

Transporting Evidence-Based Parenting Programs for Child Problem Behavior (Age 3–10) Between Countries: Systematic Review and Meta-Analysis

Frances Gardner, Paul Montgomery, and Wendy Knerr

Centre for Evidence-Based Intervention, Department of Social Policy and Intervention, University of Oxford

There has been rapid global dissemination of parenting interventions, yet little is known about their effectiveness when transported to countries different from where they originated, or about factors influencing success. This is the first systematic attempt to address this issue, focusing on interventions for reducing child behavior problems. Stage 1 identified evidence-based parenting interventions showing robust effects in systematic reviews; Stage 2 identified trials of these interventions in a new country. Systematic review/meta-analysis of transported programs was followed by subgroup analyses by trial- and country-level cultural, resource, and policy factors. We found 17 transported trials of 4 interventions, originating in United States or Australia, tested in 10 countries in 5 regions, ($n = 1,558$ children). Effects on child behavior were substantial ($SMD = .71$) in the (14) randomized trials, but nonsignificant in the (3) nonrandomized trials. Subgroup analyses of randomized trials found no association between effect size and participant or intervention factors (e.g., program brand, staffing). Interventions transported to “western” countries showed comparable effects to trials in origin countries; however, effects were stronger when interventions were transported to culturally more distant regions. Effects were higher in countries with survival-focused family/childrearing values than those ranked more individualistic. There were no differences in effects by country-level policy or resource factors. Contrary to common belief, parenting interventions appear to be at least as effective when transported to countries that are more different culturally, and in service provision, than those in which they were developed. Extensive adaptation did not appear necessary for successful transportation.

There is substantial evidence that parenting interventions can improve parent–child relationships, reduce child problem behavior, and prevent maltreatment (Barlow, Johnston, Kendrick, Polnay, & Stewart-Brown, 2006; Piquero et al., 2008). This evidence, coupled with rising concern globally about youth problem behavior (Belfer, 2008), has led many governments and international bodies (e.g., World Health Organization

[WHO], United Nations Office on Drugs and Crime [UNODC]), to promote widespread transportation and rollout of evidence-based parenting programs (UNODC, 2009; Wessels et al., 2013; WHO, 2010), for example, in the United Kingdom (Scott, 2010), Norway (Larsson et al., 2009; Ogden & Hagen, 2008), and New Zealand (New Zealand Ministry of Education, 2009).

Although some trials have reported successful “transportation” of parenting interventions across countries and cultures (Gardner, Burton, & Klimes, 2006; Hutchings et al., 2007; Larsson et al., 2009; Leung, Sanders, Leung, Mak, & Lau, 2003; Reid, Webster-Stratton, & Beauchaine, 2001; Scott et al., 2010), others have found more disappointing results (Gottfredson et al., 2006; Kumpfer, Alvarado, Smith, & Bellamy,

Correspondence should be addressed to Frances Gardner, Centre for Evidence-Based Intervention, Department of Social Policy and Intervention, University of Oxford, Barnett House—32 Wellington Square, Oxford, OX1 2ER, United Kingdom. E-mail: frances.gardner@spi.ox.ac.uk

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hcap.

2002; Malti, Ribeaud, & Eisner, 2011; Sundell et al., 2008). Although the theoretical basis and process for translating effective interventions according to cultural or national differences is not fully understood, there is a growing literature on this topic. For example, models have been developed for examining applicability of public health evidence across contexts (e.g., Bonell, Oakley, Hargreaves, Strange, & Rees, 2006; Burchett, Umoquit, & Dobrow, 2011; Wang, Moss, & Hiller, 2006), and frameworks, derived from mental health and prevention literature, to guide translation and adaptation (Barrera & Castro, 2006; Ferrer-Wreder, Sundell, & Mansoor, 2012; Kumpfer, Pinyuchon, de Melo, & Whiteside, 2008; Sussman, Unger, & Palinkas, 2008). Moreover, there is a growing body of literature—though no systematic review—on factors affecting whether transportation is likely to be appropriate or successful. Paramount among these are cultural and service contexts and how these interact with characteristics of the intervention programme.

Cultural factors that may influence parenting interventions have been examined *within* countries, especially in the United States (Kumpfer et al., 2002; Reid et al., 2001), but cultural and contextual divergence in intervention effects *between* countries has received little systematic attention. Cultural norms and values related to parenting and family practices, and political and religious factors, are all likely to influence acceptability and effectiveness of evidence-based interventions (Kumpfer et al., 2008; Lau, 2006; Palinkas et al., 2009; Webster-Stratton, 2009). Characteristics and implementation of the intervention may also be crucial, including the degree of cultural competence, cultural flexibility (Webster-Stratton, 2009) and adaptation involved, and the interplay between adaptation and fidelity (Castro, Barrera, & Holleran Steiker, 2010).

Cultural differences are especially relevant in today's increasingly heterogeneous societies, where immigration and other demographic shifts are underlining the importance of constructs such as acculturation, discrimination, and migration on processes of change (Lau, 2006). Moreover, there is burgeoning interest in transporting parenting interventions to low- and middle-income countries, where multiple factors (e.g., poverty, service capacity and resources, HIV/AIDS, inequality, and variations in family structure) influence the feasibility, acceptability, and effectiveness of interventions (e.g., Baker-Henningham & Walker, 2009; Knerr, Gardner, & Cluver, 2013; Mejia, Calam, & Sanders, 2014).

New service contexts are also likely to have a strong influence on effectiveness and appropriateness of transported interventions. Service factors include preexisting service organization, quality, and leadership, as well as prevailing values or expectations related to family services, all of which may be markedly different compared

to the settings in which many well-known parenting interventions were developed and tested (primarily the United States and Australia; Backer, 2001; Baker-Henningham & Walker, 2009; Gustle, Hansson, Sundell, & Andree-Löfholm, 2008; Hutchings, Bywater, & Daley, 2007; Sussman et al., 2008). For example, transportation of Multisystemic Therapy for young offenders from the United States to Sweden and Canada yielded disappointing results, and researchers have suggested that marked differences in youth welfare systems and sociodemographic factors may have been significant moderators of effect (Littell, 2006; Sundell et al., 2008).

Amidst this growing interest in transportability, there is a need for careful consideration of the weight given not only to cultural and contextual variation but also to the potential for commonalities across settings that may facilitate transportation. For example, some studies suggest that parenting, and indeed intervention effects, are more similar than different across cultures and countries (e.g., Albert, Trommsdorff, & Mishra, 2007; Bradford et al., 2003; Pinderhughes, Hurley, & The Conduct Problems Prevention Research Group, 2008; Reid et al., 2001). Moreover, the evidence from trials and systematic reviews is equivocal concerning the effectiveness of culturally adapted versions of interventions compared to those that have not been adapted (Barrera, Castro, & Steiker, 2011; Gottfredson et al., 2006; Huey & Polo, 2008; Wilson & Miller, 2003). It is also important to question whether “transporting” interventions is appropriate, particularly when service, cultural, and national contexts appear to diverge dramatically, for example, between developed and developing countries, “Eastern” and “Western” societies, or those with widely varying policy regimes, particularly in relation to family policies and child welfare systems.

Given the current rapid dissemination of parenting interventions across developed and developing countries, the public resources being devoted to this endeavor, and the critical questions voiced about making generalizations across contexts from randomized trials (e.g., Cartwright & Hardie, 2012), there is an urgent need to critically examine the effectiveness of interventions that have been transported and, where appropriate, to develop models for successful transportation. This will contribute to efforts aimed at reducing youth problem behavior, maltreatment, and poor parenting but also promote effective and culturally applicable public spending on youth and family services, in health; social care; education and justice; and, increasingly, international development.

To our knowledge, this is the first attempt to systematically review transportability of evidence-based parenting interventions between countries. It should be noted that transportability across cultures within countries is also a critical topic for research (Castro

et al., 2010) but falls outside the scope of this review. Moreover, there are challenges in examining these issues in countries where services (and hence trials) are open to families from many cultural groups, thus requiring analysis of data within rather than between trials. By contrast, focusing on transportability across countries means that our review not only assesses the effectiveness of transported interventions in general but also allows an exploratory analysis of how outcomes might vary by country-level cultural and family policy differences, which the literature suggests may be of key importance in the effectiveness of parenting interventions (Sundell et al., 2008). We raise some potential caveats in advance. For this first attempt at a cross-country systematic review, we focused on a subset of parenting interventions that has relatively robust evidence from high-quality systematic reviews, showing at least moderate effect sizes on child outcomes in their origin countries, namely, parenting interventions for reducing behavioral problems in children ages 3 to 10 (Furlong et al., 2012; Piquero et al., 2008), in treatment or indicated prevention samples. We chose not to include primary or selective preventive interventions, as we judged the evidence of their effects on child outcomes to be less consistent (e.g., Malti et al., 2011; Scott, Briskman, & O'Connor, 2014), and such trials would introduce considerable additional heterogeneity into the analyses. In this way, we aimed to minimize the risk of examining transportability of interventions with uncertain effects in their origin country. Second, we recognize that by subgrouping trials based on country-level cultural and policy factors, within a modest-sized meta-analysis, based on individual-level trial outcomes, we can make only very cautious exploratory inferences. Using systematic review and meta-analysis, the objectives of this review were to (a) assess the effectiveness of evidence-based parenting interventions for ameliorating youth problem behavior, when transported to countries different from that in which the intervention was originally developed and/or tested, and (b) conduct an exploratory analysis of country-level contextual factors that may influence the success of such transportation, including cultural factors related to child-rearing and family life, resources, and family policies.

METHOD

Identifying Studies for This Review

Inclusion criteria. Inclusion of studies was determined in two stages. Stage 1 involved compiling a list of evidence-based parenting interventions, which were those meeting the following criteria. First, they were designed to prevent or treat child conduct or behavior problems and aimed primarily at the parents of children

ages 3 to 10. Second, to ensure the effects were adequately robust, they had to have a minimum effect size of $d=0.5$, in one or more randomized controlled trials (RCTs), with a minimum of 25 participants per study arm. They also had to be manualized and have a clear theoretical basis (e.g., attachment or social learning theory). Third, trials of the intervention had to have at least one outcome measure related to reducing child problem behavior.

Stage 1 identified 20 parenting interventions, from six countries, that fit our criteria (Table S1). Stage 2 involved searching for all controlled trials of these 20 evidence-based parenting interventions that took place outside the country where the intervention was developed and first tested in an RCT.

In Stage 2 we included randomized, quasi-randomized, or nonrandomized trials with well-matched comparison groups, including no intervention, other interventions, or “treatment as usual.” “Well-matched” means that participants in both groups were similar in terms of demographic and behavioral characteristics, and any differences were accounted for in analyses of outcome. We planned to analyze nonrandomized trials separately, because of likely high heterogeneity and risk of bias (Higgins & Green, 2011).

Behavioral problems are most easily identified from the age of 3, and parenting interventions are most developmentally appropriate and appear to be most effective for children ages 3 to 10. Thus, participants included children (ages 3–10) identified as having conduct problems based on behavior scores (above the clinical cut-off), referral to a specialist mental health center, or diagnosis. Interventions were those identified in Stage 1 as evidence-based parenting interventions, aimed at treating child conduct or behavior problems, which have been tested in a country (transported-to country; defined by World Bank, 2012) other than that where it was originally designed and tested (origin country). Trials of interventions in transported-to countries were included if they were clear replications of the original intervention. Thus, those that combined components of an origin intervention plus components of other interventions were included if all or more than 50% of the core components of the intervention were implemented and the effects of the two approaches could be isolated from each other, *or* the additional components represented less than 20% of the components delivered, or dosage of the intervention. The primary outcome of interest was reducing child/youth conduct problems or externalizing behavior. However, trials were not excluded based on outcome measures (Higgins & Green, 2011). As well as including nonrandomized trials, we placed no lower limit on sample size. Thus, in Stage 2 we set the inclusion bar slightly lower than Stage 1, to maximize included trials.

Search methods. Electronic databases were searched for published and unpublished reports of transported trials of each of the parenting interventions identified in Stage 1 up to November 2011 (e.g., PsycINFO, MEDLINE, EMBASE, CINAHL; Global Health; and LILACS; Table S2). Reports of trials published after November 2011 were identified through grey literature searches and contacts with intervention developers. No language restrictions were imposed on any search results, although most databases were searched in English. Latin American literature (LILACS) was also searched using Spanish terms.

Unpublished or nonindexed reports were searched through Google and Google Scholar, and searches of websites for each of the evidence-based interventions. Bibliographies of articles identified through searches were examined to identify further studies. In addition, we contacted the developers of each intervention for information about transported trials. This helped to ensure identification of all relevant trials, including those not yet published.

Data collection and analysis. One author (WK) assessed abstracts, and two authors (WK, FG) assessed full text of studies that were likely to meet inclusion criteria; the third author assessed discrepancies. We assessed risk of bias in included studies (as “high,” “low,” or “unclear”) using the Cochrane Collaboration tool (Higgins & Green, 2011). We added a category: whether the trial was conducted by researchers independent of the developer (Eisner, 2009).

We analyzed data if means and standard deviations were available, or if we could calculate effect sizes from other data (e.g., *t* tests), and contacted authors for missing data. We present all continuous data as standardized mean differences (SMD) in pooled analyses from similar instruments. We used 95% confidence intervals (CIs) for individual study data and pooled estimates. Given potentially high heterogeneity, we used random-effects models. We examined effects at postintervention, which varied from 1 to 12 months. We assessed heterogeneity visually and using the I^2 statistic, which describes the approximate proportion of variation due to heterogeneity rather than sampling error. An I^2 statistic of 30%–60% is frequently interpreted as moderate, and 50%–90% as substantial heterogeneity (Higgins & Green, 2011). Sensitivity and subgroup analyses investigated possible sources of heterogeneity. When sufficient data were available, we undertook subgroup analyses to assess the extent to which effect size varied by characteristics of study participants, intervention brand or format, service offered to control group, staffing, implementation fidelity, and whether the trial tested efficacy or effectiveness. Second, we conducted subgroup analyses on country-level factors related to the

cross-country transportation of interventions, including sociocultural values, world region, and family policy context. This review follows the standards of the PRISMA Statement (Moher, Liberati, Tetzlaff, & Altman, 2009; Table S4).

RESULTS

Our Stage 2 search strategy identified 4,179 citations, of which 92 were deemed potentially eligible based on the title or abstract; we obtained full-text copies of these 92 studies. After reviewing full-text copies, we included 17 trials of four interventions (Table 1; Figure 1). The studies were conducted over a 14-year period (1998–2012) and, at the time of analysis, all but four were published (the trial by Leung & Tsang [2012] has subsequently been published [Leung, Tsang, Sin, Choi, 2015], but data for analyses were taken from the 2012 report). We excluded 46 studies because they were prevention rather than treatment trials or children were not clearly diagnosed or referred for behavior problems; children were younger than 3 or older than 10 years, the intervention included components of different interventions, or the intervention and control groups were not well matched.

Included Studies

The trials took place in 10 countries: one each in Canada, Iceland, Iran, Ireland, Sweden, Holland, and Puerto Rico; two in Norway; three in Hong Kong and five in the United Kingdom. Thus most took place in Europe or North America, and a few in Latin America, Asia, and the Middle East. All of the

TABLE 1
Included Trials, Intervention, and Implementation Country

Country	Intervention	Trial Reference
Canada	Incredible Years	Taylor et al. (1998)
Hong Kong	Triple P	Leung et al. (2003)
	PCIT	Leung and Tsang (2012)
	PCIT	Leung et al. (2009)
Iceland	PMTO	Sigmarsdóttir et al. (2013)
Iran	Triple P	Jalali et al. (2009)
Ireland	Incredible Years	McGilloway et al. (2012)
Netherlands	Incredible Years	Posthumus et al. (2012)
Norway	Incredible Years	Larsson et al. (2009)
	PMTO	Ogden and Hagen (2008)
Puerto Rico	PCIT	Matos et al., (2009)
Sweden	Incredible Years	Axberg and Broberg (2012)
UK	Incredible Years	Gardner et al. (2006)
	Incredible Years	Hutchings et al. (2007)
	Incredible Years	Scott et al. (2001)
	Triple P	Berry et al. (2012) ^a
	Incredible Years	Morpeth et al. (2012) ^a

^aSome results from these studies appeared in Little et al. (2012).

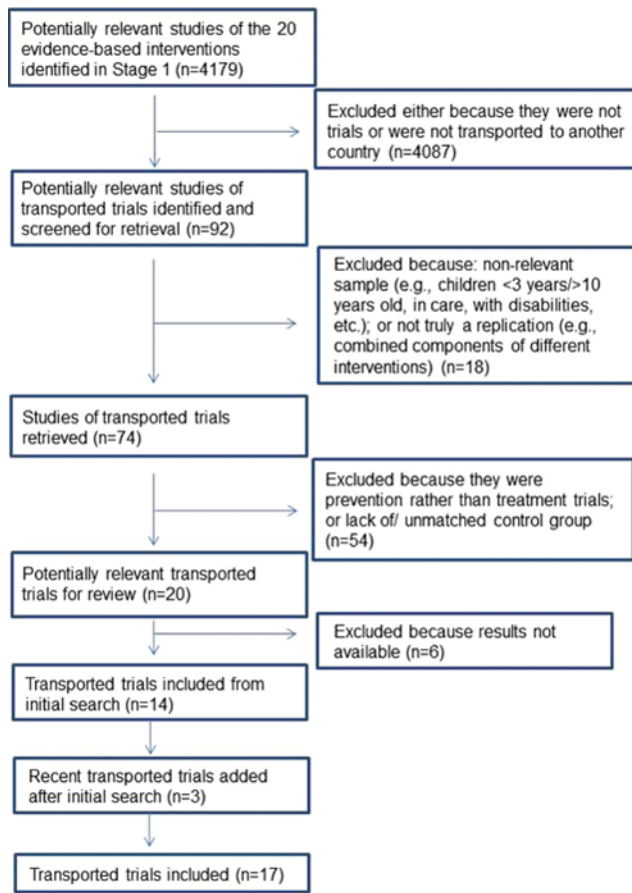


FIGURE 1 Study flow diagram for transported trials.

transported-to countries are categorized as high income (World Bank, 2012), with the exception of Iran, a middle-income country. No studies were found in low-income countries.

A total of 1,558 parent-child dyads were included in the 17 trials. Sample sizes ranged from 21 to 161 ($M = 92$). The 17 included studies tested four parenting interventions (designated as evidence-based in Stage 1, Table S1); three were originally designed and tested in the United States (Incredible Years; Parent-Child Interaction Therapy [PCIT]; Parent Management Training Oregon [PMTO]) and one in Australia (Triple P). Incredible Years and Triple P were trialed as group-based programs, whereas PMTO and PCIT were delivered to individual families.

All four interventions had similar content and theoretical underpinnings, based on social learning theory principles. They differed somewhat in the degree to which a “collaborative” model of working with parents is made explicit in the training of therapists and implementation of the program. For example, with Incredible Years, group leaders are certified according to criteria related to tailoring the program to match individual families’ goals and values; parents are encouraged to generate the parenting principles and strategies

that the group will use, rather than these being taught didactically by the therapist. Although this approach is fundamental to the program for all parents, it is especially applicable to working with mixed cultural groups (Webster-Stratton, 2009). In contrast, PCIT, an individually delivered program, has a much more “scripted” set of parenting strategies and criteria that have to be met in order to move to the next stage. However, it should be noted that in one respect PCIT is individually tailored, in that the pace at which these stages are reached is tailored to each family.

All but three of the 17 included studies used a RCT design. In the included studies, mean child age was 5.6 years (range = 3.5–8.4), and most were boys (range = 58%–100% in each study). The primary caregiver was normally the mother, and only rarely was outcome data reported from fathers. Socioeconomic status (SES) of participants varied widely among studies: Four described participant families as low income (Gardner et al., 2006; Hutchings et al., 2007; McGilloway et al., 2012; Scott et al., 2001), and four as median income to low income or status (Leung & Tsang, 2012; Leung, Tsang, Heung, & Yiu, 2009; Ogden & Hagen, 2008; Sigmarsdóttir, Degarmo, Forgatch, & Guðmundsdóttir, 2013). Two studies described families as middle to high status (Jalali, Shaeeri, Tahmasian, & Pourahmadi, 2009; Leung et al., 2003). Five trials provided no information about income or SES (Berry et al., 2012; Larsson et al., 2009; Matos, Bauermeister, & Bernal, 2009; Morpeth et al., 2012; Posthumus, Raaijmakers, Maassen, van Engeland, & Matthys, 2012). In all but three studies (Berry et al., 2012; Morpeth et al., 2012; Scott et al., 2001), the great majority of families were ethnically native to the country in which the trial took place. This homogeneity across studies limited the ability to make comparisons by ethnicity. Children were referred by health and social care professionals, schools, and parents. Most studies screened children for inclusion using clinical cutoff scores on parent report instruments for behavioral problems.

Eleven trials used a waitlist control; two used no-intervention control groups (Jalali et al., 2009; Leung et al., 2009); and three, TAU (Ogden & Hagen, 2008; Posthumus et al., 2012; Sigmarsdóttir et al., 2013). The three studies with a TAU condition provided descriptions of the services offered to control participants but few details on dosage or nature of treatment.

Outcomes

All included studies reported on child conduct problems using continuous data from parent reports, mainly using three common, well-validated instruments: ECBI Intensity and Problem subscales, CBCL Externalizing and ODD subscales, and SDQ Conduct subscale. Many trials had secondary outcomes (e.g., child internalizing

behavior, parent mental health) that were not directly relevant to this review. Several trials included outcomes based on observation of child behavior, using a variety of instruments (e.g., Dyadic Parent–Child Interaction Coding System; Eyberg & Robinson, 1981).

Risk of Bias in Included Studies (Table S3)

Our analyses showed that four RCTs (Berry et al., 2012; Hutchings et al., 2007; Morpeth et al., 2012; Ogden & Hagen, 2008) had a very low risk of bias. Three other RCTs (Gardner et al., 2006; Leung & Tsang, 2012; McGilloway et al., 2012) had a low to moderate risk of bias. Seven RCTs had an unclear risk of bias, with the most common reporting gaps related to sequence generation, allocation concealment, and blinding.

Effects of Interventions

Random-effects meta-analyses were conducted with RevMan 5.3, using the primary outcomes of parent-reported child conduct problems, using the scales just listed. Results are presented as effect sizes with 95% CI, for randomized and nonrandomized studies separately, as recommended by Higgins and Green (2011), and calculated using change scores, defined as mean difference between baseline and postintervention scores for each group. The postintervention point was normally 4 to 6 months postbaseline; additional longer term outcomes were assessed in six trials but were not analyzed, because in most trials comparison groups were lost. Narrative descriptions of results focus on analyses using change scores because they provide a better measure of effect for studies with small sample sizes (Higgins & Green, 2011). Effect sizes smaller than .20 were interpreted as indicating no evidence of effectiveness, those above .20 as small (.20–.40), moderate (.40–.75) or large (>.75).

The 14 RCTs provided data for 1,258 participants (735 parent training; 523 control). Results favored parent training, indicating significant moderate benefits to child behavior, with confidence intervals indicating a range of effect sizes from ineffective to large (random effects model: SMD -0.71), 95% CI [-0.97, -0.44], $p < .00001$. The test for heterogeneity was significant, $Q(13) = 61.21, p = .00001, I^2 = 79%$ (Figure 2). Findings were very similar using posttest scores (SMD -0.70), 95% CI [-0.99, -0.40], $p < .00001, I^2 = 83%$ (Figure S1).

Sensitivity analysis assessed the effects of different control conditions by removing two studies that used a TAU rather than waitlist control condition (Ogden & Hagen, 2008; Sigmarsdóttir et al., 2013). This can have implications for pooled results in meta-analyses, as TAU can represent a wide range of treatments of varying dosages (Löfholm, Brännström, Olsson, & Hansson, 2013). This had minimal effect on level of heterogeneity, which increases confidence in the results ($Z = 4.96, p < .00001, Q(11) = 55.37, p < .00001, I^2 = 80%$).

Direct observational measures of parent and child behavior are considered important for validating parent-report data (Aspland & Gardner, 2003). Only four of the RCTs in this review used observational measures of child behavior, and we meta-analyzed them using pooled outcomes to further assess overall effectiveness. Analysis of observed negative child behavior indicates a nonsignificant effect of parent training (SMD -0.21), 95% CI [-0.61, 0.20], $p = .32$. There were small to medium effects in the expected direction in three trials but effects in the opposite direction in the Norwegian PMTO trial. An analysis of observed positive child behavior indicates a nonsignificant effect (SMD -0.33), 95% CI [-0.80, .015], $p = .17$.

There were three nonrandomized trials, two using ECBI scales (Leung et al., 2009; Posthumus et al., 2012) and one the SDQ Conduct Scale (Scott et al., 2001), with a total

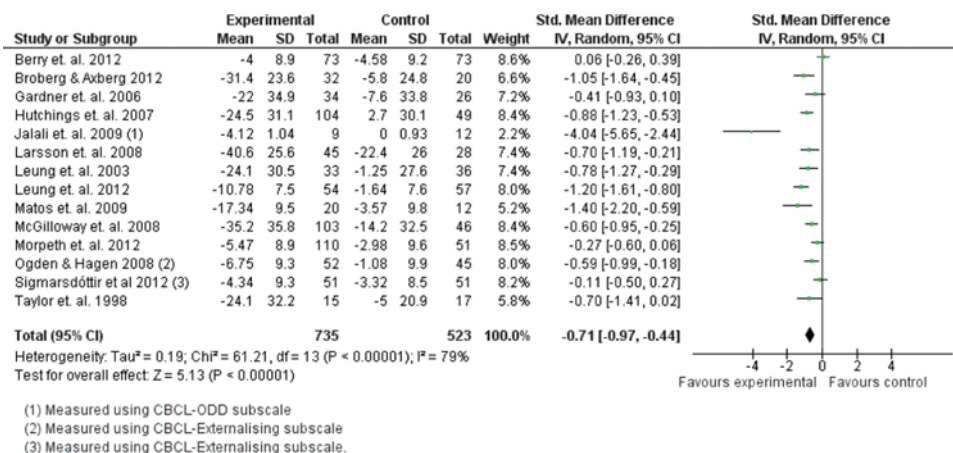


FIGURE 2 Effect sizes for pooled outcome data of randomized controlled trials, using change scores based on Eyberg Child Behavior Inventory–intensity scores or Child Behavior Checklist.

sample size of 300. The results indicated nonsignificant benefits of parent training, with CIs suggesting a range of small to large effect sizes (SMD -1.07), 95% CI $[-2.48, 0.33]$, $p = .13$. The test for heterogeneity was significant, $Q(2) = 58.15$, $p < .00001$, $I^2 = 97\%$. However, because heterogeneity was so high and the range of effects so wide (from small [$d = 0.12$] to medium [$d = 0.57$] to very large [$d = 2.55$]), findings are not interpretable, and therefore we do not report further on sensitivity or subgroup analyses for nonrandomized trials.

Subgroup analyses 1. We first analyzed potential explanatory factors related to sample and intervention characteristics. We found no significant differences in effect sizes based on SES, parent education level, or median child age, based on the 14 randomized trials. Subgroup analyses showed no differences in outcomes based on characteristics of the intervention, and control group services (TAU or no intervention), whether staff were more or less qualified, and whether the trial was conducted in an efficacy or “real-world” setting. Thirteen of the 17 studies provided some information about measures taken to enhance implementation fidelity. Eleven reported use of clinical supervision of those delivering the program, and 10 studies reported using video of therapy sessions to enhance supervision. Six studies used staff self-report, three used peer review of performance, and four used booster training. Subgroup analysis showed no difference in effect sizes by whether studies reported one or two fidelity indices ($d = -0.50$) compared to three or four ($d = -0.57$). Although trial reports for three RCTs and one nonrandomized study described direct assessment of fidelity, lack of information in other trial reports precluded synthesizing this measurement information. We found a significant difference in effect by subgroup of intervention brand, $\chi^2(3) = 11.91$, $p = .008$. Seven of the 14 RCTs tested Incredible Years; two PCIT, three Triple P, and two PMTO. Mean effect sizes were significant for Incredible Years ($d = 0.63$, $p = .000$) and PCIT ($d = 1.24$, $p = .000$), and nonsignificant for PMTO ($d = 0.35$, $p = .14$), and Triple P ($d = 1.26$, $p = .06$).

Subgroup analyses 2: Factors related to transportability. Exploratory subgroup analyses were performed to address Question 2: What factors influence the success of transportation of evidence-based parenting interventions between countries? We focus here on the randomized studies.

Country-level socioeconomic context. The United Nations Development Program’s Human Development Index (HDI) includes not only economic measures but three dimensions and four indicators of health

(life expectancy at birth), education (mean years of schooling), and living standards (gross national income per capita). All the countries represented in this review are considered Very Highly Developed, except Iran, which is ranked as having High Development, and Puerto Rico, for which no data were available. We performed a subgroup analysis of RCTs, comparing effect sizes from trials that took place in countries ranked in the top 15 on the HDI (eight studies, $MES d = 0.70$), to countries ranked between 16 and 30 (four studies, mean $ES d = 0.37$), to the trial from Iran, a country that ranks 88 ($ES d = 4.0$). The subgroup analysis showed a significant difference based on level of development (SMD -0.66), 95% CI $[-0.94, -0.39]$, $p < .0001$. However, this is hard to interpret as it depends on the trial from Iran, which had a very small sample size and a large effect size. When the trial from Iran is removed, the difference between trials in the Very Highly Developed and Highly Developed groups is nonsignificant. For comparison, the origin countries for these interventions, United States (HDI = 4) and Australia (HDI = 2), fall in the top 15 HDI group.

We also assessed the impact of socioeconomic context in terms of relative child poverty. Although comparable data could not be found for Hong Kong, Iran, and Puerto Rico, it was available for the other countries, providing a useful comparison between countries that are otherwise similar on many other factors. Using data from UNICEF, this subgroup analysis looked at percentage of children living in a household where disposable income, adjusted for family size and composition, is less than 50% of the national median income. Meta-analysis found no significant difference between RCTs in countries with a higher percentage of children living in relative poverty ($d = 0.41$), compared to those with a lower percentage ($d = 0.57$). Notably, relative child poverty was markedly lower in the countries for which data were available, compared to the United States (23%) where three of the four parenting interventions were developed. Australia’s relative child poverty rate was also in the higher category (10.9%).

Cultural context. There was a significant difference in effect sizes of RCTs that took place in countries with “Western” or Anglo/European cultural roots ($d = 0.49$) compared to studies which took place in Asia, Latin America, and the Middle East ($d = 1.50$; Figure 3). When we removed the study from Iran (which had the largest effect size and smallest sample size) the difference between the groups remained significant ($p = .004$), with higher effect sizes ($d = 1.08$) in the trials in the non-Western countries.

Cultural context was also assessed using data from the World Values Survey (WVS), which provides

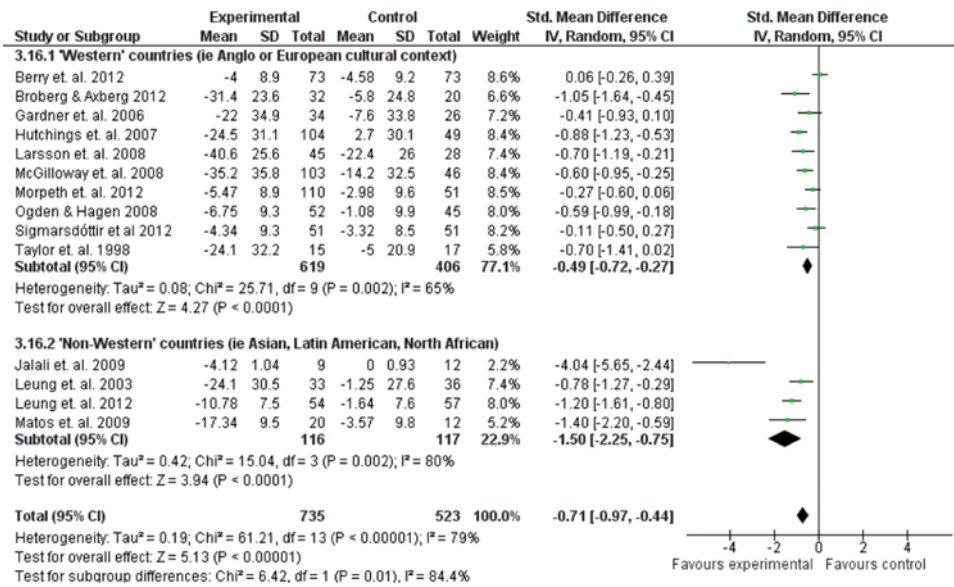


FIGURE 3 Subgroup analysis: “Western” compared to “non-Western” countries (randomized controlled trials, pooled outcome, change scores).

representative national data on sociocultural and political values and beliefs from 97 countries. As it provides data on religion, family, and social life, it is a highly relevant source of information about factors that might affect parenting and parent–child relationships, compared to purely economic measures of cross-national difference (WVS, 2008). According to the WVS (2008), two dimensions measured in the survey account for more than 70% of cross-cultural variance: “traditional versus secular-rational values” and “survival versus self-expression values.” The former dimension contrasts societies in which religion is very important and those in which it is not. Societies that are more traditional are more likely to emphasize the centrality of parent–child ties, respect for authority, and traditional family values, and to oppose divorce, abortion, and suicide. Societies that fall within the secular–rational side are less likely to rate parent–child ties and deference to authority as being influential to their values.

Most RCTs took place in countries scoring as more secular, whereas four (Iran, Puerto Rico, Ireland, Canada) scored in the more traditional range. Subgroup analysis found no significant difference ($p = .10$) between the groups with regard to child behavior outcomes (Figure S2). The group of trials in the more traditional countries had a large effect size ($d = 1.4$), whereas trials in the secular group of countries had a medium effect size ($d = 0.57$). When the study from Iran is removed, the effect size for trials in traditional countries was smaller ($d = 0.79$).

The other major dimension in the WVS is survival versus self-expression. As countries become more affluent or industrialized, people take survival for granted, which shifts attention from economic and physical security

toward subjective well-being, self-expression, and other aspects of quality of life. Countries that score highly on self-expression, for example, prioritize diversity and gender equality, and tend to demand greater participation in political decision making. Imagination and tolerance are more likely to be seen as important values in childrearing. In more survivalist-oriented societies, hard work is often seen as the most important value to teach children. Our analyses found a significant group difference ($p = .04$) in the effects of parenting programs, with larger effects in the trials in more survival-focused countries ($d = 1.62$), compared to those in the more self-expression focused group ($d = 0.54$; Figure 4). By removing the study from Iran, the effect size for trials in survival-focused countries ($d = 1.01$), remains significant ($p = .05$).

The WVS also separates societies into eight regions based on shared cultural values. The studies in this review fell into five of these regions: Protestant Europe, English speaking, Latin America, Islamic, and Confucian. There was a significant difference between the groups ($p < .0001$), with the largest effect sizes seen in the trials in countries in Islamic, Latin American, and Confucian cultural groups. By comparison, all of the interventions (Incredible Years, PCIT, PMTO, and Triple P) originated in English-speaking countries.

Family policy context. We analyzed the differences in the effects of parenting programs between countries based on two categories of “family-friendly” policies: national-level public spending on family benefits as a percentage of gross domestic product (GDP), and the number of weeks of parental paid leave. We used data from the Organisation for Economic Co-operation

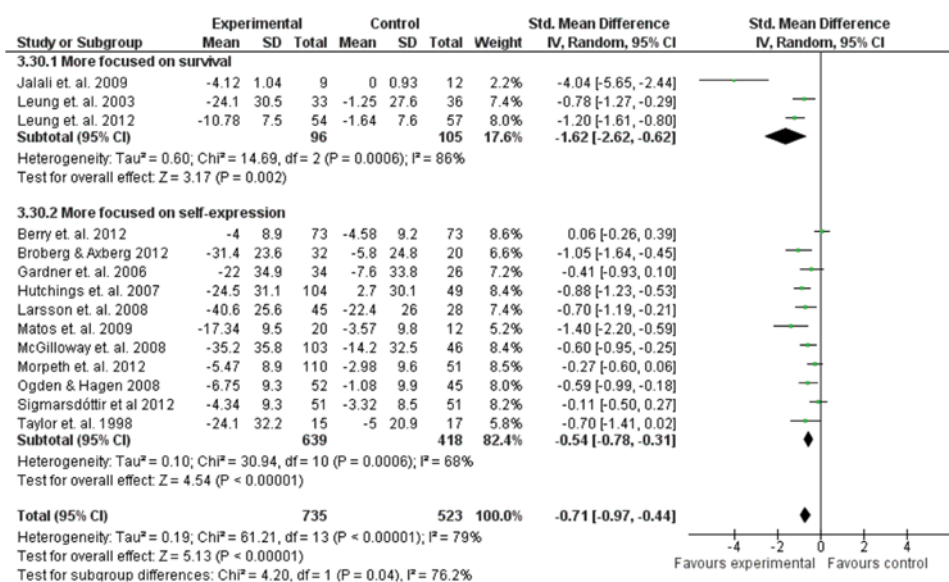


FIGURE 4 Subgroup analysis: Survival–self-expression values dimension (randomized controlled trials, pooled outcome, change scores).

and Development (OECD) Family Database, which provides raw data on the 34 OECD countries (OECD, 2012). Twelve of the trials in our review took place in OECD countries (Canada, Iceland, Ireland, the Netherlands, Norway, Sweden, United Kingdom). Available data were from the period 2007–2008. No data were available for Hong Kong, Iran, and Puerto Rico.

Nine RCTs took place in OECD countries that rank as higher spending in terms of percentage of GDP spent on family benefits, and one RCT took place in an OECD country (Canada) ranked as lower spending in terms of family benefits (Taylor, Schmidt, Pepler, & Hodgins, 1998). Our analyses found no significant difference in trial effect sizes between countries based on spending on family benefits (Figure S3). Most of the studies tested interventions which originated in the United States, where expenditures are among the lowest in the OECD (1.2% of GDP; the United States ranks 30th out of 34 OECD countries; Australia spent 2.7% of GDP on family benefits in 2007 and ranks 13th among OECD countries). Despite this difference, effect sizes remained robust across the countries represented in our review.

Five RCTs took place in OECD countries with 0 to 19 weeks paid parental leave (range = 6.6 weeks in Ireland to 12.8 weeks in the United Kingdom), and five others took place in countries with 20 to 40 weeks guaranteed paid leave (range = 20.8 weeks in Iceland to 38.8 in Norway). No significant differences were found in trial effect sizes between these groups of countries. It is notable that the OECD average is 30 weeks of paid leave, while the average number of weeks of guaranteed paid leave in the origin countries, United States and Australia, is zero. Again, despite this difference, effect

sizes remained robust across trials in the countries represented in our review.

Cross-cultural adaptation. From the trial reports, there did not appear to be extensive formal adaptation of the parenting programs. This is consistent with these programs being well-established and having clear systems of training and certification, which were then imported directly into new countries, in many cases replicating the same training systems and manuals. Only two trial reports mentioned small cultural adaptations; thus it was not possible to code variation in this factor or to conduct subgroup analysis. However, we meta-analyzed based on whether program materials were translated from their original language (English) to the language of the trial country. Although this is far from ideal, as it is a relatively superficial adaptation, it does provide some index of the effect of modification on outcome. Eight of the RCTs were conducted after translating intervention materials into the language of the trial country, compared to six RCTs that took place in English-speaking countries. There was a difference between the groups ($p = .05$): Trials using translated materials had a large effect size ($d = 0.96$), whereas those in English-speaking countries had a medium effect size ($d = .44$).

DISCUSSION

Parenting interventions are increasingly being recommended and implemented in diverse regions of the world (UNODC, 2009; WHO, 2010), as part of strategies for preventing violence—both—to and by children—and

enhancing children's development. This raises the question of whether countries should import interventions from other countries, and if so, whether they are likely to be appropriate and effective. This novel systematic review aimed to address this by examining the extent to which evidence-based parenting programs for reducing child behavioral problems are effective when transported to new countries. We also conducted exploratory subgroup analyses to examine whether effectiveness of transported interventions was associated with country-level factors, including cultural values linked to childrearing and family life, and factors related to national resources and family policies.

We found 14 randomized and three nonrandomized trials of transported evidence-based parenting interventions for reducing child behavioral problems. Four programs that originated in the United States or Australia were tested in 10 countries in five regions (Europe, Asia, North America, the Middle East, and the Caribbean), although most trials (11 of 17) were conducted in Northern Europe. In meta-analyses, data from randomized trials showed evidence of strong, highly significant effects of parenting interventions on child problem behavior in transported-to countries. Effect sizes for the three nonrandomized trials were highly variable and nonsignificant overall.

For the 14 RCTs, we examined trial-level factors and found no association between effect size and family SES, parent education level, or child age. Also, effect size was not associated with program "brand," delivery format (group vs. individual), indices of implementation fidelity, the level of service or "treatment as usual" offered to control groups, staff qualifications, or whether the trial was conducted in an "efficacy" or "effectiveness" context.

Subgroup analyses should be viewed with caution, given the number of included trials. Nevertheless, these exploratory analyses of country-level factors appear to reveal patterns that are somewhat counterintuitive, given that much literature suggests that parenting values, practices, and interventions are highly culture-dependent. Interventions transported from the United States and Australia to other "Western" countries showed comparable effect sizes to those obtained in the origin country. However, effect sizes appeared stronger when the same interventions were transported to culturally more distant regions, namely, Asia, Latin America, and the Middle East. Trials in countries with more traditional values related to childrearing and family life, as measured by the WVS, tended to show higher effect sizes than those in less traditional countries. These traditional countries are also culturally distant from the origin countries (United States, Australia), which fall into the nontraditional category. We found no differences in effect sizes for trials in countries with high versus low spending on family benefits, or with more or

less "family-friendly" policies, or by level of child poverty. All of the countries apart from Iran are classified by United Nations Development Program's HDI as Very Highly Developed. Within the very high HDI countries, there was no association between level of development and effect size. For comparison, the United States, where most of the interventions originated, has higher rates of child poverty and lower spending on family benefits than most countries in the OECD, and zero weeks of guaranteed parental paid leave. It was striking that despite huge cultural, economic, and policy differences between transported trials, and between "origin" and transported-to countries, trial effect sizes either were consistent across subgroups or tended to favor interventions in those countries that are culturally more distant compared to the origin countries.

Interpretation

Although reviews and trials present somewhat mixed conclusions on this issue (Barrera et al., 2011; Gottfredson et al., 2006; Huey & Polo, 2008; Wilson & Miller, 2003), a dominant (and plausible) view is that parenting interventions will be effective in new cultural contexts only if there is an extensive multistage adaptation process (Barrera & Castro, 2006; Kumpfer et al., 2008), or if there is limited cultural distance between the countries, as hypothesized by Sussman et al. (2008). However, our findings present a rather more optimistic view. Part of the explanation for the cross-cultural success of parenting interventions might be that their basic principles (e.g., parent-child relationship building through play and positive attention, child behavior change through social learning) are universal across cultures. It appears that in most of the trials in this review, the interventions were implemented with fidelity to the imported manual and training methods, with only minor adaptations for new countries. Unfortunately, the trial reports include few details about these minor adaptations; only one (Matos et al., 2009) made reference to literature on frameworks for cultural adaptation; another, Leung and Tsang (2012), working with Chinese parents, explained how they varied the form of praise that parents gave to children, so that it was less direct and effusive than in Western cultures.

It should also be noted that, although training was often conducted by imported experts, the parenting programs were generally initiated and implemented by senior local practitioners and researchers. It remains unclear whether these findings would be replicated in low-income countries, where there is more limited local professional capacity, and where "foreign" nongovernmental organizations and staff may often be involved in initiating, training, and implementing the intervention.

For one of the programs—Incredible Years, which was tested in half of the included trials—there is extensive discussion and data about cultural factors (Reid et al., 2001; Webster-Stratton, 2009). Incredible Years is based on a collaborative model involving explicit and flexible tailoring of parenting strategies to families' individual and cultural needs. Thus adaptation occurs at the level of each individual family, rather than by developing specific program versions for each country or culture. This “culturally flexible” approach is likely to be more suitable for the ethnically mixed communities found in many large urban areas (Scott et al., 2010; Webster-Stratton, 2009), especially where services are not built around specific ethnic groups. Moreover, its effectiveness is supported by data pooled from several U.S. trials (Reid et al., 2001), which found remarkably few differences by ethnicity in outcomes, satisfaction, or engagement. Just as the interventions in this review mostly appeared to transport well to countries with very different cultural values, they also transported successfully to countries with very different levels of resources and policy regimes. This suggests that their success may not be dependent on a close match of policy traditions or practices between the origin country and the new country, as long as the program is suitably flexible.

Our results differ somewhat from other (narrative) reviews on cross-country transportability. For example, Ferrer-Wreder et al. (2012) commented that imported evidence-based interventions made up 40% of psychosocial interventions tested in Sweden, and yet in some cases these imported programs failed to show expected benefits, including Sundell's trial (Gustle et al., 2008; Sundell et al., 2008) of Multisystemic Therapy, which was imported from the United States and aimed at teenage delinquents. Ferrer-Wreder et al. suggested this might be related, in part, to the high quality of regular services for young offenders in Sweden compared to the United States. This could mean that other trials comparing a new intervention to existing TAU might be less likely to show effects and that new approaches would be needed to improve outcomes for troubled youth in Sweden. We were able to compare interventions where the control group received an active TAU versus ones that received nothing. Of interest, we did not find any significant difference between the subgroups, although of course we were not able to make nuanced distinctions based on the quality of the comparison treatment.

We found very few examples of evidence-based parenting interventions that had been replicated in low- and middle-income countries, despite the fact that various other generic parenting interventions have been tested in small randomized trials and found to be promising in these settings (Knerr et al., 2013). However, these trials were not included in the present review,

because they either had not been transported from another country or had not been rigorously tested in their origin country. An important question, though, is why the included interventions tended to be more effective in non-Western countries, many of which offer less well-developed family services. Of interest, the findings of our review have at least some parallels with a recent meta-analysis (Panagiotou, Contopoulos-Ioannidis, & Ioannidis, 2013) comparing the effects of medical interventions in high- versus middle- and low-income countries. That study found higher effect sizes in lower income countries, which the study authors interpreted as a sign that risk of bias in the trials was higher in those settings. Although we cannot rule out this possibility in our review, we suggest that more attention be given to reasons why interventions might genuinely work better in settings where the culture or resources differ from the countries in which those interventions were developed. One possible explanation is that parents in more traditional cultures might be more responsive and respectful to perceived experts and therefore engage more willingly, and learn more, from the intervention. Alternatively, where fewer family services are available, under conditions of rapid social transition, or where there are higher levels of stress caused by living in lower resource contexts, parents may be more receptive or there may be more room for change. This is consistent with recent analyses of moderator effects in parenting intervention trials, where low-income parents experiencing more stress and higher levels of child problem behavior, often show more change following intervention (Gardner, Hutchings, Bywater, & Whitaker, 2010; Leijten, Raaijmakers, de Castro, & Matthys, 2013).

The study has several limitations. First, given that the subgroup analyses are based on 14 trials, with quite high levels of heterogeneity and low power, we view these comparisons as exploratory and interpret their findings with caution. Despite the limitations imposed by high heterogeneity, we note the consistent pattern of trends toward higher effects in countries with lower levels of development and in countries that are culturally more distant from the origin countries. Some of the studies were very small or poorly reported, and thus had unclear risk of bias. The study from Iran, with the highest effect size, was also the smallest and had an unknown or high risk of bias. We explored the effect of removing this study from relevant analyses and generally found a similar pattern of results. The quality of implementation was not always clear; although most trials reported on efforts to enhance fidelity, and we were able to subgroup studies by this factor, few directly assessed fidelity of delivery, meaning this may have been an unmeasured source of variation. Second, there are many possible ways to classify country-level characteristics along cultural dimensions and in terms of resources and policies;

but most of the classification schemes available were not applicable to all the countries in our review. We chose systems that were well used and validated; were representative; covered as many of the countries as possible; and, in the case of the WVS, had the most relevance to parenting and family values. We recognize that country-level cultural classifications reflect an average value for a large population, and we cannot tell if the families included in the trials share those average values. It is worth noting that in 12 of the 14 RCTs, almost all participants were from the ethnic majority for that country. Unfortunately, numbers of families from ethnic minorities were too small for further analysis. Finally, it was beyond the scope of this review to conduct meta-analyses of the many trials of these four interventions in their origin countries; rather, the primary aim was to examine the effectiveness of transported interventions, and factors predicting variation in these effects. However, to allow some comparison between transported-to and origin countries, we were able to use existing systematic reviews to provide an estimate of effect sizes for each of the four interventions in their countries of origin.

CONCLUSIONS

These limitations need to be seen in the context of the study's strengths. It is, to our knowledge, the first systematic attempt for any psychosocial intervention to address questions about the extent of cross-country transportability and factors predicting its success, using a well-defined question and rigorous methods of review. The question is highly topical for policy and practice internationally (Wessels et al., 2013; WHO, 2010). The findings are intriguing in that they appear to be at odds with the common, and arguably highly plausible, view that interventions will be most effective when transported to countries that are more similar culturally, and in terms of service provision, to those in which they were first developed. Given that these were well-established, manualized interventions, with training and certification methods which appear to have been imported largely intact from another country, we cautiously suggest that there may be no need for deep or extensive adaptation of the programs when transported from one country to another. Although these results may only be applicable to high-income countries, our review included trials from various world regions, which differed widely in cultural traditions and family policy regimes. Therefore this review should be viewed as an encouraging contribution to the growing international literature on cultural transportability of psychosocial interventions for child well-being and development.

FUNDING

This research was funded by Knut Sundell and the Swedish Board of Health and Welfare.

SUPPLEMENTAL MATERIALS

Supplemental data for this article can be accessed on the publisher's website at <http://dx.doi.org/10.1080/15374416.2015.1015134>.

REFERENCES

(References marked with an asterisk indicate studies included in the meta-analysis.)

- Albert, I., Trommsdorff, G., & Mishra, R. (2007). Parenting and adolescent attachment in India and Germany. In G. Zheng, K. Leung, & J. Adair (Eds.), *Perspectives and progress in contemporary cross-cultural psychology* (pp. 97–108). Beijing, China: China Light Industry Press.
- Aspland, H., & Gardner, F. (2003). Observational measures of parent-child interaction: An introductory review. *Child and Adolescent Mental Health*, 8, 136–143. doi:10.1111/1475-3588.00061
- *Axberg, U., & Broberg, A. G. (2012). Evaluation of “The Incredible Years” in Sweden: The transferability of an American parent-training program to Sweden. *Scandinavian Journal of Psychology*, 53, 224–232. doi:10.1111/j.1467-9450.2012.00955.x
- Backer, T. (2001). *Finding the balance—program fidelity and adaptation in substance abuse prevention: A state-of-the-art review*. Rockville, MD: CSAP.
- Baker-Henningham, H., & Walker, S. (2009). A qualitative study of teacher's perceptions of an intervention to prevent conduct problems in Jamaican pre-schools. *Child: Care, Health and Development*, 35, 632–642. doi:10.1111/j.1365-2214.2009.00996.x
- Barlow, J., Johnston, I., Kendrick, D., Polnay, L., & Stewart-Brown, S. (2006). Individual and group-based parenting programmes for the treatment of physical child abuse and neglect. *Cochrane Database of Systematic Reviews*, 2006(3), 1–23.
- Barrera, M., & Castro, F. G. (2006). A heuristic framework for the cultural adaptation of interventions. *Clinical Psychology: Science and Practice*, 13, 311–316. doi:10.1111/j.1468-2850.2006.00043.x
- Barrera, M., Jr., Castro, F. G., & Steiker, L. K. H. (2011). A critical analysis of approaches to the development of preventive interventions for subcultural groups. *American Journal of Community Psychology*, 48, 439–454. doi:10.1007/s10464-010-9422-x
- Belfer, M. L. (2008). Child and adolescent mental disorders: The magnitude of the problem across the globe. *Journal of Child Psychology and Psychiatry*, 49, 226–236. doi:10.1111/j.1469-7610.2007.01855.x
- *Berry, V., Blower, S., Morpeth, L., Tobin, K., Taylor, R., ... Lehtonen, M. (2012). *The effectiveness and micro-costing analysis of the Triple P Parenting Programme: An independent UK randomised controlled trial*. Dartington, UK: Dartington Social Research Centre.
- Bonell, C., Oakley, A., Hargreaves, J., Strange, V., & Rees, R. (2006). Assessment of generalisability in trials of health interventions: Suggested framework and systematic review. *BMJ*, 333, 346–349. doi:10.1136/bmj.333.7563.346
- Bradford, K., Barber, B. K., Olsen, J. A., Maughan, S. L., Erickson, L. D., Ward, D., & Stolz, H. E. (2003). A multi-national study of interparental conflict, parenting, and adolescent functioning: South Africa, Bangladesh, China, India, Bosnia, Germany, Palestine,

- Colombia, and the United States. *Marriage & Family Review*, 35, 107–137. doi:10.1300/J002v35n03_07
- Burchett, H., Umoquit, M., & Dobrow, M. (2011). How do we know when research from one setting can be useful in another? A review of external validity, applicability and transferability frameworks. *Journal of Health Services Research & Policy*, 16, 238–244. doi:10.1258/jhsrp.2011.010124
- Cartwright, N., & Hardie, J. (2012). *Evidence-based policy: A practical guide to doing it better*. New York, NY: Oxford University Press.
- Castro, F. G., Barrera, M., & Holleran Steiker, L. K. (2010). Issues and challenges in the design of culturally adapted evidence-based interventions. *Annual Review of Clinical Psychology*, 6, 213–239. doi:10.1146/annurev-clinpsy-033109-132032
- Eisner, M. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology*, 5, 163–183. doi:10.1007/s11292-009-9071-y
- Eyberg, S., & Robinson, E. (1981). *Dyadic parent-child interaction coding system*. Seattle: Parenting Clinic, University of Washington.
- Ferrer-Wreder, L., Sundell, K., & Mansoor, S. (2012). Tinkering with perfection: Theory development in the intervention cultural adaptation field. *Child & Youth Care Forum*, 41, 149–171. doi:10.1007/s10566-011-9162-6
- 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 of Systematic Reviews*, 2012, CD008225. 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*, 47, 1123–1132. doi:10.1111/j.1469-7610.2006.01668.x
- Gardner, F., Hutchings, J., Bywater, T., & Whitaker, C. (2010). Who benefits and how does it work? Moderators and mediators of outcome in an effectiveness trial of a parenting intervention. *Journal of Clinical Child & Adolescent Psychology*, 39, 568–580. doi:10.1080/15374416.2010.486315
- Gottfredson, D., Kumpfer, K., Polizzi-Fox, D., Wilson, D., Puryear, V., Beatty, P., & Vilmenay, M. (2006). The Strengthening Washington D.C. Families project: A randomized effectiveness trial of family-based prevention. *Prevention Science*, 7, 57–74. doi:10.1007/s11121-005-0017-y
- Gustle, L.-H., Hansson, K., Sundell, K., & Andree-Löfholm, C. (2008). Implementation of evidence-based models in social work practice: Practitioners' perspectives on an MST trial in Sweden. *Journal of Child & Adolescent Substance Abuse*, 17, 111–125. doi:10.1080/15470650802071713
- Higgins, J., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions version 5.1.0* [Updated March 2011]. Retrieved from <http://www.cochrane-handbook.org>
- Huey, S. J., Jr., & Polo, A. J. (2008). Evidence-based psychosocial treatments for ethnic minority youth. *Journal of Clinical Child & Adolescent Psychology*, 37, 262–301. doi:10.1080/15374410701820174
- Hutchings, J., Bywater, T., & Daley, D. (2007). Early prevention of conduct disorder: How and why did the North and Mid-Wales Sure Start study work? *Journal of Children's Services*, 2, 4–15. doi:10.1108/17466660200700012
- *Hutchings, J., Bywater, T., Daley, D., Gardner, F., Whitaker, C., Jones, K., ... 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. doi:10.1136/bmj.39126.620799.55
- *Jalali, M., Shaeeri, M. R., Tahmasian, K., & Pourahmadi, E. (2009). The effects of Triple P Positive Parenting Program on 7–10 year old children with oppositional defiant disorder (ODD). *Daneshvar Raftar*, 1, 29–38.
- Knerr, W., Gardner, F., & Cluver, L. (2013). Improving positive parenting skills and reducing harsh and abusive parenting in low- and middle-income countries: A systematic review. *Prevention Science*, 14, 352–363. doi:10.1007/s11121-012-0314-1
- Kumpfer, K. L., Alvarado, R., Smith, P., & Bellamy, N. (2002). Cultural sensitivity and adaptation in family-based prevention interventions. *Prevention Science*, 3, 241–246. doi:10.1023/A:1019902902119
- Kumpfer, K. L., Pinyuchon, M., de Melo, A. T., & Whiteside, H. O. (2008). Cultural adaptation process for international dissemination of the strengthening families program. *Evaluation & the Health Professions*, 31, 226–239. doi:10.1177/0163278708315926
- *Larsson, B., Fossum, S., Clifford, G., Drugli, M. B., Handegård, B. H., & Mørch, W.-T. (2009). Treatment of oppositional defiant and conduct problems in Norwegian children: Results of a randomized controlled trial. *European Child & Adolescent Psychiatry*, 18, 42–52. doi:10.1007/s00787-008-0702-z
- Lau, A. S. (2006). Making the case for selective and directed cultural adaptations of evidence-based treatments: Examples from parent training. *Clinical Psychology: Science and Practice*, 13, 295–310. doi:10.1111/j.1468-2850.2006.00042.x
- Leijten, P., Raaijmakers, M. A. J., de Castro, B. O., & Matthys, W. (2013). Does socio-economic status matter? A meta-analysis on parent training effectiveness for disruptive child behavior. *Journal of Clinical Child & Adolescent Psychology*, 42, 384–392. doi:10.1080/15374416.2013.769169
- *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
- *Leung, C., & Tsang, S. (2012). *Parent-Child Interaction Therapy (PCIT) Service in Hong Kong: An efficacy and effectiveness study report*. Hong Kong: Tung Wah Hospitals.
- *Leung, C., Tsang, S., Heung, K., & Yiu, I. (2009). Effectiveness of parent-child interaction therapy (PCIT) among Chinese families. *Research on Social Work Practice*, 19, 304–313. doi:10.1177/1049731508321713
- Leung, C., Tsang, S., Sin, T. C. S., & Choi, S. (2015). The efficacy of parent-child interaction therapy with Chinese families: Randomized controlled trial. *Research on Social Work Practice*, 25(1), 117–128. doi:10.1177/1049731513519827
- Littell, J. H. (2006). The case for multisystemic therapy: Evidence or orthodoxy? *Children and Youth Services Review*, 28, 458–472. doi:10.1016/j.childyouth.2005.07.002
- *Little, M., Berry, V., Morpeth, L., Blower, S., Axford, N., Taylor, R., ... Tobin, K. (2012). The impact of three evidence-based programmes delivered in public systems in Birmingham, UK. *International Journal of Conflict & Violence*, 6, 261.
- Löfholm, C. A., Brännström, L., Olsson, M., & Hansson, K. (2013). Treatment-as-usual in effectiveness studies: What is it and does it matter? *International Journal of Social Welfare*, 22, 25–34. doi:10.1111/j.1468-2397.2012.00870.x
- Malti, T., Ribeaud, D., & Eisner, M. P. (2011). The effectiveness of two universal preventive interventions in reducing children's externalizing behavior: A cluster randomized controlled trial. *Journal of Clinical Child & Adolescent Psychology*, 40, 677–692. doi:10.1080/15374416.2011.597084
- *Matos, M., Bauermeister, J. J., & Bernal, G. (2009). Parent-child interaction therapy for Puerto Rican preschool children with ADHD and behavior problems: A pilot efficacy study. *Family Process*, 48, 232–252. doi:10.1111/j.1545-5300.2009.01279.x
- *McGilloway, S., Ni Mhaille, G., Bywater, T., Furlong, M., Leckey, Y., Kelly, P., ... Donnelly, M. (2012). A parenting intervention

- for childhood behavioral problems: A randomized controlled trial in disadvantaged community-based settings. *Journal of Consulting and Clinical Psychology*, 80, 116–127. doi:10.1037/a0026304
- Mejia, A., Calam, R., & Sanders, M. R. (2014). Examining delivery preferences and cultural relevance of an evidence-based parenting program in a low-resource setting of Central America: Approaching parents as consumers. *Journal of Child and Family Studies*. Advance online publication. doi:10.1007/s10826-014-9911-x
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *Annals of Internal Medicine*, 151(4), 264–269. doi:10.7326/0003-4819-151-4-200908180-00135
- *Morpeth, L., Berry, V., Blower, S., Tobin, K., Taylor, R., Edwards, R. T., . . . Lehtonen, M. (2012). *The effectiveness of the Incredible Years (IY) pre-school parenting programme in Birmingham, UK: A randomised controlled trial*. Dartington, UK: Dartington Social Research Centre.
- New Zealand Ministry of Education. (2009). *Positive behaviour for learning*. Wellington, NZ: Author.
- OECD. (2012). *OECD family database*. Retrieved from <http://www.oecd.org/social/family/database>.
- *Ogden, T., & Hagen, K. A. (2008). Treatment effectiveness of parent management training in Norway: A randomized controlled trial of children with conduct problems. *Journal of Consulting and Clinical Psychology*, 76, 607–621. doi:10.1037/0022-006X.76.4.607
- Palinkas, L. A., Aarons, G. A., Chorpita, B. F., Hoagwood, K., Landsverk, J., & Weisz, J. R. (2009). Cultural exchange and the implementation of evidence-based practices: Two case studies. *Research on Social Work Practice*, 19, 602–612. doi:10.1177/1049731509335529
- Panagiotou, O. A., Contopoulos-Ioannidis, D. G., & Ioannidis, J. P. A. (2013). Comparative effect sizes in randomised trials from less developed and more developed countries: Meta-epidemiological assessment. *BMJ*, 346, f707. doi:10.1136/bmj.f707
- Pinderhughes, E. E., Hurley, S., & The Conduct Problems Prevention Research Group. (2008). Disentangling ethnic and contextual influences among parents raising youth in high-risk communities. *Applied Developmental Science*, 12, 211–219. doi:10.1080/10888690802388151
- Piquero, A. R., Farrington, D., Jennings, W. G., Tremblay, R., Piquero, A., & Welsh, B. (2008). Effects of early family/parent training programs on antisocial behavior and delinquency: A systematic review. *Campbell Systematic Reviews*, 4(11), 1–122. doi:10.4073/csr.2008.11
- *Posthumus, J. A., Raaijmakers, M. A. J., Maassen, G. H., van Engeland, H., & Matthys, W. (2012). Sustained effects of Incredible Years as a preventive intervention in preschool children with conduct problems. *Journal of Abnormal Child Psychology*, 40, 487–500. doi:10.1007/s10802-011-9580-9
- Reid, M. J., Webster-Stratton, C., & Beauchaine, T. P. (2001). Parent training in head start: A comparison of program response among African American, Asian American, Caucasian, and Hispanic mothers. *Prevention Science*, 2, 209–227. doi:10.1023/A:1013618309070
- Scott, S. (2010). Nationwide dissemination of effective parenting interventions: Building a parenting academy for England. In J. Weisz & A. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (2nd ed., 500–518). New York, NY: Guilford.
- Scott, S., Briskman, J., & O'Connor, T. G. (2014). Early prevention of antisocial personality: Long-term follow-up of two randomized controlled trials comparing indicated and selective approaches. *American Journal of Psychiatry*, 171, 649–657. doi:10.1176/appi.ajp.2014.13050697
- *Scott, S., Spender, Q., Doolan, M., Jacobs, B., Aspland, H., & Webster-Stratton, C. (2001). Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice commentary: Nipping conduct problems in the bud. *BMJ*, 323, 194. doi:10.1136/bmj.323.7306.194
- Scott, S., Sylva, K., Doolan, M., Price, J., Jacobs, B., Crook, C., & Landau, S. (2010). Randomised controlled trial of parenting for child antisocial behaviour targeting multiple risk factors: The SPOKES project. *Journal of Child Psychology and Psychiatry*, 51, 48–57. doi:10.1111/j.1469-7610.2009.02127.x
- *Sigmarsdóttir, M., Degarmo, D. S., Forgatch, M. S., & Guðmundsdóttir, E. V. (2013). Treatment effectiveness of PMTO for children's behavior problems in Iceland: Assessing parenting practices in a randomized controlled trial. *Scandinavian Journal of Psychology*, 54, 468–476. doi:10.1111/sjop.12078
- Sundell, K., Hansson, K., Lofholm, C. A., Olsson, T., Gustle, L.-H., & Kadesjo, C. (2008). The transportability of multisystemic therapy to Sweden: Short-term results from a randomized trial of conduct-disordered youths. *Journal of Family Psychology*, 22, 550–560. doi:10.1037/a0012790
- Sussman, S., Unger, J. B., & Palinkas, L. A. (2008). Country prototypes and translation of health programs. *Evaluation & the Health Professions*, 31, 110–123. doi:10.1177/0163278708315918
- *Taylor, T. K., Schmidt, F., Pepler, D., & Hodgins, H. (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
- United Nations Office on Drugs and Crime. (2009). *Guide to implementing family skills training programmes for drug abuse prevention*. New York, NY: United Nations.
- Wang, S., Moss, J. R., & Hiller, J. E. (2006). Applicability and transferability of interventions in evidence-based public health. *Health Promotion International*, 21, 76–83. doi:10.1093/heapro/dai025
- Webster-Stratton, C. (2009). Affirming diversity: Multi-cultural collaboration to deliver the Incredible Years parent programs. *International Journal of Child Health and Human Development*, 2, 17–32.
- Wessels, I., Mikton, C., Ward, C., Kilbane, T., Alves, R., Campello, G., . . . Madrid, B. (2013). *Preventing violence: evaluating outcomes of parenting programmes*. Geneva, Switzerland: WHO.
- Wilson, B. D. M., & Miller, R. L. (2003). Examining strategies for culturally grounded HIV prevention: A review. *AIDS Education and Prevention*, 15, 184–202. doi:10.1521/aeap.15.3.184.23838
- World Bank. (2012). *Databank—countries (database)*. Retrieved from <http://data.worldbank.org/country/>
- World Health Organization. (2010). *Violence prevention: the evidence*. Geneva, Switzerland: WHO.
- World Values Survey. (2008). *Values change the world: World values survey*. Retrieved from http://www.worldvaluessurvey.org/wvs/articles/folder_published/article_base_110