaviour (de Graaf et al., 2008; Fletcher et al., 2011; Kaminski et al., 2008; Lundahl et al., 2006; Nowak & Heinrichs, 2008), child behaviour (Kaminski et al., 2008; Lundahl et al., 2006; Maughan et al., 2005; Nowak & Heinrichs, 2008; Piquero et al., 2009), and parent adjustment outcomes (Nowak & Heinrichs, 2008). Child age has previously been examined as a possible moderator with regards to all three outcome variables, but has demonstrated mixed results (de Graaf et al., 2008; Lundahl et al., 2006; Maughan et al., 2005; Nowak & Heinrichs, 2008). Child formal diagnostic status has been found to moderate child behaviour outcomes (Lundahl et al., 2006), but not parent behaviour outcomes (de Graaf et al., 2008; Lundahl et al., 2006). This variable has not previously been examined with respect to parent adjustment outcomes. Similarly, methodological quality has been found to be a significant moderator of both parent and child behaviour outcomes (Nowak & Heinrichs, 2008) but its role as a moderator has not been explored with respect to parent adjustment. Despite publication status having not been previously found to be a significant moderator of child behaviour outcomes (Maughan et al., 2005; Piquero et al., 2009), its possible moderating role with respect to parent behaviour and parent adjustment have not been explored. Furthermore, experimenter allegiance and comparison group have not been previously examined as a possible moderator of any of these three outcomes.

Prior to testing for moderator variables, I examined the homogeneity of effects for each type of outcome variable using the Q statistic. Notably, most conditions were found to exhibit significant homogeneity, which has not been found in many previous meta-analyses of behavioural parenting interventions. For those outcome conditions which exhibited significant heterogeneity (i.e., parent behaviour and parent adjustment, both when compared to a no-

treatment comparison group), I examined the role of possible moderators for those outcomes. A variety of procedures have been used for examining moderators in previous meta-analyses, and although it is generally accepted that moderators only be examined when significant heterogeneity is found (e.g., de Graaf et al., 2008; Kaminski et al., 2008; Lundahl et al., 2006; Maughan et al., 2005; McCart et al., 2006; Norcross, Krebs, & Prochaska, 2011; Nowak & Heinrichs, 2008; Piquero et al., 2009), in past meta-analyses moderator analyses have been conducted without a prior test for homogeneity (e.g., Serketich & Dumas, 1996) or only conducting after-the-fact tests of homogeneity (e.g., Reyno & McGrath, 2006).

Parent behaviour.

In this meta-analysis, significant homogeneity was found for parent behaviour outcomes in those studies that compared behavioural parenting interventions with an active comparison group. This finding echoes those of Lundahl et al. (2008) who found evidence of homogeneity when examining parent behaviour outcomes. It is important to note that the sample of studies that compared behavioural parenting interventions to an active comparison group was quite small, and therefore may not have been robust enough to detect differences. However, when behavioural parenting interventions were compared to a no-treatment comparison group, parent behaviour outcomes were found to be heterogeneous, a finding which is consistent with previous meta-analyses (i.e., de Graaf et al., 2008; Fletcher et al., 2011; Lundahl et al., 2006; Kaminski et al., 2008; Nowak & Heinrichs, 2008). One variable was found to moderate the effects of behavioural parenting interventions on parent behaviour outcomes

Child age, when examined continuously, was a significant moderator of parent behaviour outcomes, a finding which was contrary to the conclusions of the majority of previous meta-analyses on behavioural parenting interventions (i.e., de Graaf et al., 2008; Lundahl et al., 2006).

It therefore seems as though age effects are too sensitive to be examined categorically, as doing so appears to mask differential effects. According to my findings, as child age increases, so do effect sizes. It is important to note that child age only increased to 9 years, and so these findings cannot be generalized to children and adolescents over the age of 10. My findings with respect to child age partially support those of Nowak and Heinrichs (2008) who, in examining age as a continuous variable, found age effects. However, Nowak and Heinrichs (2008) found negative effects such that younger children had greater effects than older children. It may be that the particular intervention program examined by Nowak and Heinrichs (2008) may be more effective for younger children than older. It is important to note that their sample included children aged 0 to 14, so it may be that the eldest of these children skewed the results so that the interventions appeared less effective for older children. It is also noteworthy that the great majority of children examined in that meta-analysis did not have clinically significant or diagnosed problems (Nowak & Heinrichs, 2008). Therefore, one possibility is that for children with clinically significant problems, parents of older children derive greater benefit with regards to their parenting behavior from such interventions than parents of younger children, but that the opposite may be true for prevention or early intervention programs. This is supported by the idea that prevention and early intervention programs should be initiated as early as possible to derive the greatest benefits (Durlack, 1998) and suggests that program implementers need to be cognizant of child age when they are looking to implement either a prevention/early intervention-based program or a treatment program so as to ensure that participants derive the greatest benefits from such interventions.

Child formal diagnostic status did not moderate parent behaviour outcomes. This finding supports the findings of previous meta-analyses (de Graaf et al., 2008; Lundahl et al., 2006) and

suggests that parenting interventions are efficacious across the range of disruptive behaviour disorders. It is important to note that the sample of studies examined by de Graaf et al. (2008) included very few children with formally diagnosed problems, and although Lundahl et al. (2006) examined such outcomes, their sample included no children with formal diagnoses. In the present sample, although children with diagnosed problems were studied, they were often studied alongside children with undiagnosed problems, thus potentially masking the true moderating effects of child formal diagnostic status on parenting behaviour. It may be that effects of child formal diagnostic status were overshadowed by the many children with clinically significant but undiagnosed problems, and may suggest that children with diagnosed problems may represent a different clinical entity than those that are not diagnosed. Future research should attempt to tease apart these groups so as to determine whether parenting behaviour may be moderated by formal diagnostic status when looking at pure samples.

Most surprisingly, methodological quality was not a significant moderator of either outcome (i.e., parent behaviour or parent adjustment) for which moderating effects were examined. It may be that the inclusion criteria used to select studies may have restricted the range of methodological quality ratings. Methodological quality ratings in the present meta-analysis ranged from one to seven. However, only five studies had ratings between one and three, 29 had ratings between six and seven out of seven. In particular, selection for study design features (i.e., the presence of a comparison, particular minimum group sizes) may have unwittingly created a more homogeneous group in terms of methodological quality. One previous meta-analysis (i.e., Nowak & Heinrichs, 2008) found methodological quality to be a significant moderator of parent behaviour outcomes, but it is possible that the studies included were more variable in methodological quality. It is important to note that the studies included in

that meta-analysis were limited to one particular intervention program (i.e., Triple P) and that no limits were put on sample size or design. Furthermore, Nowak & Heinrichs (2008) examined methodological quality using the Downs and Black (1998) scale, which focuses a great deal on the reporting of information. By contrast, the measure of methodological quality I developed in this meta-analysis focused on group randomization and the quality of the measures used. It may be that, had I used a more elaborate measure of methodological quality, I may have found methodological quality to be a significant moderator of parent behaviour outcomes. Future meta-analysts and readers of the body of literature on behavioural parenting interventions may thus benefit from verifying that minimum quality standards are being followed to ensure that the conclusions being put forth are based on sound research.

Contrary to predictions based on findings of researcher allegiance effects reported by other studies (e.g., Eisner, 2009; Miller, Wampold, & Varhely, 2008), those studies with experimenter allegiance did not differ from those with no allegiance. It is notable that a number of the studies with experimenter allegiance included comorbid samples. It has been shown that comorbidity presents a complicating factor to treatment and generally leads to worse prognoses (Connor & Doerfler, 2009; Greene et al., 2002). Such problem complexity may have reduced the overall effects of behavioural parenting interventions and masked any moderating effects of experimenter allegiance. Due to the low reliability in the coding this variable, it is also important that these results are interpreted with caution. Greater consistency in coding this variable may have identified support for researcher allegiance bias, and should be further examined in future studies.

Finally, publication status did not differentially affect behavioural parenting interventions on parent behaviour outcomes. This finding is inconsistent with the concept of publication bias

(Staines & Cleland, 2007), such that those studies that were published had significantly larger effect sizes compared to those that were unpublished. As previously mentioned, the use of several methodological characteristics as inclusion criteria may have unwittingly led to the exclusion of a number of unpublished studies which were unable to fulfill those criteria and led to the selection of a more cohesive group of studies overall; published or unpublished. I found nonsignificant differences in methodological quality between published and unpublished studies, suggesting that the quality of such studies does not account for the noted effect size differences. This is the first time that publication status has been examined with respect to parent behaviour outcomes, and thus would benefit from replication by future meta-analyses.

Child behaviour.

Child behaviour outcomes, both when compared to no-treatment and active comparison groups, were found to be homogenous, which is consistent with the findings of Lundahl et al. (2008), but which contradicts the findings of a number of other meta-analyses in this area (e.g., Kaminski et al., 2008; Lundahl et al., 2006; Maughan et al., 2005; McCart et al., 2006; Nowak & Heinrichs, 2008; Piquero et al., 2009). In attempting to explain the discrepant findings, it may be that inclusion of children with only clinically significant problems may have led to this sample being homogeneous in terms of child behaviour outcome, whereas many previous study samples examined groups of children with varying levels of problems. Thus, children with clinically significant and diagnosed problems may represent a homogeneous group of children who evidence similar effects of behavioural parenting interventions on their levels of externalizing behaviours. It may therefore be that combining children with clinical and nonclinical problems decreases the homogeneity of a group of children with respect to externalizing behavior outcomes. Many of the previous meta-analyses in this area that found heterogeneity in child

behaviour outcomes (e.g., Kaminski et al., 2008; Lundahl et al., 2006; McCart et al., 2006; Nowak & Heinrichs, 2008) examined both clinical and nonclinical samples of children. It is reasonable to conclude that the current meta-analysis indicates that the effects of behavioural parenting programs are robust in terms of the child behaviour outcomes among sample of children with clinically significant problems.

Parent adjustment.

Homogeneous effects were found for parent adjustment outcomes when behavioural parenting interventions were compared to an active comparison group. This is supported by the findings of Furlong et al. (2012) who found no evidence of heterogeneity for parent adjustment outcomes when behavioural parenting programs were compared to both no-treatment and active comparison groups. However, effects were found to be heterogeneous when behavioural parenting interventions were compared to a no-treatment group. It may be that active comparison groups offer a more similar group of conditions so as to create more unified experience for those whereas active comparison groups may vary more widely in the types of comparison groups they offer and are thus less cohesive. This finding lends partial support to the results of a previous meta-analysis (i.e., Nowak & Heinrichs, 2008), who found parent adjustment outcomes to be heterogeneous. None of the methodological variables I examined moderated the effects of behavioural parenting interventions on parent adjustment. It is clear that such outcomes are unaffected by the methodological (comparison group, publication status, methodological quality, and experimenter allegiance) variables I examined. It is notable that, similar to the results found for parent behaviour outcomes, methodological quality of the studies did not affect parent adjustment outcomes. It still remains unclear as to what variables might account for the significant heterogeneity of effects in parent adjustment outcomes. One possible explanation may

be the variety of different measures which were combined to look at parent adjustment outcomes. These included measures of stress, parenting stress, general health, and depression. However, when this was examined, effects were found to be similar for the differing measures. It is also possible that the samples in different studies had varying levels of distress, leading them, as a whole to be heterogeneous.

Child age, examined continuously, was found to be a significant moderator of parent adjustment outcomes such that as child age increased parent adjustment effects also increased. The moderating effect of child age on parent adjustment has not been previously examined in meta-analyses. It may be that younger and older children represent conceptually different groups, with younger children with clinically significant problems representing an early onset group. Research has shown that those children who develop conduct problems earlier tend to have worse prognoses and increased conduct problems (Loeber, 1991) and have a number of other stressors present in the environment (McCabe et al., 2001). Therefore, those children who are demonstrating clinically significant conduct problems at a later age may have parents who are more well-adjusted and thus more likely to experience change in their level of distress. Thus it may be important for practitioners who are working with parents of younger, clinically distressed, children to have add-on supplements to programs that are specifically targeted at addressing both the increased stress these parents may be experiencing as a result of their child's difficulties, as well as to aid them to cope more effectively with general life stressors which may be impairing their ability to parent in a positive way.

With regards to child formal diagnostic status, those studies including families of children diagnosed with ADHD were found to have greater effects of behavioural parenting interventions on parent adjustment than those studies where children were not diagnosed with

ADHD. It may be that the particular intervention strategies employed in behavioural parenting programs are particularly effective for addressing the concerns and thus improving the well-being of parents of children with ADHD. This is supported by previous research that found both a reduction in ADHD symptoms and an increase in maternal well-being with attending a behavioural parenting program (Sonuga-Barke et al., 2001).

In sum, it appears as though the child variables I examined were significant moderators of parent behaviour (child age) and parent adjustment (child age and ADHD diagnosis) outcomes when behavioural parenting interventions were compared to no-treatment groups. The homogeneity found for child behaviour outcomes, as well as parent behaviour and adjustment when compared to an active comparison group, indicates that the inclusion criteria I set yielded a cohesive set of studies in which families who experienced robust effects of behavioural parenting interventions. Finally, it remains to be determined whether unexplored moderators are responsible for any unaccounted for variability in parent adjustment and behaviour outcomes. These results would benefit from replication in future studies to support these conclusions.

Grey Literature

With calls for inclusion of grey literature in meta-analysis (McAuley et al., 2000), debate about its quality (McAuley et al., 2000), and the dearth of grey reports available through the internet, I sought to identify the most comprehensive body of grey literature on behavioural parenting interventions. This was undertaken to properly assess the contribution of grey literature to the conclusions of meta-analyses on parenting interventions.

In the second manuscript, I focused on the contribution of grey literature to meta-analysis in the current meta-analysis of behavioural parenting interventions. In that manuscript, I examined how much grey literature was found, how many reports were duplicated in the

published literature, looked at whether effect sizes differed for the published and grey literature, and also examined whether there were differences in the methodological quality of grey versus published literature. In total, 537 grey reports were identified, with 18 of those meeting the inclusion criteria of the meta-analysis. Of those, seven were identified as being multiple reports on the same dataset, either duplicating data that was later published, or other grey reports that were identified. A further five documents did not contain sufficient information to allow me to calculate effect sizes. Therefore, the final sample included six grey reports. Five of these six reports were dissertations which could have been located without the use of time-consuming search strategies. Similar effect sizes were found for published and unpublished literature for all outcomes. Notably, no significant difference was found between the methodological quality of unpublished and published literature, contrary to the theoretical concerns raised by researchers such as Roth (2010), Pappas and Williams (2011), and Seymour (2010). However, a large portion of the unpublished studies I included were masters theses and doctoral dissertations, which is in contrast to the grey literature sources identified by these researchers, including unpublished clinical trials, conference proceedings, unpublished studies, reports by agencies and private institutions. Dissertations have been shown to be of better quality and are rigorously reviewed (McLeod & Weisz, 2004). It is possible that if more unpublished studies had been reports from mental health agencies, unpublished trial data and conference proceeding, my findings may have supported the idea of an overestimation of effect sizes which was previously hypothesized. It is notable that these concerns were raised with regards to literature in fields such as Anthropology (Roth, 2010), Archaeology (Seymour, 2010), and medical research (Pappas & Williams, 2011). It is possible that as a large amount of research in these areas are conducted by private institutions and not published in traditional peer-reviewed journals, these bodies of

literature may be more susceptible to differences in the quality of literature disseminated through a peer-reviewed journal and nonpeer-reviewed sources, than that of behavioural parenting interventions. However, my stringent inclusion criteria may have created very limited opportunity for methodologically poor studies to be included in my meta-analysis. Also, it is important to note that these concerns were not based on research findings, but were rather theoretical in nature, and so it will be important to determine whether my nonsignificant findings extend to other areas in psychology and, possibly, other disciplines.

Unique Contributions of this Meta-Analysis

The current meta-analysis extends the literature by the inclusion of 19 studies which have not previously been included in meta-analyses. These studies include those that have been published since the most recent meta-analysis, those that extend the age range, those that extended the types of comparison groups used, and studies with nonsignificant results. It also includes six unpublished reports which were not included in previous meta-analyses of behavioural parenting interventions.

The results of this meta-analysis both supported the findings of previous studies as well as extended these findings in several important ways. First, I examined the findings across childhood and teenage years in order to examine the effects of behavioural parenting interventions over child age so as to examine these effects comprehensively and determine whether there are age-related effects. I found evidence of age-related effects on both parent behaviour and parent adjustment, suggesting that parents of older school-age children with clinically significant behaviour problems seem to be deriving greater benefits from such interventions than parents of younger children. Second, I conducted a priori tests of homogeneity that have not been consistently undertaken in the past (e.g., Dretzke et al., 2009; Serketich &

Dumas, 1996; Thomas & Zimmer-Gembeck, 2007) or were only conducted after examining moderators (Reyno & McGrath, 2006). In some cases I confirmed previous moderator findings (e.g., with respect to parent behaviour and parent adjustment) and in other cases (i.e., with respect to child behaviour, as well as parent behavior and parent adjustment when compared to an active comparison group) I determined that such analyses were not warranted. I also examined previously unexamined moderator variables with respect to parent behaviour outcomes: comparison group, experimenter allegiance, and publication status. My findings with respect to these variables provided no support for their role as moderators of the most proximal outcome variable of behavioural parenting interventions – parent behaviour.

I conducted an in-depth examination of grey literature, by employing generic search engines such as Google and Yahoo. In my search for grey literature, I expanded those searches used in previous meta-analyses of behavioural parenting interventions, such as Proquest Dissertations and Theses and, in one instance each, GoogleScholar and a limited use of Google. GoogleScholar was not searched due to its focus on yielding published articles which were accessible through a number of other search strategies used. However, the use of Proquest Dissertations and Theses yielded a number of useful results with very little effort required. Finally, I was involved in the development of an innovative program (ReSearch; Posner & Bellefontaine, 2011) to facilitate searches using the Yahoo interface. This considerably reduced the amount of time spent conducting searches using that search engine by removing duplicate search results, as well as putting results in order from most to least relevant.

Limitations

One limitation of this study is that, in error, the sources of approximately half of the grey literature studies found were not recorded. It is thus impossible for me to determine the extent to

which the use of basic search engines may have added to included grey literature in this meta-analysis. It is notable, however, that one included report, four reports which had no data for calculating effect sizes, as well as five reports that were identified as multiples, were not dissertations or book chapters and are likely to have been identified using a basic search engine. Thus, although it appears as though searches of Google and Yahoo may have added some novel sources, future research should formally examine the actual yield of these sources.

Unfortunately a number of articles, books, theses or conference proceedings (226; 13% of those resources that were identified) could not be obtained to adequately assess for inclusion into the present meta-analysis. Four of these were requested through the interlibrary loan service, but have not been received as of over one year after the request was made. The rest were not obtainable through the interlibrary loan service, or had incorrect citations which could not be corrected by internet searches, many of which came from non-North American sources, and were unable to be located using generic search engines such as Google or Yahoo. 137 of these were dissertations which were identified primarily (68.6%) using PsycINFO and Proquest Dissertations and Theses. Many were from unidentified or non-Canadian universities and 129 of them had been defended at least 15 years prior. Although most meta-analyses examined did not examine such data, one previous meta-analysis (Dretzke et al., 2009) did report that they were unable to obtain 9% of the articles they had identified. Although my analyses suggested no notable bias in the results which may stem from these inaccessible resources, this certainly does point to a gap in researchers' ability to access all reports.

As with all reviews and meta-analyses, the current findings have a limited shelf-life and should be updated in the future to incorporate new studies. Many meta-analysts and reviewers struggle with the question of when a review or meta-analysis is due for an updating. Despite the

proposal of involved methods of determining whether null meta-analyses are ready to be updated (e.g., Barrowman, Fang, Sampson, & Moher, 2003), research by Shojania et al. (2007) found that the need to update may be as early as after one year. This is likely warranted given that a number of the failsafe Ns calculated in this meta-analysis were relatively low.

Implications for Evidence-Based Services for Parenting

Meta-analyses have long been recognized as a very important tool in making clinical decisions within the field of evidence-based practice (e.g., Nestoriuc, Kriston, & Rief, 2010). My main contribution to the area of evidence-based behavioural parenting interventions is the provision of a meta-analysis with a heterogeneous sample of children clinically significant problems. This will allow practitioners and parents who wish to provide such services to such children with a clear idea of the benefits and limits of such interventions, and what characteristics may or may not impact outcomes.

Further, this meta-analysis provided evidence that behavioural parenting interventions, in addition to their impact on parent and child behaviours, have a significant impact on parent adjustment. This outcome has only recently been examined in one systematic review (Barlow, Smailagic, Huband, Roloff, & Bennett, 2012) and two previous meta-analyses, one focused on Triple P behavioural parenting interventions (Nowak & Heinrichs, 2008) and the other on RCTs and quasi-controlled studies of group parenting interventions (Furlong et al., 2012).

Implications for Knowledge Translation

For knowledge translation to work best, key stakeholders (e.g., policy-makers, shareholders, practitioners, and consumers) need to receive information quickly and in accessible ways, which can include strategies such as conference presentations, brief reports, and workshops (Barwick et al., 2005). Some researchers have estimated that the implementation of

research into clinical practice can take 10 to 30 years (Fitzpatrick, 2008). The advent of the Internet means that information on evidence-based practice is more readily available than ever before, but key stakeholders, and particularly consumers, need to be able to discern which information is of good quality in order to make good, informed decisions (Barwick et al., 2008). A survey of executive directors and practitioners involved with Children's Mental Health Ontario by Barwick et al. (2008) revealed that less than 40% of the organizations surveyed had access to academic library evidence bases. This implies that over half of respondents do not have a way to easily access current research findings in order to base their policies and practices on, leaving them to rely on potentially outdated practices or knowledge. The National Institute for Health Research's Centre for Research Dissemination has created a new prospective register of systematic reviews (PROSPERO; <http://www.crd.york.ac.uk/prospero/index.asp>) in order to partially address this problem. PROSPERO is intended to provide a singular place where international researchers can register their systematic reviews and meta-analyses in order to make it easier to locate such resources. However, it remains that once such resources have been located, their quality should be assessed.

Psychologists are trained to critically examine the quality of research. However, many practitioners have great difficulty in assessing the quality of meta-analyses, due at least in part to reporting practices, despite their reliance on their conclusions for best practice recommendations (Liberati et al., 2009). The Quality of Reporting of Meta-Analyses [QUOROM] statement (Moher et al., 1999) was developed in order to attempt to address these concerns and improve the reporting standards of meta-analyses. More recently, these recommendations of reporting standards were revised and expanded, and are now known as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses [PRISMA] statement (Moher, Liberati, Tetzlaff, &

Altman, The PRISMA Group, 2009). Such reporting standards offer suggested ways of, for example, identifying studies as meta-analyses, presenting the protocol, eligibility criteria, sources, search strategy, screening process used for identifying studies, as well as how results should be presented. Shercliffe, Stahl, and Tuttle (2009) also support the use of reporting standards similar to those suggested by PRISMA to allow researchers consulting meta-analyses and systematic reviews to be able to properly assess the methodological quality of such research overviews.

Implications for Grey Literature

In its ability to be quickly produced, grey literature seems like a natural fit for quick dissemination of up-to-date research so as to be able to implement findings into practice in a timely manner. This would necessitate an adaptation of the peer-review process in order to determine the quality of such studies, but would greatly improve access to some interesting, quality studies. As of yet, finding easily accessible, reputable sources of grey literature remains difficult. This prevents many people from going to the grey literature in order to find the most up-to-date information. The development of a database for grey literature has been suggested by a number of researchers in order to address this concern (e.g., Okoroma, 2011). Launched in 2011, OpenGrey (<http://www.opengrey.eu>) has attempted to address the demand for a grey literature repository in the European Union and currently contains over 690,000 citations. Although OpenGrey has partnered with other similar sites in different parts of the world (e.g., GreyNet; <http://www.greynet.org>), no such repositories currently exist for North American grey literature.

Frater, Myohanen, Taylor, and Keith (2007) suggested an alternative way to address grey literature that makes it theoretically as accessible as traditional literature. They suggested

registering grey literature with Digital Object Identifiers in order to have documents be available in perpetuity (Frater et al., 2007). This may facilitate the suggestions of researchers such as Banks (2010) who have advocated for the integration of access to grey literature into traditional databases which researchers typically consult. Banks (2010) advocated for “continuum of evidence” which would allow grey works the opportunity to become as well-regarded as peer-reviewed journal articles.

Searching for grey literature using generic search engines presents challenges. Google and Yahoo both produced an overwhelming number of search results, despite my attempts to limit those to the most relevant through the pairing of search terms. Most of the search results returned by these were not relevant to the present research question. The development and subsequent use of the ReSearch program (Posner & Bellefontaine, 2011) led to both an overall reduction in the number of unique results returned and an increase in the relevance of those results. Such search programs are relatively easy to develop and make the task of searching for grey literature in these search engines much easier and less time-consuming. It may be of use for meta-analysts and systematic reviewers to evaluate the utility of such programs in conducting searches.

My research showed that most grey literature is eventually published in peer-reviewed sources, calling into question whether the search for grey literature sources is necessary in meta-analyses. The difficulty in locating as well as the time-consuming nature of such searches to locate grey literature make it so that many meta-analysts prefer to simply not include such resources (Coad et al., 2006; Roth, 2010). However, reviews of newer research areas (e.g., Benzies et al., 2006) suggest that grey literature may be of significant benefit in newer research areas where much literature might not be published as of yet. In the case of research on

behavioural parenting interventions, an area which has been widely studied and published, my research showed no real benefit to including such literature.

References

- Abby, M., Massey, M. D., Galandiuk, S., & Polk, Jr., H. C. (1994). Peer review is an effective screening process to evaluate medical manuscripts. *Journal of the American Medical Association*, 272, 105-107. DOI:10.1001/jama.1994.03520020031008
- Abidin, R. R. (1995). *Parenting Stress Index: Professional Manual* (3rd. ed.). Odessa, FL: Psychological Assessment Resources.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Ahmed, I., Sutton, A. J., & Riley, R. D. (2012). Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: A database survey. *BMJ*, 344. DOI: <http://dx.doi.org/10.1136/bmj.d7762>
- American Academy of Child and Adolescent Psychiatry [AACAP]. (1997). Practice parameter for the assessment and treatment of children and adolescents with conduct disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36, 122S-139S.
- AACAP. (2006). Policy statement on Evidence-based Practice. Retrieved September 10, 2012 from http://aacap.org/cs/root/policy_statements/evidence_based_practice.
- AACAP. (2007a). Practice parameter for the assessment and treatment of children and adolescents with oppositional defiant disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46, 126-141.
- AACAP. (2007b). Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46, 894-921.

- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text revision). Washington, DC: Author.
- American Psychological Association. (2005). *Policy statement on evidence-based practice in psychology*. Retrieved from <http://www.apa.org/practice/resources/evidence/evidence-based-statement.pdf>
- American Psychological Association. (APA). (2010). *Publication manual of the American Psychological Association*. Washington, DC: APA.
- American Psychological Association, Task Force on the Future of Psychology. (2009). *2009 Presidential task force on the future of psychology practice final report*. Retrieved from <http://www.apa.org/pubs/info/reports/future-practice.pdf>
- Anastopoulos, A. D., Shelton, T. L., & Barkley, R. A. (2005). Family-based psychosocial treatments for children and adolescents with Attention-Deficit/Hyperactivity Disorder. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for children and adolescent disorders: Empirically based strategies for clinical practice (Second edition)*, pp.327-350. Washington, DC: American Psychological Association.
- *Anastopoulos, A. D., Shelton, T. L., DuPaul, G. J., & Guevremont, D. C. (1993). Parent training for attention-deficit hyperactivity disorder: its impact on parent functioning. *Journal of Abnormal Child Psychology*, 21, 581-596.
- APA Publications and Communications Board Working Group on Journal Article Reporting Standards (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist*, 63, 839-851. DOI: 10.1037/0003-066X.63.9.839
- Auger, C. P. (1998). *Information sources in grey literature* (4th ed.). London: Bowker-Saur.

Bandura, A. (1971). *Social learning theory*. New York: General Learning Press.

Banks, M. A. (2010). Blog posts and tweets: The next frontier for grey literature. In D. J. Farace, & J. Schopf (Eds.), *Grey literature in library and information studies*. Berlin: De Gruyter Saur.

Barkley, R. A. (1987). *Defiant children*. New York: Guilford Press.

*Barkley, R. A., Guevremont, D. C., Anastopoulos, A. D., & Fletcher, K. E. (1992). A comparison of three family therapy programs for treating family conflicts in adolescents with attention-deficit hyperactivity disorder. *Journal of Consulting and Clinical Psychology*, 60, 450-462. DOI: 10.1037/0022-006X.60.3.450

*Barkley, R. A., Shelton, T. L., Crosswait, C., Moorehouse, M., Fletcher, K., Barrett, S., ... Metevia, L. (2000). Multi-method psycho-educational intervention for preschool children with disruptive behavior: Preliminary results at post-treatment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41, 319-332. DOI: 10.1023/A:1005144304071

Barlow, J., Smailagic, N., Huband, N., Roloff, V., & Bennett, C. (2012). Group-based parent training programmes for improving parental psychosocial health. *Cochrane Database of Systematic Reviews*, 6. DOI: 10.1002/14651858.CD002020.pub3

Barrowman, N. J., Fang, M., Sampson, M., & Moher, D. (2003). Identifying meta-analyses that are ripe for updating. *BMC Medical Research Methodology*, 3, 13. DOI:10.1186/1471-2288-3-13

Barwick, M. A., Boydell, K. M., Stasiulis, E., Ferguson, H. B., Blase, K., & Fixsen, D. (2005). *Knowledge transfer and implementation of evidence-based practices in children's mental health*. Toronto, ON: Children's Mental Health Ontario.

Barwick, M., Boydell, K. M., Stasiulis, E., Ferguson, H. B., Blase, K., Fixsen, D. (2008).

Research utilization among children's mental health providers. *Implementation Science*, 3, 19. DOI: 10.1186/1748-5908-3-19

Baskin, T. W., Tierney, S. C., Minami, T., & Wampold, B. E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology*, 71, 973-979.

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology*, 1, 311-320.

Beck, A. T., & Steer, R. A. (1990). *Manual for the Beck Anxiety Inventory*. San Antonio, TX: Psychological Corporation.

Bellefontaine, S., & Lee, C. M. (2013a). *A meta-analytic examination of behavioral parent training in the treatment of children's disruptive behavior*. Manuscript submitted for publication.

Bellefontaine, S., & Lee, C. M. (in press). *Between black and white: Examining grey literature in meta-analyses of psychological research*.

Benzies, K. M., Premji, S., Hayden, K. A., & Serrett, K. (2006). State-of-the-evidence reviews: Advantages and challenges of including grey literature. *Worldviews on Evidence-based Nursing*, 3, 55-61. DOI: 10.1111/j.1741-6787.2006.00051.x

*Bor, W., Sanders, M. R., & Markie-Dadds, C. (2002). The effects of the triple p-positive parenting program on preschool children with co-occurring disruptive behavior and attentional/hyperactive difficulties. *Journal of Abnormal Child Psychology*, 30, 571-587.

- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive Meta-analysis Version 2*. Englewood, NJ: Biostat.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to Meta Analysis*. United Kingdom: John Wiley & Sons Ltd.
- *Braet, C., Meerschaert, T., Merlevede, E., Bosmans, G., Van Leeuwen, K., & De Mey, W. (2009). Prevention of antisocial behaviour: Evaluation of an early intervention programme. *European Journal of Developmental Psychology*, 6, 223-240. DOI: 10.1080/17405620601033194
- Brauner, C. B., & Stephens, C. B. (2006). Estimating the prevalence of early childhood serious emotional/behavioral disorders: Challenges and recommendations. *Public Health Reports*, 121, 303-310.
- Briesmeister, J. M., & Schaefer, C. E. (Eds.). (2007). *Handbook of parent training: Helping parents prevent and solve problem behaviors* (3 ed.). Hoboken, NJ: John Wiley and Sons.
- Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review*, 26, 17-31. DOI: 10.1016/j.cpr.2005.07.003
- Centre for Reviews and Dissemination [CRD]. (2008). *Systematic reviews: CRD's guidance for undertaking reviews in health care*. York, UK: CRD, University of York.
- Chalmers (1993). The Cochrane Collaboration: Preparing, maintaining, and disseminating systematic reviews of the effects of health care (pp. 153-163). In K. S. Warren, & F. Mosteller (Eds.), *Doing More Good Than Harm: The Evaluation of Health Care Interventions*. New York, NY: Annals of the New York Academy of Sciences Publishers.
- Chamberlain, P., & Smith, D. K. (2003). Antisocial behavior in children and adolescents:

- The Oregon multidimensional treatment foster care model. In A. E. Kazdin & J. R. Weisz (Eds.) *Evidence-based psychotherapies for children and adolescents* (pp. 282-300). New York: Guilford.
- Children's Mental Health Ontario (2001, May). *Evidence based practices for conduct disorder in children and adolescents*. Toronto, ON: Author.
- Children's Mental Health Ontario (2008, January). *Evidence-based practice consultation paper*. Toronto, ON: Author.
- Clark, E., Donovan, E., & Schoettker, P. (2006). From outdated to updated, keeping clinical guidelines valid. *International Journal for Quality in Health Care*, 18, 165-166.
- Clarke, M., & Oxman, A. D. (Eds.). (2003). *The Cochrane Reviewers Handbook 4.2.0* [updated March 2003]. Oxford: Cochrane Collaboration.
- Coad, J., Hardicare, J., & Devitt, P. (2006). Searching for and using grey literature. *Nursing Times*, 102, 35-36.
- Conn, V. S., Valentine, J. C., Cooper, H. M., & Rantz, M. J. (2003). Grey literature in meta-analyses. *Nursing Research*, 52, 256-261.
- *Connolly, L., Sharry, J., & Fitzpatrick, C. (2001). Evaluation of a group treatment programme for parents of children with behavioural disorders. *Child & Adolescent Mental Health*, 6, 159-165. DOI: 10.1017/S1360641701002714
- Connor, D. F., & Doerfler, L. A. (2009). Attention-Deficity/Hyperactivity Disorder and comorbid Oppositional Defiant Disorder or Conduct Disorder. *Current Attention Disorder Report*, 1, 5-11.
- Cook, D. J., Greengold, N. L., Ellrodt, A. G., & Weingarten, S. R. (1998). The relation between systematic reviews and practice guidelines. In C. D. Mulrow and D. Cook (Eds.),

Systematic Reviews: Synthesis of Best Evidence for Health Care Decisions. Philadelphia: American College of Physicians.

Cook, D.J., Guyatt, G.H., Ryan, G., Clifton, J., Buckingham, L., Willan, A., ... Oxman, A. D. (1993). Should unpublished data be included in meta-analyses? Current convictions and controversies. *The Journal of the American Medical Association*, 269, 2749-2753. DOI: 10.1001/jama.1993.03500210049030

*Costin, J., Lichte, C., Hill-Smith, A., Vance, A., & Luk, E. (2004). Parent group treatments for children with oppositional defiant disorder. *Australian e-journal for the Advancement of Mental Health*, 3, 1-8.

Council of Editors. (2011). Summary report of journal operations, 2010. *American Psychologist*, 66, 405-406. DOI: 10.1037/a0023978

CPA Task Force on Evidence-Based Practice of Psychological Treatments. (2012). *Evidence-based practice of psychological treatments: A canadian perspective*. Retrieved from http://www.cpa.ca/docs/file/Practice/Report_of_the_EBP_Task_Force_FINAL_Board_Aproved_2012.pdf

Craig, J. V., & Smyth, R. L. (Eds.). (2002). *The evidence-based practice manual for nurses*. London: Elsevier Limited.

*Cruz, M. (2009). *Behavior parent training and fathers of young children with ADHD* (Unpublished doctoral dissertation). University at Buffalo, State University of New York, New York.

Cunningham, C. E., Bremner, R., & Secord-Gilbert, M. (1998). *The community parent education (COPE) program: A school-based family systems oriented course for parents of children with disruptive behavior disorders*. Unpublished manual, McMaster University and

Chedoke-McMaster Hospitals.

Davidson, L. A., & Douglas, K. (1998). Digital object identifiers: Promise and problems for scholarly publishing. *Journal of Electronic Publishing*, 4. DOI: [10.3998/3336451.0004.203](https://doi.org/10.3998/3336451.0004.203)

Davies, P. T., & Boruch, R. (2001). The Campbell Collaboration. *British Medical Journal*, 323, 294-295.

Decullier, E., Lhéritier, V., & Chapuis, F. (2005). Fate of biomedical research protocols and publication bias in France : Retrospective cohort study. *BMJ*, 331, 19. DOI: <http://dx.doi.org/10.1136/bmj.38488.385995.8F>

De Graaf, I., Speetjens, P., Smit, F., de Wolff, M., & Tavecchio, L. (2008). Effectiveness of the Triple P positive parenting program on parenting: A meta-analysis. *Family Relations*, 57, 553-566. DOI: [10.1177/0145445508317134](https://doi.org/10.1177/0145445508317134)

Dickersin, K., & Rennie, D. (2003). Registering clinical trials. *Journal of the American Medical Association*, 290, 516-523.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomized studies of health care interventions. *Journal of Epidemiology Community Health*, 52, 377-384.

Dretzke, J., Davenport, C., Few, E., Barlow, J., Stewart-Brown, S., Bayliss, S., ... Hyde, C. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: A systematic review of randomized controlled trials. *Child and Adolescent Psychiatry and Mental Health*, 3, 1-10. DOI: [10.1186/1753-2000-3-7](https://doi.org/10.1186/1753-2000-3-7)

Durlack, J. A. (1998). Primary prevention programs for children and adolescents are effective. *Journal of Mental Health*, 7, 463-469.

- Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A-W., Cronin, E., ... Williamson, P. R. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE*, 3, e3081. DOI:10.1371/journal.pone.0003081
- Eisner, M. (2009). No effects in independent prevention trials: Can we reject the cynics' hypothesis? *Journal of Experimental Criminology*, 5, 163-183.
- Emrick, C.D. (1975). A review of psychologically oriented treatment of alcoholism. *Journal of Studies on Alcohol*, 36, 88-108.
- Evidence-Based Medicine Working Group. (1992). Evidence-based medicine. A new approach to teaching the practice of medicine. *Journal of the American Medical Association*, 268, 420-425.
- Eyberg, S. M., & Calzada, E. J. (1998). *Parent-child interaction therapy: Procedures manual*. Unpublished manuscript, University of Florida.
- Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology*, 37, 215-237. DOI: 10.1080/15374410701820117
- Eyberg, J., & Pincus, D. (1999). *The ECBI and SESBI-R: Eyberg Child Behavior and Sutter-Eyberg Student Behavior Inventory – Revised: Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting and Clinical Psychology*, 16, 319-324.
- Eysenck, H. J. (1978). An exercise in mega-silliness. *American Psychologist*, 33, 517. DOI: 10.1037/0003-066X.33.5.517.a

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries.

Scientometrics, 90, 891-904. DOI: 10.1007/s11192-011-0494-7

Fletcher, R., Freeman, E., & Matthey, S. (2011). The impact of behavioural parent training on fathers' parenting: A meta-analysis of the triple p – positive parenting program.

Fathering, 9, 291-312. DOI: 10.3149/fth.0903.291

Forehand, R. L., & McMahon, R. (1981). *Helping the noncompliant child*. New York: Guilford Press.

Frater, J., Myohanen, L., Taylor, E., & Keith, L. (2007). What would you tell me if I said grey literature? The New York academy of medicine grey literature report. *Journal of Electronic Resources in Medical Libraries*, 4, 145-153.

Frick, P. J., & McMahon, R. J. (2008). Child and adolescent conduct problems. In J. Hunsley & E. J. Mash (Eds.), *A Guide To Assessments That Work* (pp. 41-68). New York: Oxford University Press.

Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S. M., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database for Systematic Reviews*, 2012(2), 1-344. DOI: 10.1002/14651858.CD008225.pub2.

*Gardner, F., Burton, J., & Klimes, I. (2006). Randomised controlled trial of a parenting intervention in the voluntary sector for reducing child conduct problems: Outcomes and mechanisms of change. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 47, 1123-1132. DOI: 10.1111/j.1469-7610.2006.01668.x

Gilgun, J. F. (2005). The four cornerstones of evidence-based practice in social work. *Research on Social Work Practice*, 15, 52-61. DOI: 10.1177/1049731504269581

- Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research Education, 5*, 351-379. DOI: 10.3102/0091732X005001351
- Gokhale, P. (1997). Grey Literature Varieties - Definitional Problems. In *Third International Conference on Grey Literature: Perspectives on the Design and Transfer of Scientific and Technical Information, 13-14 November 1997*.
- Grant, S., Montgomery, P., & Mayo-Wilson, E. (2012). Development of a CONSORT extension for interventions in public health and related disciplines. *The Lancet, 380*, S14.
- *Greene, R. W., Ablon, J. S., Monuteaux, M. C., Goring, J. C., Henin, A., Raezer-Blakely, L., . . . Rabbitt, S. (2004). Effectiveness of collaborative problem solving in affectively dysregulated children with oppositional-defiant disorder: Initial findings. *Journal of Consulting & Clinical Psychology, 72*, 1157-1164. DOI: 10.1037/0022-006X.72.6.1157
- Greene, R. W., Biederman, J., Zerwas, S., Monuteaux, M. C., Goring, J. C., & Faraone, S. V. (2002). Psychiatric comorbidity, family dysfunction, and social impairment in referred youth with oppositional defiant disorder. *American Journal of Psychiatry, 159*, 1214-1224. DOI: 10.1176/appi.ajp.159.7.1214
- Hamilton, J. (2005). Clinician's guide to evidence-based practice. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 494-498.
- Harbour, R., & Miller, J. (2001). A new system for grading recommendations in evidence based guidelines. *British Medical Journal, 323*, 334-336.
- Harris, C. R. (2003). A review of sex differences in sexual jealousy, including self-report data, psychophysiological responses, interpersonal violence, and morbid jealousy. *Personality and Social Psychology Review, 7*, 102-128.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando:

Academic Press.

Hedges, L.V., & Pigott, T.D. (2001). The power of statistical tests in meta-analysis.

Psychological Methods, 6, 203-217.

Hedges, L.V., & Pigott, T.D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9, 426-445. DOI: 10.1037/1082-989X.9.4.426

Hempel, S., Suttorp, M. J., Miles, J. N. V., Wang, Z., Maglione, M., Morton, S., ... Shekelle, P.

G. (2011, June). *Empirical evidence of associations between trial quality and effect size*

(AHRQ Publication No. 11-EHC045-EF). Retrieved from

http://effectivehealthcare.ahrq.gov/ehc/products/319/710/Assessing-Empirical-Evidence_Final-Report_20110621.pdf

Higgins J. P. T., Green S. (editors). *Cochrane Handbook for Systematic Reviews of Interventions*

Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Hopewell, S., & Clarke, M. (2001). Methodologists and their methods: Do methodologists write up their conference presentations or is it just 15 minutes of fame? *International Journal of Technology Assessment in Health Care*, 17, 601-603.

Hopewell, S., McDonald, S., Clarke, M. J., & Egger, M. (2008). Grey literature in meta-analyses of randomized trials of health care interventions (Review). *The Cochrane Library*, 1-18.

Hudziak, J. J., Achenbach, T. M., Althoff, R. R., & Pine, D. S. (2007). A dimensional approach to developmental psychopathology. *International Journal of Methods in Psychiatric Research*, 16, S16-S23.

Hunsley, J., & Mash, E. J. (2008). Developing criteria for evidence-based assessment: An introduction to assessments that work. In J. Hunsley & E. J. Mash (Eds.), *A Guide*

- To Assessments That Work* (pp. 3-14). New York: Oxford University Press.
- Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, 8, 275-292.
- *Hutchings, J., Bywater, T., Daley, D., Gardner, F., Daley, D., Whitaker, C., . . . Edwards, R. T. (2007). Parenting intervention in Sure Start services for children at risk of developing conduct disorder: Pragmatic randomised controlled trial. *British Medical Journal*, 334, 678-682. DOI: 10.1136/bmj.39126.620799.55
- Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials*, 17, 1-12. DOI: 10.1016/0197-2456(95)00134-4
- Jensen, P. S. (2003). What is the evidence for evidence-based treatments? *Emotional & Behavioral Disorders in Youth*, 3, 37-41.
- Johnston, C., & Mah, J. W. T. (2008). Child attention-deficit/hyperactivity disorder. In J. Hunsley & E. J. Mash (Eds.), *A Guide To Assessments That Work* (pp. 17-40). New York: Oxford University Press.
- Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology*, 36, 567-589. DOI: 10.1007/s10802-007-9201-9
- Kamphaus, R.W., & Campbell, J.C. (2006). Psychodiagnostic assessment of children: Dimensional and categorical methods. New York, NY: Wiley.
- Kazdin, A. E. (1987). Treatment of antisocial behavior in children: Current status and future

directions. *Psychological Bulletin*, 102, 187-203.

Kazdin, A. E. (2005). Child, parent, and family-based treatment of aggressive and antisocial child behavior. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial Treatments for Child and Adolescent Disorders: Empirically Based Strategies for Clinical Practice* (2nd ed.; pp. 445-476). Washington: American Psychological Association.

Kazdin, A. E. (2010). Problem-solving skills training and parent management training for oppositional defiant disorder and conduct disorder. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (2nd ed.). New York, NY: The Guilford Press.

Kazdin, A. E., & Bass, D. (1989). Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting and Clinical Psychology*, 57, 138-147.

*Kazdin, A. E., Siegel, T. C., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. *Journal of Consulting and Clinical Psychology*, 60, 733-747.

*Kling, A., Forster, M., Sundell, K., & Melin, L. (2010). A randomized controlled effectiveness trial of parent management training with varying degrees of therapist support. *Behavior Therapy*, 41, 530-542. DOI: 10.1016/j.beth.2010.02.004

*Landy, S., & Menna, R. (2006). An evaluation of a group intervention for parents with aggressive young children: Improvements in child functioning, maternal confidence, parenting knowledge and attitudes. *Early Child Development and Care*, 176, 605-620. DOI: 10.1080/03004430500147425

- *Larsson, B., Fossum, S., Clifford, G., Drugli, M. B., Handegard, B. H., & Morch, W. T. (2009). Treatment of oppositional defiant and conduct problems in young Norwegian children: Results of a randomized controlled trial. *European Child & Adolescent Psychiatry, 18*, 42-52. DOI: 10.1007/s00787-008-0702-z
- *Lehner-Dua, L. L. (2001). *The effectiveness of Russell A. Barkley's parent training program on parents with school-aged children who have ADHD on their perceived severity of ADHD, stress, and sense of competence* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3030505)
- *Leung, C., Sanders, M. R., Leung, S., Mak, R., & Lau, J. (2003). An outcome evaluation of the implementation of the Triple P Positive Parenting Program in Hong Kong. *Family Process, 42*, 531-544. DOI: 10.1111/j.1545-5300.2003.00531.x
- Levant, R. F. (2005). *Report of the 2005 presidential task force on evidence-based practice*. Retrieved from <http://www.apa.org/practice/resources/evidence/evidence-based-report.pdf>
- Levant, R. F. (2006). Making psychology a household word. *American Psychologist, 61*, 383-395.
- Leykin, Y., & DeRubeis R. J. (2009). Allegiance in psychotherapy outcome research: Separating association from bias. *Clinical Psychology: Science and Practice, 16*, 54-65.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., ... Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *British Medical Journal, 339*, b2700. DOI: 10.1136/bmj.b2700

- Light, R. J., & Smith, P. V. (1971). Accumulating evidence: Procedures for resolving contradictions among different research studies. *Harvard Educational Review*, 41, 429-471.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *The American Psychologist*, 48, 1181-1209.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.
- Loeber, R. (1991). Antisocial behavior: More enduring than changeable? *Journal of the American Academy of Child and Adolescent Psychiatry*, 30, 393-397.
- Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*, 26, 86-104. DOI: 10.1016/j.cpr.2005.07.004
- Lundahl, B. W., Tollefson, D., Risser, H., & Lovejoy, M. C. (2008). A meta-analysis of father involvement in parent training. *Research on Social Work Practice*, 18, 97-106. DOI: 10.1177/1049731507309828
- Mahood, Q., van Eerd, D. & Irvin, E-L. (2010). [P165] Grey literature in systematic reviews: a practical approach to searching and including in systematic reviews. *Cochrane Colloquium Abstracts Journal*.
- Manchikanti, L. (2008). Evidence-based medicine, systematic reviews, and guidelines in interventional pain management, part I: Introduction and general considerations. *Pain Physician*, 11, 161-186.

- Martin, J. L. R., Perez, V., Sacristan, M., & Alvarez, E. (2005). Is grey literature essential for a better control of publication bias in psychiatry? An example from three meta-analyses of schizophrenia. *European Psychiatry*, 20, 550-553. DOI: 10.1016/j.eurpsy.2005.03.011
- Mash, E. J., & Wolfe, D. A. (2005). *Abnormal child psychology* (3rd ed.). Belmont, CA: Thomson Wadsworth.
- Maughan, D. R., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review*, 34, 267-286.
- Mays, G. P., & Hogg, R. A. (2012). Expanding delivery system research in public health settings: Lessons from practice-based research networks. *Journal of Public Health Management Practice*, 18, 485-498. DOI: 10.1097/PHH.0b013e31825f75c9
- McAuley, L., Pham, B., Tugwell, P., & Moher, D., (2000). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *The Lancet*, 356, 1228-1231.
- McAuley, L. M. (1999). *The influence of 'grey' literature on meta-analysis* (Masters Thesis). Retrieved from uO Research.
- McCabe, K. M., Hough, R., Wood, P. A., & Yeh, M. (2001). Childhood and adolescent onset conduct disorder: A test of developmental taxonomy. *Journal of Abnormal Child Psychology*, 29, 305-316.
- *McCabe, K. M., & Yeh, M. (2009). Parent-child interaction therapy for mexican americans: A randomized clinical trial. *Journal of Clinical Child & Adolescent Psychology*, 38, 753-759. DOI: 10.1080/15374410903103544

- McCart, M. R., Priester, P. E., Davies, W. H., & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology*, 34, 527-543. DOI: 10.1007/s10802-006-9031-1
- *McGilloway, S., Bywater, T., Mhaille, G. N., Furlong, M., O'Neill, D., Comiskey, C., ... Donnelly, M. (n.d.). *Proving the power of positive parenting: A randomised control trial to investigate the effectiveness of the Incredible Years Basic Parent training programme in an Irish context (short-term outcomes)*. Retrieved from http://eprints.nuim.ie/1816/1/Research_Report_-_Incredible_Years_Ireland_Study.pdf
- McLeod, B. D., & Weisz, J. R. (2004). Using dissertations to examine potential bias in child and adolescent clinical trials. *Journal of Consulting and Clinical Psychology*, 72, 235-251. DOI: 10.1037/0022-006X.72.2.235
- Miller, S., Wampold, B., Varhely, K. (2008). Direct comparisons of treatment modalities for youth disorders: A meta-analysis. *Psychotherapy Research*, 18, 5-14. DOI: 10.1080/10503300701472131
- *Moharreri, F., Shahrivar, Z., Tehrani-doost, M., & Mahmoudi-Gharaei, J. (2008). Efficacy of the positive parenting program (triple p) for parents of children with attention deficit/hyperactivity disorder. *Iranian Journal of Psychiatry*, 3, 59-63.
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet*, 354, 1896-1900. DOI: 10.1016/S0140-6736(99)04149-5
- Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., ... Altman, D. G., for the CONSORT Group. (2010). CONSORT 2010 Explanation and Elaboration:

- updated guidelines for reporting parallel group randomised trial. *British Medical Journal*, 340, c869-897. DOI: 10.1016/j.jclinepi.2010.03.004
- Moyer, A., Schneider, S., Knapp-Oliver, S. K., & Sohl, S. J. (2010). Published versus unpublished dissertations in psycho-oncology intervention research. *Psycho-oncology*, 19, 313-317. DOI: 10.1002/pon.1561
- Munder, T., Gerger, H., Trelle, S., & Barth, J. (2011). Testing the allegiance bias hypothesis: A meta-analysis. *Psychotherapy Research*, 21, 670-684. DOI: 10.1080/10503307.2011.602752
- National Institute for Health and Clinical Excellence (2006, July). *Parent-training/education programmes in the management of children with conduct disorders* (NICE technology appraisal guidance 102). London, UK: Author.
- National Institute for Health and Clinical Excellence (2008, September). *Attention deficit hyperactivity disorder: Diagnosis and management of ADHD in children, young people and adults* (NICE clinical guideline 72). London, UK: Author.
- Nay, W. R. (1975). A systematic comparison of instructional techniques for parents. *Behavior Therapy*, 6, 14-21.
- Nestoriuc, Y., Kriston, L., & Rief, W. (2010). Meta-analysis as the core of evidence-based behavioral medicine: Tools and pitfalls of a statistical approach. *Current Opinion in Psychiatry*, 23, 145-150. DOI: 10.1097/YCO.0b013e328336666b
- Niemeyer, H., Musch, J., & Pietrowsky, R. (2012, December 17). Publication bias in meta-analyses of the efficacy of psychotherapeutic interventions for depression. *Journal of Consulting and Clinical Psychology*. Advance online publication. DOI: 10.1037/a0031152

- *Nixon, R. D. V., Sweeney, L., Erickson, D. B., & Touyz, S. W. (2003). Parent-child interaction therapy: A comparison of standard and abbreviated treatments for oppositional defiant preschoolers. *Journal of Consulting & Clinical Psychology, 71*, 251-260. DOI: 10.1037/0022-006X.71.2.251
- Norcross, J. C., Krebs, P. M., & Prochaska, J. O. (2011). Stages of change. *Journal of Clinical Psychology, 67*, 143-154. DOI: 10.1002/jclp.20758
- Nowak, C., & Heinrichs, N. (2008). A comprehensive meta-analysis of triple p – positive parenting program using hierarchical linear modeling: Effectiveness and moderating variables. *Clinical Child and Family Psychology Review, 11*, 114-144. DOI: 10.1007/s10567-008-0033-0
- *Ogden, T., & Amlund-Hagen, K. (2008). Treatment effectiveness of parent management training in Norway: A randomized controlled trial of children with conduct problems. *Journal of Consulting & Clinical Psychology, 76*, 607-621. DOI: 10.1037/0022-006X.76.4.607
- Okoroma, F. N. (2011). Towards the effective management of grey literature for higher education, research, and national development. *Library Review, 60*, 789-802. DOI: 10.1108/00242531111176808
- Paley, G., & Shapiro, D. A. (2002). Lessons from psychotherapy research for psychological interventions for people with schizophrenia. *Psychology and Psychotherapy, 75*, 5-17.
- Pappas, C., & Williams, I. (2011). Grey literature: Its emerging importance. *Journal of Hospital Librarianship, 11*, 228-234. DOI: 10.1080/15323269.2011.587100

- Patsopoulos, N. A., Analatos, A. A., & Ioannidis, J. P. (2005). Relative citation impact of various study designs in the health sciences. *The Journal of the American Medical Association*, 293, 2362-2366.
- Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.
- Pelham Jr., W. E., & Fabiano, G. A. (2008). Evidence-based psychosocial treatments for attention-deficit/hyperactivity disorder: An update. *Journal of Clinical Child and Adolescent Psychology*, 37, 184-214. DOI: 10.1080/15374410701818681
- Petrosino, A. & Soydan, H. (2005). The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1, 435-450.
- Piquero, A. R., Farrington, D. P., Welsh, B. C., Tremblay, R., & Jennings, W. G. (2009). Effects of early family/parent training programs on antisocial behavior and delinquency. *Journal of Experimental Criminology*, 5, 83-120. DOI 10.1007/s11292-009-9072-x
- *Pisterman, S., Firestone, P., McGrath, P., Goodman, J. T., Webster, I., Mallory, R., & Goffin, B. (1992). The role of parent training in treatment of preschoolers with ADDH. *American Journal of Orthopsychiatry*, 62, 397-408.
- *Pisterman, S., McGrath, P., Firestone, P., & Goodman, J. T. (1989). Outcome of parent-mediated treatment of preschoolers with Attention Deficit Disorder with hyperactivity. *Journal of Consulting & Clinical Psychology*, 57, 628-635.
- *Plant, K. M., & Sanders, M. R. (2007). Reducing problem behavior during care-giving in families of preschool-aged children with developmental disabilities. *Research in Developmental Disabilities*, 28, 362-385. DOI: 10.1016/j.ridd.2006.02.009
- Posner, B. P., & Bellefontaine, S. (2011). *ReSearch* [Computer Software].

- Provos, N., McClain, J., Wang, K. (2006, November). *Search worms*. Paper presented at the 4th Workshop on Recurring Malcode, Fairfax, VA. Retrieved from http://www.citi.umich.edu/u/provos/papers/search_worms.pdf
- Quintana, S. M., & Minami, T. (2006). Guidelines for meta-analysis of counseling psychology research. *Counseling Psychologist*, 34, 839–877. DOI:10.1177/0011000006286991
- Rescorla, L., Achenbach, T., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., ... Verhulst, F. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and Behavioral Disorders*, 15, 130-142. DOI: 10.1177/10634266070150030101
- Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems – a meta-analytic review. *Journal of Child Psychology and Psychiatry*, 47, 99-111. DOI: 10.1111/j.1469-7610.2005.01544.x
- Rosenberg, M. S., Adams, D. C., & Gurevitch, J. (2000). *MetaWin. statistical software for meta-analysis: Version 2.0*. Sunderland: Sinauer Associates.
- Rosenthal, R., and DiMatteo, M.R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, 52, 59-82.
- *Ross, C. M. (2002). *Bridging the gaps in services for families of children with attention-deficit/hyperactivity disorder: Examining the effectiveness of parent management training groups in a rural community* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. MQ85908)
- Roth, B. J. (2010). An academic perspective on grey literature. *Archaeologies: Journal of the World Archaeological Congress*, 6, 337-345. DOI: 10.1007/s11759-010-9141-9

- Royle, J., & Blythe, J. (1998). Promoting research utilisation in nursing: The role of the individual, organisation, and environment. *Evidence-Based Nursing*, 1, 71-72. DOI: 10.1136/ebn.1.3.71
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence-based medicine: What it is and what it isn't. *British Medical Journal*, 312, 71-72. DOI: 10.1136/bmj.312.7023.71
- Sackett, D. L., Straus, S., Richardson, S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-based medicine: How to practice and teach EBM* (2nd ed.). London, UK: Churchill Livingstone.
- Sacks, H. S., Reitman, D., Pagano, D., & Kupelnick, B. (1996). Meta-analysis: An update. *Mount Sinai Journal of Medicine*, 63, 216-224.
- Sanders, M. R. (1999). Triple P- positive parenting program: Towards an empirically validated multilevel parenting and family support strategy for the prevention of behavior and emotional problems in children. *Clinical Child and Family Psychology Review*, 2, 71-90.
- *Sanders, M. R., Markie-Dadds, C., Tully, L. A., & Bor, W. (2000). The Triple P-Positive Parenting Program: A comparison of enhanced, standard, and self-directed behavioral family intervention for parents of children with early onset conduct problems. *Journal of Consulting and Clinical Psychology*, 68, 624-640.
- Schmidt, L. M., & Gotzsche, P. C. (2005). Of mites and men: reference bias in narrative review articles – a systematic review. *Journal of Family Practice*, 54, 334-338.

- *Schuhmann, E. M., Foote, R. C., Eyberg, S. M., Boggs, S. R., & Algina, J. (1998). Efficacy of parent-child interaction therapy: Interim report of a randomized trial with short-term maintenance. *Journal of Clinical Child Psychology*, 27, 34-45.
- Schulz, K. F., Altman, D. G., Moher, D., for the CONSORT Group. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *British Medical Journal*, 340, 698-702. DOI: 10.1136/bmj.c332
- *Scott, S., Spendor, Q., Doolan, M., Jacobs, B., & Aspland, H. (2001). Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice. *British Medical Journal*, 323, 1-7.
- Scullion, P.A. (2002). Effective dissemination strategies. *Nurse Researcher*, 10, 65-67.
- Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy*, 27, 171-186. DOI: 10.1016/S0005-7894(96)80013-X
- Seymour, D. J. (2010). Sanctioned inequity and accessibility issues in the grey literature in the United States. *Archaeologies: Journal of the World Archaeological Congress*, 6, 233-269. DOI: 10.1007/s11759-010-9144-6
- Shadish, W. R., & Baldwin, S. (2005). Effects of behavioral marital therapy: A meta-analysis of randomized controlled trials. *Journal of Consulting and Clinical Psychology*, 73, 6-14.
- *Sheeber, L. B., & Johnson, J. H. (1994). Evaluation of a temperament-focused, parent-training program. *Journal of Clinical Child Psychology*, 23, 249-259. DOI: 10.1207/s15374424jccp2303_3

- Shercliffe, R. J., Stahl, W., & Tuttle, M. P. (2009). Meta-analysis as used by psychology: A casting of old wine in new bottles or a superior vintage? *Theory and Psychology*, 19, 413-430.
- Shojania, K. G., Sampson, M., Ansari, M. T., Ji, J., Doucette, S., & Moher, D. (2007). How quickly do systematic reviews go out of date? A survival analysis. *Annals of Internal Medicine*, 147, 224-233.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 36, 753-760.
- Smith, M., Glass, G., & Miller, T. (1980). *The Benefits of Psychotherapy*. Baltimore, MD: John Hopkins University Press.
- *Sofronoff, K., Jahnel, D., & Sanders, M. (2011). Stepping Stones Triple P seminars for parents of a child with a disability: A randomized controlled trial. *Research in Developmental Disabilities*, 32, 2253-2262. DOI: 10.1016/j.ridd.2011.07.046
- *Sofronoff, K., Leslie, A., & Brown, W. (2004). Parent management training and Asperger syndrome. A randomized controlled trial to evaluate a parent based intervention. *Autism*, 8, 301-317. DOI: 10.1177/1362361304045215
- Solomon, D. J. (2007). The role of peer review for scholarly journals in the information age. *The Journal of Electronic Publishing*, 10. DOI: <http://dx.doi.org/10.3998/3336451.0010.107>
- *Sommer, J. L. (2007). *Efficacy of parent training for children with attention-deficit/hyperactivity disorder: A therapeutic component analysis* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 1443241)
- *Sonuga-Barke, E. J. S., Daley, D., Thompson, M., Laver-Bradbury, C., & Weeks, A. (2001). Parent-based therapies for preschool attention-deficit/hyperactivity disorder: A

- randomized controlled trial with a community sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40, 402-408. DOI: 10.1097/00004583-200104000-00008
- Spring, B. (2007). Evidence-based practice in clinical psychology: What it is, why it matters: what you need to know. *Journal of Clinical Psychology*, 63, 611-631. DOI: 10.1002/jclp.20373
- Staines, G. L., & Cleland, C. M. (2007). Bias in meta-analytic estimates of the absolute efficacy of psychotherapy. *Review of General Psychology*, 11, 329-347. DOI: 10.1037/1089-2680.11.4.329
- Stanley, T. D. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, 15, 131-150.
- Taylor, T. K., & Biglan, A. (1998). Behavioral family interventions for improving child-rearing: A review of the literature for clinicians and policy makers. *Clinical Child and Family Psychology Review*, 1, 41-60. DOI: 10.1023/A:1021848315541
- *Taylor, T. K., Schmidt, F., Pepler, D., & Hodgins, C. (1998). A comparison of eclectic treatment with Webster-Stratton's parents and children series in a children's mental health center: A randomized controlled trial. *Behavior Therapy*, 29, 221-240. DOI: 10.1016/S0005-7894(98)80004-X
- Thomas, R., & Zimmer-Gembeck, M. J. (2007). Behavioral outcomes of parent-child interaction therapy and triple p-positive parenting program: A review and meta-analysis. *Journal of Abnormal Child Psychology*, 35, 475-495. DOI 10.1007/s10802-007-9104-9

- *Thomas, R., & Zimmer-Gembeck, M. J. (2011). Accumulating evidence for parent-child interaction therapy in the prevention of child maltreatment. *Child Development*, 82, 177-192. DOI: 10.1111/j.1467-8624.2010.01548.x
- *Thompson, R. W., Ruma, P. R., Schuchmann, L. F., & Burke, R. V. (1996). A cost-effectiveness evaluation of parent training. *Journal of Child & Family Studies*, 5, 415-429. DOI: 10.1007/BF02233863
- *Thorell, L. B. (2009). The community parent education program (COPE): Treatment effects in a clinical and community-based sample. *Clinical Child Psychology and Psychiatry*, 14, 373-387. DOI: 10.1177/1359104509104047
- Tricco, A. C., Tetzlaff, J., Sampson, M., Fergusson, D., Cogo, E., Horsley, T., & Moher, D. (2008). Few systematic reviews exist documenting the extent of bias: A systematic review. *Journal of Clinical Epidemiology*, 61, 422-434. DOI: 10.1016/j.jclinepi.2007.10.017
- *Turner, K. M., Richards, M., & Sanders, M. R. (2007). Randomised clinical trial of a group parent education programme for Australian Indigenous families. *Journal of Paediatrics and Child Health*, 43, 429-437. DOI: 10.1111/j.1440-1754.2007.01053.x
- United Nations Department of Crime (2012). *Compilation of evidence-based family skills training programmes*. Retrieved from http://www.unodc.org/docs/youthnet/Compilation/10-50018_Ebook.pdf
- *van de Wiel, N. M. H., Matthys, W., Cohen-Kettenis, P. T., Maassen, G. H., Lochman, J. E., & van Engeland, H. (2007). The effectiveness of an experimental treatment when compared to care as usual depends on the type of care as usual. *Behavior Modification*, 31, 298-312. DOI: 10.1177/0145445506292855

- *Van Den Hoofdakker, B. J., Van Der Veen-Mulders, L., Sytema, S., Emmelkamp, P. M. G., Minderaa, R. B., & Nauta, M. H. (2007). Effectiveness of behavioral parent training for children with ADHD in routine clinical practice: A randomized controlled study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46, 1263-1271. DOI: 10.1097/chi.0b013e3181354bc2
- *Van Zeijl, J., Mesman, J., Van Ijzendoorn, M. H., Bakermans-Kranenburg, M. J., Juffer, F., Stolk, M. N., ... Alink, L. R. A. (2006). Attachment-based intervention for enhancing sensitive discipline in mothers of 1- to 3-year-old children at risk for externalizing behavior problems: A randomized control trial. *Journal of Consulting and Clinical Psychology*, 74, 994-1005.
- Von Elm, E., Costanza, M. C., Walder, B., & Tramer, M. R. (2003). More insight into the fate of biomedical meeting abstracts: A systematic review. *BMC Medical Research Methodology*, 3, 12-23. DOI: 10.1186/1471-2288-3-12
- Wahler, R. C., Winkel, G. H., Peterson R. F., & Morrison, D. C. (1965). Mother as behavior therapists for their own children. *Behaviour Research and Therapy*, 3, 113-134.
- Wampold, B. E. (2001). *The great psychotherapy debate: Model, methods, and findings*. Mahwah, NJ: Erlbaum.
- *Ward, S. L. (1997). *Glasser's parent training model: Effects on child and parent functioning* (Unpublished doctoral dissertation). University of Arizona, Arizona.
- *Webster-Stratton, C. (1992). Individually administered videotape parent training: "Who benefits?" *Cognitive Therapy and Research*, 16, 31-35.
- Webster-Stratton, C., & Herbert, M. (1994). *Troubled families-problem children: Working with parents: A collaborative process*. Chichester, UK: Wiley.

*Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct-problem children: Comparison with two cost-effective treatments and a control group. *Journal of Consulting & Clinical Psychology*, 56, 558-566.

*Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology*, 33, 105-124.

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin*, 132, 132-149. DOI: 10.1037/0033-2909.132.1.132

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology*, 55, 542-549. DOI: 0022-006X/87/J00.75

*Whittingham, K., Sofronoff, K., Sheffield, J. K., & Sanders, M. R. (2009). Stepping stones triple P: An RCT of a parenting program with parents of a child diagnosed with an autism spectrum disorder. *Journal of Abnormal Child Psychology*, 37, 469-480. DOI 10.1007/s10802-008-9285-x

*Wiggins, T. L., Sofronoff, K., & Sanders, M. R. (2009). Pathways triple P-positive parenting program: Effects on parent-child relationships and child behavior problems. *Family Process*, 48, 517-530.

Wilson, D. B., & Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Bulletin*, 6, 413-429.

Wood, J. (2008). Methodology for dealing with duplicate study effects in meta-analysis.

Organizational Research Methods, 11, 79-95. DOI: 10.1177/1094428106296638

World Health Organization (2001). *The world health report: 2001: Mental health: New understanding, new hope*. Geneva: Author.

Table 1

Meta-Analyses of Behavioral Parenting Interventions for Children's Disruptive Behavior

Authors	Diagnoses (or status)	# Studies	Inclusion Criteria	Child Ages	Conclusions (ES)	Moderators
de Graaf, Speetjens, Smit, de Wolff, & Tavecchio (2008)	Clinical: Y (26.32%); Dx: Y (5.26%)	19	1975-2006; Triple P level 4 intervention; PS and PSOC	n/a	P Style = .68 parent competence = .65	C & F variables; Tx Characteristics
Dretzke et al. (2009)	Clinical: Y (71.64%); Dx: Y (14.93%)	24	until 2006; RCT; ECBI or CBCL	0 to 18	SMD CB = -.67 (favors Interv. Grp)	n/a
Furlong et al. (2012)	Clinical: Y (100%) Dx: Y (46%)	15	Until January 2011; RCTs and Quasi-randomized studies	3 to 12	SMD CB = -.53 SMD PB = -.77 SMD PA = -.36	C & F variables; Tx Characteristics; Method. Characteristics
Fletcher, Freeman, & Matthey (2011)	Clinical: n/a Dx: n/a	28	1950-2011; RCTs; fathers targeted or eligible; Triple P	n/a	P Style = .77	Method. Characteristics
Kaminski, Valle, Filene, & Boyle (2008)	Clinical: Y (n/a); Dx: Y (n/a)	77	1990-2002	0 to 7	CB = .30 PB = .39	Tx Characteristics; Method. Characteristics

Authors	Diagnoses (or status)	# Studies	Inclusion Criteria	Child Ages	Conclusions (ES)	Moderators
Lundahl, Risser, & Lovejoy (2006)	Clinical: Y (67.16%); Dx: N	63	1974- August 2003	M = 6.79 years	CB = .42 PB = .45	C & F variables; Tx Characteristics
Lundahl, Tollefson, Risser, & Lovejoy (2008)	Clinical: n/a Dx: n/a	32	Until 2005	n/a	CB = .48 PB = .54	n/a
Maughan, Christiansen, Jenson, Olympia, & Clark (2005)	Clinical: n/a Dx: n/a	79	1966 – 2001	3 to 16	CB = .30	C & F variables; Tx Characteristics; Method. Characteristics
McCart, Priester, Davies, & Azen (2006)	Clinical: N; Dx: Y (20%)	30	Until 2004	under 18 years	CB = .38 PA+PS = .33 PB = .30	C & F variables; Tx Characteristics
Nowak & Heinrichs (2008)	Clinical: Y (16.36%); Dx: Y (5.45%)	55	1970-2007; Triple P	0 to 14	PB = .38; CB = .35; PA = .17	C & F variables; Tx Characteristics; Method. Characteristics
Piquero, Farrington, Welsh, Tremblay, & Jennings (2009)	Clinical: n/a Dx: n/a	55	Until 2008	1 to 5 (or M = ~5)	CB = .35	Tx Characteristics; Method. Characteristics

Authors	Diagnoses (or status)	# Studies	Inclusion Criteria	Child Ages	Conclusions (ES)	Moderators
Reyno & McGrath (2006)	Clinical: n/a Dx: n/a	31	1980 to 2004	1 to 16	n/a	C & F variables
Serketich & Dumas (1996)	Clinical: n/a Dx: n/a	26	Until 1995	M = 6.05	CB = 0.86	C & F variables; Tx Characteristics; Method. Characteristics
Thomas & Zimmer-Gembeck (2007)	Clinical: Y (79.17%) Dx: n/a	24	1980-2004; RCTs and single group follow-up; PCIT or Triple P	3 to 12	triple P = -.73 PCIT = -1.31	n/a

Note. ASB=Antisocial Behavior; BP=Behavior Problems; C=Child; CB=Child Behavior; CBCL=Child Behavior Checklist;

CL=Clinical; DBD=Disruptive Behavior Disorders; DX=Diagnosed; ECBI=Eyberg Child Behavior Inventory EP=Externalizing

Problems; ES=Effect Size; F=Family; M=Mean; Method. = Methodological; N = No; n/a=Not reported; NC= Non-clinical;

ODD=Oppositional Defiant Disorder; P=Parent; PA=Parent Adjustment; PB=Parent Behavior; PCIT=Parent-Child Interaction

Therapy; PS=Parenting Scale; PSOC=Parenting Sense of Competence Scale; RCT=Randomized Control Trial; SMD=Standardized

Mean Difference; Tx=Treatment; Y = Yes.

Table 2

Hypotheses for Moderator Variables

Moderator Variable	Parent Behaviour	Child Behaviour	Parent Adjustment
Child formal diagnostic status and comorbidity	No effect	No effect	Higher ES for diagnosed problems
Child age	Categorically: no difference	Categorically: no difference	Categorically: no difference
	Continuously: Higher ES for younger children	Continuously: Higher ES for younger children	Continuously: Higher ES for younger children
Comparison group	Higher ES for comparison to No-treatment vs. Active Comparison	Higher ES for comparison to No-treatment vs. Active Comparison	Higher ES for comparison to No-treatment vs. Active Comparison
Overall quality	Higher quality = lower ES	Higher quality = lower ES	E
Experimenter allegiance	E	E	E
Publication status	E	E	E

Note. E=Exploratory; ES=Effect Size.

Appendix A

List of Key Journals

List of Key Journals

Adolescence

Adolescent Psychiatry

American Journal of Psychiatry

American Psychologist

Annals of Clinical Psychiatry

Annual Review of Clinical Psychology

Annual Review of Psychology

Applied Psychology

Archives of General Psychiatry

Behavior Therapy

Behavior Research and Therapy

British Journal of Clinical Psychology

British Journal of Developmental Psychology

British Journal of Psychiatry

Canadian Psychology

Child and Adolescent Mental Health

Child Development

Child Psychology and Psychiatry Review

Clinical Child and Family Psychology Review

Clinical Psychology & Psychotherapy

Clinical Psychology Review

Clinical Psychology: Science and Practice

Cognitive and Behavioral Practice

Development and Psychopathology

European Journal of Development Research

European Journal of Developmental Psychology

Evidence Based Mental Health

The Family Journal

Family Relations

Journal of Abnormal Child Psychology

Journal of Abnormal Psychology

Journal of Adolescence

Journal of Adolescent Research

Journal of Attention Disorders

Journal of Child Psychology and Psychiatry and Allied Disciplines

Journal of Clinical Child and Adolescent Psychology

Journal of Clinical Psychology

Journal of Consulting and Clinical Psychology

Journal of Developmental and Physical Disabilities

The Journal of Early Adolescence

Journal of Family Psychology

Journal of Marital and Family Therapy

Journal of Marriage and the Family

Journal of Pediatric Psychology

Journal of Psychopathology and Behavioral Assessment

Journal of Research on Adolescence

Journal of the American Academy of Child & Adolescent Psychiatry

Journal of the Canadian Academy of Child & Adolescent Psychiatry

Journal of Youth and Adolescence

Merrill-Palmer Quarterly

New Directions for Child and Adolescent Development

Professional Psychology: Research and Practice

Psychological Bulletin

Psychotherapy Research

Scientific Review of Mental Health Practice

Appendix B

List of Search Terms

List of Search Terms

Behavioral/behavioural family interaction
Behavior/behaviour modification
Behavioral/behavioural parent training
Behavior/behaviour therapy
Contingency management
Parent-child interaction
Parent Effectiveness Training
Parent Management Training
Parent training
Community Parent Education
COPE
Defiant Children
Incredible Years
Parent-Child Interaction Therapy
PCIT
Positive Parenting Program
Triple P

ADD
ADHD
Aggressive
Antisocial
Attention deficit disorder
Attention deficit hyperactivity disorder
Behavior/behaviour disorder
CD
Conduct disorder
Conduct problem
DBD
Defiant
Disruptive
Disruptive behavior/behaviour disorder
Externalizing problem
Hyperkinetic
ODD
Oppositional defiant disorder

Appendix C

Study Inclusion Coding Manual

Criteria for Study Inclusion

Coder must address the following criteria, *in order*. If a study does not meet the requirements of an established criterion, it is immediately excluded. *It is not necessary to code any further criteria once the study has been rejected.*

1. Behavioural Parenting Intervention

- Intervention in which parents are taught behavioural parenting skills (e.g., differential reinforcement, time-out, praise, response cost, point systems) through groups or individually, using methods such as: didactic teaching, re-enactment, on-the-spot training, videotaped teaching, telephone or internet consultation.
- Intervention must have been carried out in the study being coded (i.e., cannot be a review of such an intervention)

2. Design.

- Eligible studies will have to have a comparison group, either a waitlist control group, no treatment group, or a group in which another therapeutic modality was employed.
- A sample size of at least 15 participants in each group (at time of analysis) will be required. If at least 2 eligible groups in a given study contain 15 or more participants, the study is included.

3. Child age.

- Parents of children targeted for treatment between the ages of 2 and 18 years.
- Age range of children can be assumed to be within the limits if clear grade levels are given (e.g., Kindergarten to Grade 9 are definitely within the range)

4. Clinically significant problems.

- Categorical: the target child for whom services are offered must have met criteria established by DSM-III, DSM-IV, or DSM-IV-TR for a Disruptive Behaviour Disorder (Attention Deficit/Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD), or Conduct Disorder (CD)).
- Dimensional: externalizing problem behaviour will be determined by a mean group score which is above the clinical cut-off on at least one externalizing problem behaviour on a psychometrically sound behavioural rating scale or clinical interview
 - Studies will be included if only the group receiving a behavioural parenting intervention has an elevated score.

5. Outcome measures.

- The study must include outcomes such as *child behaviour, parent behaviour, or parental adjustment*. Studies that will be included in the meta-analysis will examine at least one of these outcome variables through the course of their study.

Appendix D

Study Inclusion Coding Sheet

Appendix E

Multiple Articles Coding Manual

Multiple Articles Coding Manual

Often a single study yields multiple publications. This may occur when a series of papers are written presenting first overall findings, then follow-up data. Duplicate reports also occur when results on a subset of participants is reported (e.g., a study of children with ODD is reported, then another article describes the children within the sample who also had intellectual deficits). In a meta-analysis, it is essential that each sample is counted only once—so it is necessary to group articles that may be based on the same sample, and to identify which report is the most complete.

If two studies are suspected of being based on the same data set, this coding sheet will help to determine if they potentially are duplicates. If all criteria are met, then the studies examined are duplicates. Any unduplicated feature (going down the list in order) means that the studies are based on unique data sets.

1. Please note if the studies in question have any authors in common, judging by first initials and last names (not middle initials, as they are not always included).
2. Note if the following study characteristics are sufficiently different.
Duplicate possible if:
 - The population descriptions are the same or similar: age range of children, source of referral
 - The method used to analyze data is the same
 - The studies cover the same time periods (e.g., same years, number of years)
3. Note if the following sample characteristics are sufficiently different.
Duplicate possible if:
 - The response rates (from individuals and/or groups) are the same or similar
 - The demographic data (e.g., average income, average age of children or parents) is the same or similar
4. Note if the construct definitions and measures are sufficiently different.
Duplicate possible if:
 - The same treatment conditions are used
 - The same measures are used to examine outcomes
 - The constructs are *measured* in the same way
 - The constructs are *defined* in the same way
5. Note if the study effects and findings are sufficiently different.
Duplicate possible if:
 - The study effects are the same or very similar

Appendix F

Multiple Articles Coding Sheet

Multiple Articles on the Same Data-set Detection Coding Sheet

Based on Wood 2008

Suspected multiples:

Study 1: _____

Study 2: _____

1. Does the study share 1 or more authors with another study?

1 or more common coauthors PROBABLY NOT DUPLICATED

2. Are the study characteristics sufficiently different?

Same/similar populations Same method of analysis Same time periods

Same sample description NOT DUPLICATED

3. Are sample characteristics sufficiently different?

Same/similar response rates

Same/similar demographics data NOT DUPLICATED

4. Are constructs' definitions and measures sufficiently different to exclude study?

Same treatments Same scales used Same measures of constructs

Same construct definitions NOT DUPLICATED

5. Are matched study effects sufficiently different to exclude study?

Very similar/same study effects NOT DUPLICATED

DUPLICATES

NOT DUPLICATES

Appendix G

Methodological Quality Coding Manual

Methodological Quality Coding Criteria

The following codes are used to help determine a rating of the quality of the study.

1. *Group Assignment*: To determine the method used to assign participants to groups

Random Assignment: participants are assigned to groups randomly by the experimenters

Equivalency: participants are matched on certain characteristics and then assigned to groups

Unable to obtain/None: either participants decide on group placement themselves, are assigned based on position coming into the study, or by need for training.

2. *Psychometric Soundness of Measures (refer to associated coding sheet)*: Use the associated coding sheet to determine the soundness of measures being used. This info can be obtained from either the study measure description, or from independent information on the measure (e.g., other studies, measure information provided by developer). Provide soundness information for up to 5 measures. If there are more than 5 appropriate measures used in the study, pick the 5 best known or most psychometrically sound. Then, based on this information, provide an average rating of the psychometric soundness.

3. *Total Methodological Quality*: Provide the overall rating of the methodological quality by adding up the totals of the previous sections

Appendix H

Methodological Quality Coding Sheet

Coder Initials: _____

Authors: _____

Year: _____

Methodological Quality Coding Sheet

1. Group Assignment

Random Assignment (2) Equivalency (1) Contact author Unable to obtain (0)

Score _____ /2

Page number for this information: _____

2. Psychometric Soundness of Measures

Measure Name _____ /5

Measure Name _____ /5

Measure Name _____ /5

Measure Name _____ /5

Average Psychometric Soundness of Measures: _____ /5

Page number for this information: _____

(refer to checklist for psychometric soundness)

Total Methodological Quality: _____ /17

Appendix I

Moderator Variable Coding Manual

Moderator Variables Coding Criteria

This coding procedure provides information necessary for conducting moderator analyses. Each variable should be coded *if the information is available in the article*. If information is *not* available, please code N/A for information not available.

Child Variables

Mean Child Age: Record the mean child age in years. Using the mean child age, classify children according to the age ranges provided. Code “P” for Preschool (ages 2-5); “E” for Elementary School (ages 6-12); “H” for High School (ages 13-18). If the mean age falls between two age ranges:

.01-.49 → lower age range is marked

.50-.99 → higher age range is marked

Child Formal Diagnostic Status: mark whether children were formally diagnosed with (or met diagnostic criteria for) ADHD, ODD, Conduct Disorder (CD), or comorbid (i.e., more than one) conditions. If not, mark as having clinically significant externalizing problems (E).

Methodological Variables

Comparison Group: Note if the group that a behavioural parenting intervention is being compared to is a *waitlist control group* (participants are waiting for treatment), a *no treatment group* (no treatment is received or expected by participants), or *active comparison group* (e.g., psychotherapy, medication, or attention to participants).

Experimenter Allegiance: Note (Y/N) Whether one or more of the authors has a vested interest in the program being tested (e.g., author was involved in the development of the behavioural parenting intervention). Some examples of potential authors’ last names and their program allegiances are as follows:

C. Webster-Stratton: Incredible Years

M. Sanders: Triple P

R. Barkley: Defiant Children

Publication Status: Note if the study was published in a peer-reviewed journal

Appendix J

Moderator Variable Coding Sheet

Moderator Variable Coding Sheet

Study	Mean Child Age (Years)	Mean Child Age Range	Child Formal Diagnostic Status	Comparison Group	Experimenter Allegiance	Publication Status
-------	------------------------	----------------------	--------------------------------	------------------	-------------------------	--------------------