

PARENTING SUPPORT

Parenting Support for Diverse Populations

Dana Sheshko

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Doctorate in Philosophy degree in Clinical Psychology

School of Psychology
Faculty of Social Sciences
University of Ottawa

© Dana Sheshko, Ottawa, Canada, 2021

Abstract

The overarching goal of this dissertation is to inform evidence-based supports for diverse parents. Past meta-analyses (e.g., Dimitrova, Chasiotis, & van de Vijver, 2016; Porter & Haslam, 2005) have examined adjustment in migrant children, adolescents, and adults. To date, reviews have often examined only immigrants or refugees, with few quantitative comparisons between groups. Further, few include information on parenting or parental adjustment (e.g., van Os, Kalverboer, Zijlstra, Post, & Knorth, 2016). To fill this gap, I conducted a systematic review and meta-analysis of studies that used psychometrically strong measures of adjustment and parenting with newcomer immigrant and refugee families to update the literature and to provide a quantitative comparison between migrant groups. Database searches yielded 18,139 abstracts for screening; 4,626 full text articles were reviewed; 31 independent samples 25 of which have not been included in previous reviews met inclusion criteria. Results highlight important differences between migrant groups and informants: parent reports of youth adjustment generally reported greater problems among both immigrant and refugee youth compared to normative samples, whereas youth self-reports varied by migrant group (with refugee youth reporting greater problems than immigrant youth) and adjustment construct. In my second study, I developed and evaluated a self-report measure of adherence to an evidence-based parenting program that is used with diverse families. This tool, the Practitioner Session Reflection Tool (PSRT; Sheshko, Lee, & Gagné, 2015), was designed to support practitioners' adherence to both the session content and to the primary process of self-regulation theorized to underpin the Triple P – Positive Parenting Program (Sanders, 1999, 2012). Fifty-two practitioners from 12 community agencies submitted 361 self-reports; a subset of 23 practitioners submitted 63 audio-recordings to permit additional ratings by a coding team providing a multi-informant and

multimethod evaluation of adherence. Results provided preliminary evidence of good levels of reliability and validity for scores on the adherence measures. Consistent with previous research there were low correlations between informants: practitioners reported both higher content (84.6% compared to 49.9%) and process (80.0 – 95.0% compared to 30 – 55%) adherence than that rated by coders. Possible explanations for this discrepancy are discussed. Taken together, the results of these two studies highlight both the importance of offering supports that can be delivered flexibly to suit the needs of diverse families, and the need to support practitioners' self-reflection on adherence to program content and underlying processes in delivering those supports.

Acknowledgements

I would like to express my deep appreciation for the following individuals who have influenced my development throughout my education. First and foremost, I thank my thesis supervisor, Dr. Catherine M. Lee. Thank you for our many conversations, both about research and about life, that supported me in navigating my graduate training. Your guidance and kindness throughout these years has helped to shape me as both a researcher and a clinician and continues to inspire me in working with diverse families.

I would like to thank my thesis committee for their support: your guidance and expertise were invaluable throughout my training. Thank you to my research collaborators from Quebec City, Montreal, and Australia for your support and encouragement. To the families, practitioners, and agencies that generously contributed to these research projects, and to the research assistants that helped it to continually move forward, I express my sincerest gratitude. I would also like to extend my gratitude to my first research supervisor as an undergraduate student, Dr. John R. Z. Abela, who inspired my early engagement in research and encouraged me to start down the path of graduate training in clinical psychology.

I extend my thanks to the University of Ottawa, the School of Psychology, and to the Clinical Psychology program for your commitment to rigorous and engaging training. I thank the many clinical supervisors who helped to shape me as a clinician as a practicum student in Ottawa and as a resident in Winnipeg: Thank you for creating environments that supported my self-reflection, challenged me, and promoted my personal and professional growth.

And finally, to my family and friends: You bring so much joy to my life and I am grateful for each of you. Thank you for your encouragement, support, and helping me to maintain balance in my life. You have helped to make this journey unforgettable.

Statement of Co-Authorship

This dissertation consists of two manuscripts, both prepared in collaboration with my dissertation supervisor, Dr. Catherine M. Lee. I was primary author and Dr. Lee was the secondary author for both manuscripts, entitled “Adjustment and Parenting Among Immigrant and Refugee Families: A Systematic Review and Meta-Analysis” and “Multimethod Adherence Measurement in an Evidence-Based Parenting Program.” I was responsible for research ethics board communication, data collection for components of my studies, managing research assistants, observational coding, data coding and analyses, and manuscript preparation and submission. Dr. Lee was actively involved in research ethics board communication, training research assistants, coding data, and communications with collaborators. Dr. Lee provided guidance at all steps for these projects, including refining research questions and study methods, statistical analyses, and manuscript preparation.

Table of Contents

Abstract	ii
Acknowledgements	iv
Statement of Co-Authorship	v
List of Tables and Figures.....	xii
General Introduction	1
Canadian Context.....	3
Migration and Family Diversity.....	7
Adjustment Among Immigrant and Refugee Children	8
Adjustment Among Immigrant and Refugee Parents	12
Use of Measures with Evidence of Psychometric Adequacy	15
Incorporating Multiple Informants	16
Limitations of Past Research	18
Evidence-Based Interventions with Diverse Families	19
Evidence-Based Parenting Support.....	22
Diversity in Parents’ Preferences and a Public Health Approach to Parenting Supports	25
Triple P – Positive Parenting Program.....	27
Change Process in Triple P	28
Flexible Delivery of Triple P	29
Measurement of Adherence in Evidence-Based Parenting Programs	31

Assessment of Content Adherence	33
Assessment of Process Adherence.....	34
Limitations of Past Research	36
Dissertation Objectives	37
Table 1	38
Table 2	42
Study 1	43
Abstract	44
Adjustment and Parenting Among Immigrant and Refugee Families	45
Child and Adolescent Adjustment	45
Adult Adjustment.....	49
Parenting	51
Examining the Perspectives of Different Raters of Adjustment and Parenting.....	52
Measures of Adjustment and Parenting with Evidence of Psychometric Adequacy	53
Limitations of Past Research	55
Goals of the Meta-Analysis	56
Method	57
Meta-Analysis Overview	57
Operationalization of Terms	57
Inclusion and Exclusion Criteria.....	58

Selection of Studies.....	59
Coding the Studies	61
Development of Coding Protocols.....	61
Methodological Quality	65
Data Analysis Plan.....	66
Unit of Analysis	67
Meta-Analytic Model.....	68
Results.....	69
Demographic Data	69
Adjustment Among Immigrant and Refugee Youth.....	70
Adjustment Among Immigrant and Refugee Parents	73
Parenting in Immigrant and Refugee Families	74
Discussion.....	75
Limitations, Strengths and Future Research	85
References.....	91
Table 1	112
Table 2	118
Table 3	120
Table 4	123
Table 5	125

Table 6	126
Table 7	128
Table 8	130
Table 9	131
Figure 1	132
Figure 2	133
Figure 3	134
Figure 4	135
Figure 5	136
Study 2	137
Abstract	138
Multimethod Adherence Measurement in an Evidence-Based Parenting Program.....	139
Measurement of Adherence in Evidence-Based Programs.....	140
Flexible Delivery and Adherence in Triple P	141
Multiple Assessment Perspectives	143
Method	144
Participant Characteristics	144
Sampling Procedures	145
Measures	146
Results.....	148

Feasibility of PSRT in Community Implementation	148
Reliability of Adherence Measures.....	149
Convergent Validity of the Coder Session Review Tool.....	149
Multiple Informants of Adherence.....	149
Self-Report Measurement of Adherence	150
Coder Ratings of Adherence.....	151
Discussion	152
Clinical Implications	154
Future Research	155
Conclusions.....	157
References	158
Table 1	166
Table 2	167
Integrated Discussion.....	168
Findings: Adjustment and Parenting in Immigrant and Refugee Families.....	169
Adjustment among Immigrant and Refugee Children and Adolescents.....	169
Adjustment among Immigrant and Refugee Parents	174
Parenting among Immigrant and Refugee Families.....	175
Limitations	178
Implications for Research and Practice.....	178

Future Research	183
The Public Health Approach: Working with Diverse Families	185
Findings: Adherence	186
Adherence Measurement in Parenting Programs.....	186
Limitations	190
Implications for Research and Practice.....	192
Future Research	194
Conclusion	195
References	197
Appendix A: Sample Triple P Session Checklist	223
Appendix B: Search Terms	226
Appendix C: Meta-analysis Inclusion Protocol	229
Appendix D: Meta-analysis Data Extraction Protocol.....	231
Appendix E: Meta-analysis Measures List	239
Appendix F: Methodological Quality Coding Criteria	248
Appendix G: Sample Practitioner Session Reflection Tool.....	251
Appendix H: Practitioner and Coder Ratings of Adherence.....	254
Appendix I: Power Analysis	256

List of Tables and Figures

General Introduction

Table 1: Summary of Reviews of Adjustment Among Immigrant and Refugee

Children and Adolescents.....	38
-------------------------------	----

Table 2: Summary of Reviews of Adjustment Among Immigrant and Refugee

Adults.....	42
-------------	----

Study 1

Table 1: Literature Reviews of Child Adjustment.....	112
--	-----

Table 2: Literature Reviews of Adult Adjustment.....	118
--	-----

Table 3: Characteristics of Child Adjustment Studies.....	120
---	-----

Table 4: Characteristics of Parental Adjustment Studies.....	123
--	-----

Table 5: Characteristics of Parenting Studies.....	125
--	-----

Table 6: Overall Effect Sizes of Child Adjustment, Parental Adjustment, and

Parenting across Informants.....	126
----------------------------------	-----

Table 7: Effect Sizes of Adjustment in Immigrant, Refugee, and Combined Samples

of Children/Adolescents.....	128
------------------------------	-----

Table 8: Effect Sizes of Adjustment in Immigrant, Refugee, and Combined Samples

of Parents.....	130
-----------------	-----

Table 9: Effect Sizes of Parenting in Immigrant, Refugee, and Combined Samples of

Families.....	131
---------------	-----

Figure 1: Flowchart of Included and Excluded Studies.....	132
---	-----

Figure 2: Individual & Summary Effect Sizes: Child Adjustment (Youth Self-

Reports).....	133
---------------	-----

Figure 3: Individual & Summary Effect Sizes: Child Adjustment (Parent and Teacher Reports).....	134
Figure 4: Individual & Summary Effect Sizes: Parental Adjustment.....	135
Figure 5: Individual & Summary Effect Sizes: Parenting.....	136
Study 2	
Table 1: Practitioner characteristics ($N = 52$) who delivered Group Triple P.....	166
Table 2: Content adherence: Self-reported frequency and type of modifications, PSRT part one ($N = 361$).....	167

General Introduction

Canada is an ethnically diverse country with a high proportion of recent immigrants and refugees (Statistics Canada, 2016). Stressors associated with migration may increase children and parents' risk of developing adjustment problems following migration (e.g., Porter & Haslam, 2005). These in turn may contribute to challenges in parenting and family relationships (Barlow, Smailagic, Huband, Roloff, & Bennett, 2014). Reviews of this literature provide information about immigrant (e.g., Dimitrova, Chasiotis, & van de Vijver, 2016) and refugee groups (e.g., Blackmore et al., 2020), as well as on combined samples that do not provide direct comparisons between these groups (e.g., Curtis, Thompson, & Fairbrother, 2018). Little information is known about parenting within these families (van Os, Kalverboer, Zijlstra, Post, & Knorth, 2016). Having a clearer understanding of adjustment and parenting among newcomer families could inform the provision of parenting supports to diverse families. The provision of services to diverse families requires flexibility to meet families' needs (Huey, Tilley, Jones, & Smith, 2014; Richard et al., 2019) and minimize barriers to service access (e.g., Richard et al. 2019; Miller et al., 2020), as well as adherence to the evidence base of the program (Kendall, Gosch, Furr, & Sood, 2008). Evaluation of adherence among parenting programs has focused on adherence to program content with less attention paid to the underlying processes theorized to underpin them (Sanders & Kirby, 2015).

To provide the context for the thesis: first, I describe cultural and ethnic diversity within Canada's population. I then highlight migration on a global scale, reviewing adjustment and parenting among immigrant and refugee children, adolescents, and adults. In this context, I discuss the importance of using psychometrically sound measures of adjustment and parenting, and of incorporating reports from multiple raters to obtain a more comprehensive understanding

of these families' potential needs. I subsequently review evidence-based practice and present an overview of evidence-based parenting interventions that are helpful to diverse parents. Next I discuss diversity across Canadian parents' preferences in accessing supports and describe the public health framework for parenting support proposed by Sanders and colleagues (Sanders, 2008, 2010; Sanders & Prinz, 2018) that aims to increase programs' reach to families who may need them. I then describe the Triple P–Positive Parenting Program (Triple P). An essential component of this program is its core process of promoting self-regulation, an ingredient which is especially relevant in working with diverse families. Following this, I review the research on adherence assessment, which is the extent to which an intervention is delivered in accordance with the core elements of its evidence base (Forgatch, Patterson, & DeGarmo, 2005). I present available measures for assessing adherence across parenting programs, emphasizing adherence to both a program's content and to its underlying processes. Finally, I describe the specific objectives of the thesis.

In this thesis I examine evidence-based parenting supports for diverse parents by presenting data from two sources. The first is a manuscript describing a systematic review and meta-analysis of the literature on immigrant and refugee families examining child adjustment, parental adjustment, and parenting. This study was designed to synthesize research findings on adjustment and parenting in immigrant and refugee youth and their parents. The second source of data is from a tool I created to measure adherence to a parenting intervention that is offered to diverse families. The article describing measure development and preliminary data from its use in a randomized controlled trial conducted in Quebec was published in the peer-reviewed journal *Practice Innovations* (Sheshko, Lee, & Gagné, 2020). Together, these two manuscripts of which I am first author will contribute to knowledge about the provision of evidence-based services to

diverse parents. Two themes are prominent in the two studies that comprise this thesis: sound assessment and self-regulation. First, in assessing needs across diverse families, it is essential to use reliable and valid measures. Second, it is also crucial to evaluate program implementation with reliable and valid methods. Finally, in line with a public health approach in which parenting supports are offered to meet a range of needs, the promotion of self-regulation supports families in identifying their own needs, setting goals that are pertinent to them, and taking credit for the changes they make.

Canadian Context

The number of globally displaced persons has increased rapidly in the last decade, reaching over 20 million refugees in 2019, 38% of whom are under the age of 18 (United Nations High Commissioner for Refugees, 2020). Migration refers to the movement of people, either within a region or across countries, and includes the migration of refugees, displaced people, or immigrants (Perruchoud & Redpath-Cross, 2011). The United Nations High Commissioner for Refugees (2006) defined an economic migrant as an individual who migrates for economic reasons, seeking improved livelihood, whereas a refugee is defined as a person forced to flee for well-founded fears for safety.

According to the 2011 National Household Survey, immigrants and refugees make up 20.6% of the population in Canada (Statistics Canada, 2013). Canada is among the top refugee resettlement countries in the world (United Nations High Commissioner for Refugees, 2016). As migration will be a primary source of the growth of Canada's population (Statistics Canada, 2016), understanding individuals' and families' adjustment will provide important information to inform the provision of services these families may receive.

The New Canadian Children and Youth Study (NCCYS; Beiser, 2010; Beiser, Ogilvie,

Rummens, Armstrong, & Oxman-Martinez, 2005) was designed to examine mental health among immigrant and refugee children (aged four to six) and preadolescents' (aged 11 to 13) living in six major cities in Canada. The NCCYS was developed to address a shortcoming of previous Canadian research that did not include a sufficient sample of immigrant and refugee youth populations to examine their post-migratory adjustment (Beiser, 2010; Beiser et al., 2005). Beiser and colleagues (Beiser et al. 2005, 2010, 2011, 2012, 2014, 2015, 2016; George & Bassani, 2013, 2016, 2018; George, Bassani, & Armstrong, 2012; Hamilton, Marshall, Rummens, Fenta, & Simich, 2011; Oxman-Martinez & Choi, 2014; Oxman-Martinez et al., 2012) recruited over 4,000 families in which children were foreign-born or had at least one parent who had immigrated within the previous 10 years. Immigrant parents completed surveys reporting on their child, themselves, and their families; the 11 – 13-year-old cohort completed a self-report. The final sample included data collected between 2001 and 2006 on families identified as either refugees or immigrants from 16 ethnocultural communities (Beiser & Hou, 2016; Beiser et al., 2005, 2015).

The survey used questions from the National Longitudinal Study of Children and Youth (NLSCY; Statistics Canada, 1998) to address ethnic and religious backgrounds, parenting behaviours, family functioning, as well as child emotional and physical aggression problems. Analyses of the NCCYS data-set have been presented for subsamples, ranging from small samples of 64 preadolescent youth from Ethiopia (Beiser et al., 2012), to 2,031 children and preadolescents from Hong Kong, Mainland China and the Philippines (Beiser et al., 2010, 2011, 2014), and 2,074 preadolescents from all 16 ethnocultural groups (Beiser et al., 2015). Collectively, these reports highlight the contribution of demographic variables (such as age, gender, and family characteristics), migration-specific variables (such as ethnic background and

settlement region), and post-migration variables (such as exposure to perceived discrimination and school environment) to emotional problems and physical aggression among immigrant youth (e.g., Beiser et al., 2010, 2011, 2014) as described below.

Most of the published reports focused on approximately 2,000 children and preadolescents from Hong Kong, the Philippines, and Mainland China (Beiser, 2010; Beiser et al., 2010, 2011, 2014; George & Bassani, 2013; Hamilton et al., 2011; Oxman-Martinez & Choi, 2014; Oxman-Martinez et al., 2012). Five articles (Beiser & Hou, 2016; Beiser et al., 2012; George & Bassani, 2013, 2016, 2018; George et al., 2012) presented data from additional smaller subsamples from Afghanistan, El Salvador, Ethiopian, India, Iran, Serbia, Sri Lanka, and Vietnam. Among children and preadolescents from Hong Kong, the Philippines, and Mainland China subsamples ($n = 2,031$) parental depressive symptoms, family dysfunction, and country of origin were associated with higher child and preadolescent emotional problems and physical aggression (Beiser et al., 2010, 2011). Supportive parenting was significantly associated with fewer emotional problems among the child subsamples from Hong Kong, the Philippines, and Mainland China ($n = 2,031$); harsh parenting was significantly associated with greater mental health problems in both children and preadolescents (Beiser et al., 2014). It is important to note that only the supportive parenting measure demonstrated adequate internal consistency, whereas caution must be exercised in interpreting the results examining harsh parenting. Higher levels of perceived discrimination were associated with (i) greater emotional problems among preadolescents from Ethiopia ($n = 64$: Beiser et al., 2012), (ii) lower self-perceived social competence, and (iii) lower self-esteem among preadolescents from Hong Kong, the Philippines, and Mainland China ($n = 1,053$: Oxman-Martinez et al., 2012). Furthermore, preadolescent youth from all 16 ethnocultural groups ($n = 2,074$; Beiser et al., 2015) who reported higher levels

of social competence also reported fewer emotional problems. Beiser and Hou (2016) found that refugee preadolescents reported greater levels of emotional problems and aggressive behaviour, pre-migratory traumatic events, and discrimination than did immigrant youth who migrated from the same countries. Interpretation of data in a separate report on children and preadolescents from Afghanistan, Iran, and the Punjab region of India was limited by inadequate internal consistency in the measures of family discrimination and child behaviour ($n = 1,081$; George et al., 2012).

The NCCYS data indicate similar patterns among immigrants to those found in nonmigrant samples in that higher rates of parental depression, family dysfunction, harsh parenting, and perceived discrimination were associated with poorer outcomes and supportive parenting was associated with improved outcomes (e.g., Azar & Wolfe, 2006; Raby et al., 2015; Ryan & Ollendick, 2018; Slagt, Dubas, Dekovic, & van Aken, 2016). The presentation of the data to date, however, does not reflect the full diversity of the sample's 16 ethnocultural groups, and with the exception of the data presented by George et al. (2012) which included an Afghani subsample of refugees ($n = 181$) and Beiser and Hou (2016), reports to date have focused only on immigrant youth or did not differentiate between immigrant and refugee youth. Further, the total number of participating refugee (as compared to immigrant) families has not been reported. As the first wave of NCCYS data were collected between 2001 to 2006 and the majority of the participants arrived in Canada in the 1990s (Beiser et al., 2015), it is unclear whether the patterns that were found in that cohort would also apply to immigrants and refugees who have arrived in Canada in the last 20 years. The number of forcibly displaced people worldwide has nearly doubled in the last decade (United Nations High Commissioner for Refugees, 2020). Given Canada's status as one of the top receiving countries (United Nations High Commissioner for

Refugees, 2016), it is important to continue to integrate results of studies examining individuals' and families' adjustment to inform the provision of services these families may receive.

Migration and Family Diversity

Researchers have examined the adjustment of immigrants who freely pursue international migration and that of refugees, including asylum seekers, detainees, and (externally) displaced persons, who have experienced forced international migration (European Migration Network, 2018; Perruchoud & Redpath-Cross, 2011). It is important to consider how long migrants have been in their new country. The “immigrant paradox” or the “healthy immigrant effect” refers to a finding of better physical and mental health among immigrants, compared to their native-born counterparts, upon their arrival (e.g., Cunningham, Ruben, & Narayan, 2008; Stafford, Newbold, & Ross, 2011; Vang, Sigouin, Flenon, & Gagnon, 2017). This health advantage, however, appears to decline over time (De Maio & Kemp, 2010; Newbold, 2005; Ng, Wilkins, Gendron, & Berthelot, 2005; Vang, Sigouin, Flenon, & Gagnon, 2015). The initial advantage among immigrants, at least in a Canadian context, may be partly explained by the strict health screening immigrants undergo as part of the selection process to enter Canada (Vang et al., 2015). The subsequent decline in health remains poorly understood (Vang et al., 2015), but is hypothesized to be explained by a host of factors including the development of unhealthy behaviours, lower rates of access to health care services, and increasing willingness to report symptoms (Newbold, 2005; Ng et al., 2005; Vang et al., 2017).

Current understanding of the immigrant paradox is largely based on research with immigrant adults. Far less is known about this phenomenon among child and adolescent immigrants, and among refugees (Vang et al., 2017). Research in Canada has found that immigrant youth report better overall mental health than do their native-born counterparts,

although some subsamples report more severe psychosocial distress (Vang et al., 2017). Vang and colleagues (2015, 2017) highlighted that the data examining health are rarely presented separately for refugee and nonrefugee migrants. When these data are available, however, refugees report poorer overall health than do economic migrants or their Canadian-born peers.

Other researchers have observed that rather than an “immigrant paradox” the data reflect “migration morbidity,” in which migrants display increased vulnerability to adjustment problems post-migration. A meta-analysis of studies conducted on immigrant youth in Europe found overall adjustment among immigrant youth was slightly poorer than that of their native-born counterparts with regards to internalizing and externalizing problems (Dimitrova et al., 2016). Research examining adjustment from this perspective considers challenges to post-migratory adjustment among children, adolescents, and adults, and potential contributors or moderators in an attempt to understand variability across migrant groups (e.g., Vang et al., 2017).

Adjustment Among Immigrant and Refugee Children

The adjustment experiences of migrant youth have been synthesized in systematic reviews (e.g., Bronstein & Montgomery, 2011; Curtis et al., 2018; Fazel, Reed, Panter-Brick, & Stein, 2012; Kien et al., 2019; Marley & Mauki, 2019; Reed, Fazel, Jones, Panter-Brick, & Stein, 2012; Stevens & Vollebergh, 2008; Tam, Houlihan, & Melendez-Torres, 2017; van Os et al., 2016;) and meta-analyses (e.g., Blackmore et al., 2020; Dimitrova et al., 2016; Fazel, Wheeler, & Danesh, 2005; Porter & Haslam, 2005). Many systematic reviews and meta-analyses focused exclusively on either refugee (Blackmore et al., 2020; Bronstein & Montgomery, 2011; Fazel et al., 2005; Fazel et al., 2012; Kien et al., 2019; Marley & Mauki, 2019; Porter & Haslam, 2005; Reed et al., 2012; Tam et al., 2017; van Os et al., 2016) or immigrant youth (Dimitrova et al., 2016; Stevens & Vollebergh, 2008), with far fewer including both (Curtis et al., 2018). Across

these reviews, distress has been examined at the level of reported symptoms (e.g., Bronstein & Montgomery, 2011) and diagnoses based on semi-structured interviews (e.g., Blackmore et al., 2020). Studies have included general measures of overall distress and of externalizing and/or internalizing behaviours, and specific measures of symptoms of anxiety, depression, and trauma/posttraumatic stress disorder (PTSD); symptoms of attention-deficit/hyperactivity disorder (ADHD) and oppositional-defiant disorder (ODD) are less frequently examined (e.g., Blackmore et al., 2020). First, I first present meta-analyses and systematic reviews on refugee samples, then immigrant, and then for combined samples; methodological considerations and issues are summarized in [Table 1](#).

Refugee rates of mental health problems are generally found to be higher than for nonmigrant peers. An early large-scale meta-analysis conducted by Porter and Haslam (2005) found that refugees' mental health was poorer than that of nonrefugee comparison groups (weighted mean effect size $d = 0.41$). Results were not presented specifically for children or adolescents, other than that they appeared to exhibit better adjustment than adults; however, there were no statistical comparisons of the effect sizes calculated for child/adolescent samples compared to adult samples, so the magnitude of this difference is unknown. The first meta-analysis to present results uniquely for children and adolescents by Fazel and colleagues (2005) confirmed that refugee children and adolescents experienced greater adjustment problems: the overall prevalence rate of PTSD was 11% among refugee youth, which is higher than prevalence rates among nonmigrants (e.g., Merikangas et al., 2010). Expanding beyond PTSD, Bronstein and Montgomery (2011) found prevalence rates of symptoms of PTSD, symptoms of depression, and general psychological distress were also higher among refugee youth than nonmigrant peers. These reviews included studies with wide ranges of how long refugee participants had been in

the host country. Addressing this issue, van Os and colleagues (2016) conducted a systematic review that focused on recently arrived refugee and asylum-seeking children. They reported that the most common mental health problems reported upon arrival included symptoms of PTSD, depression, and anxiety disorders. The persistence of PTSD among refugee youth was then explored by Tam and colleagues (2017) in a systematic review of longitudinal studies among forcibly displaced children, who reported that high PTSD prevalence rates remained relatively stable over time. Building on past reviews, Kien and colleagues' (2019) systematic review found that up to one-third of refugee youth reported symptoms consistent with diagnoses of depression or anxiety, and up to half reported PTSD. In addition to increased levels of symptoms, recent research indicates higher prevalence rates of adjustment problems at the level of diagnosis among refugee and asylum-seeking youth across a variety of disorders, including PTSD, anxiety and major depressive disorders, and ADHD (but not ODD; Blackmore et al., 2020) compared to worldwide prevalence rates (Polanczyk, Salum, Sugaya, Caye, & Rohde, 2015).

Unlike reviews centered on refugee youth, reviews specific to immigrant youth have yielded mixed findings. Stevens and Vollebergh (2008) were unable to draw firm conclusions about rates of adjustment problems due to the high variability across youth self-reports, parent, and teacher reports and inconsistency in how terms, such as "migrant," were defined. Dimitrova and colleagues (2016) conducted a meta-analysis of externalizing and internalizing problems in first-generation immigrant youth. Overall, they reported that first-generation immigrant youth exhibited somewhat poorer adjustment in terms of internalizing ($d = -0.15$; lower scores reflect worse outcomes) and externalizing ($d = -0.13$) problems than comparison groups. Finally, Curtis and colleagues' (2018) systematic review of immigrant and refugee youth found that symptoms of depression and anxiety were elevated among immigrant youth and symptoms of PTSD were

elevated among refugee youth (compared to nonmigrant peers); however, there was no direct comparisons of the experiences of immigrant compared to refugee youth.

Across these studies, authors reported a high degree of variability in the reported rates of adjustment problems and explored factors that may account for this heterogeneity. Meta-analytic moderator analyses identified time and location as significant variables. Refugee youth who migrated within two years reported greater distress than did those who had been displaced for longer (Blackmore et al., 2020). Levels of distress also varied by geographical host region. Larger effect size for externalizing problems were found for youth in Northern than in Southern Europe; for internalizing symptoms the effect was larger for youth in Western Europe than for youth in Northern or Southern Europe (Dimitrova et al., 2016); differences in effect sizes across these regions were statistically significant. Reviews consistently found that children who experienced stressful life events of greater frequency, intensity, and duration reported higher symptoms (van Os et al., 2016). Further, children who directly experienced violence (prior to and following migration), had poor financial support or unstable post-migration housing, experienced discrimination, parents with poorer mental health (Fazel et al., 2012), resettled in a refugee camp, or were internally displaced (Reed et al., 2012) reported greater distress. Finally, some evidence from a systematic review of refugee youth suggested that – mirroring the epidemiological patterns found in nonmigrants – girls were at a heightened risk for internalizing problems, whereas boys exhibited increased risk for externalizing problems (Reed et al., 2012).

In addition to reviews that focused on adjustment problems, reviews conducted by Fazel, Reed, and colleagues (2012; 2012) and Marley and Mauki (2019) identified protective factors: youth who reported stronger social support (family cohesion, support from peers, positive school experiences), more stable resettlement conditions in the host country (Fazel et al., 2012), and

those who were repatriated to their country of origin (Reed et al., 2012) appeared to be at a decreased risk for mental health problems. Marley and Mauki (2019) found that despite the hardship they experienced, refugee youth reported more positive mental health outcomes when they also had greater levels of social support (among peers, family and from ethnically similar community members), higher self-esteem, higher parental education, greater opportunities to maintain their cultural identity, increased school and community connectedness, and access to innovative social care services.

Taken together, the evidence suggests that problems with adjustment are more consistently reported among refugee youth, whereas results among immigrant youth are mixed. Drawing firm conclusions comparing groups, however, is limited by few reviews incorporating both immigrant and refugee youth to better understand potential moderators of adjustment, and the degree to which immigrant and refugee youth may be differentially impacted by pre- and post-migratory experiences.

Adjustment Among Immigrant and Refugee Parents

Systematic reviews and meta-analyses on adults who have migrated are summarized in [Table 2](#). These studies have synthesized the experiences of adult refugee (Fazel et al., 2005; Porter & Haslam, 2005), immigrant (Falah-Hassani, Shiri, Vigod, & Dennis, 2015), or combined immigrant and refugee samples (Anderson, Hatch, Comacchio, & Howard, 2017; Fellmeth, Fazel, & Plugge, 2017). Apart from reviews examining perinatal adjustment among immigrant and refugee mothers (Anderson et al., 2017; Falah-Hassani et al., 2015; Fellmeth et al., 2017), the parental status of participating adults was unspecified, and none of these reviews included information specific to migrant fathers' experiences. Evaluation of PTSD, depression (including in the perinatal period), and anxiety were frequently represented.

Among adult refugees, both Fazel et al. (2005) and Porter and Haslam (2005) found that refugees reported high rates of adjustment problems. Specifically, Fazel and colleagues (2005) reported overall prevalence rates for PTSD of 9%, 5% for major depression, and 4% for generalized anxiety disorder (GAD). Further, Porter and Haslam (2005) reported that refugees overall disclosed poorer adjustment than nonmigrant comparison groups ($d = 0.41$).

Among immigrant mothers, Falah-Hassani and colleagues' (2015) meta-analysis of postpartum depression indicated that immigrant mothers were approximately one-and-a-half to two times more likely to report symptoms of postpartum depression than were nonmigrant comparison groups. Anderson and colleagues' (2017) meta-analysis of immigrant and refugee mothers in the perinatal period found that samples in Canada ($k = 9$) displayed increased risk for antenatal and postnatal depression whereas those in the United States ($k = 5$) and Australia ($k = 3$) did not. In exploring these differences, the authors proposed that variability across nonmigrant comparison groups used in each country may account for this difference. Across the American studies, the comparison groups may have been more comparable to migrant mothers: both groups were often from ethnic minority groups, reported comparably low levels of economic status, and were potentially exposed to comparable levels of stress associated with being members of minority groups. Alternatively, immigrant mothers in the Canadian studies tended to be compared to nonmigrants from majority ethnic groups, who may have been exposed to fewer of the challenges reported by migrant mothers. These variables may have augmented the differences between the groups, yielding greater differences in depressive symptoms. In another examination of immigrants and refugees, Fellmeth and colleagues (2017) examined perinatal mental health among mothers from low- and middle-income countries. Due to few included studies assessing refugees, data were presented for both immigrant and refugee mothers

together and revealed higher rates of depression symptoms among migrants compared to nonmigrants.

As in the syntheses of child adjustment, researchers attempted to clarify the high levels of variability in the reported rates of adjustment problems among adult immigrants and refugees by examining potential moderating variables. In one review, moderator analyses did not explain the heterogeneity of effect sizes across the included samples (Anderson et al., 2017). Other meta-analyses identified a number of factors that appeared to have a protective or deleterious effect on adults' mental health: migrants who lived longer in the host country tended to report better outcomes than did more recent migrants (Falah-Hassani et al., 2015; Porter & Haslam, 2005), and women reported poorer outcomes than did men (Porter & Haslam, 2005). Further, in line with patterns among nonmigrants, women in the perinatal period who were low in social support (Falah-Hassani et al., 2015; Fellmeth et al., 2017), with poorer marital adjustment, perceived insufficient household income (Falah-Hassani et al., 2015), and with a previous history of depression (Fellmeth et al., 2017) were more likely to report symptoms of perinatal depression.

In Porter and Haslam's (2005) meta-analysis, region of origin emerged as a potential moderator, as refugees from Europe reported more negative outcomes than did refugees from Asia and the Middle East; however, authors proposed that these results may have been confounded by the comparison groups' regions of origin, as the European refugees may have been compared to nonrefugee groups that were more likely to come from comparatively peaceful regions, whereas the refugees from Asia and the Middle East may have been more likely to be compared to nonmigrants from regions with a history of civil unrest, and may have more closely resembled the refugee groups. Further, refugees reported poorer adjustment than did immigrants; however, when exposure to violence was accounted for, immigrants who had been

exposed to violence reported adjustment that was comparable to that of refugee samples (Porter & Haslam, 2005). Both Fazel et al. (2005) and Porter and Haslam (2005) found that studies with greater methodological rigour yielded lower estimates of adjustment problems than did lower quality studies.

Research on migrant adults indicates higher rates of adjustment problems among migrant adults than in nonmigrants. Finer-grained analyses comparing immigrants and refugee experiences are rare (a notable exception is Porter & Haslam, 2005). The parental status of adults in these studies is rarely reported, apart from studies of the perinatal period, and no reviews included perspectives from fathers. Given these gaps, it is important to understand the adjustment experiences of immigrant and refugee parents, including both mothers and fathers. Before exploring potential supports for these families, it is essential to review issues of measurement associated with evaluating adjustment and parenting.

Use of Measures with Evidence of Psychometric Adequacy

In measure development and evaluation of their reliability and validity, psychometric properties refer to a measure as used with a specific sample. It cannot be assumed that a tool that is psychometrically adequate in the sample for which it was originally developed is also adequate for the participants in other studies (Hunsley & Mash, 2018). A challenge in research with immigrant and refugee participants is that often a tool with demonstrated psychometric strength does not have evidence of psychometric adequacy with the specific ethnocultural group with which it is used. Translation and back-translation do not guarantee identical meaning across languages. Nor does use of a measure developed with one cultural group, when applied to a second group, guarantee that the second group will interpret the items' meaning in the same way, or describe symptoms or behaviour in the same manner. Both concerns may potentially reduce

the reliability and validity of scores on the measure. An alternative could be to create a new measure specific to a cultural group, however this may mean larger-scale studies to establish the reliability and validity of scores on the measure across various groups, which may not be consistently feasible. At a minimum, it is important for researchers to use measures with demonstrated psychometric adequacy. It is also essential to report how the measure performed in each sample that uses it, as it cannot be assumed that evidence of reliability and validity found in a normative sample will be replicated in other culturally diverse samples. Some systematic reviews or meta-analyses (e.g., Bronstein & Montgomery, 2011; Fellmeth et al., 2017; Falah-Hassani et al., 2015; Kien et al., 2019; Stevens & Vollebergh, 2008) stressed the necessity of using instruments with evidence of reliability and validity to measure distress. However, many other reviews included weak or untested measures (e.g., Fazel et al., 2012; Marley & Mauki, 2019; Reed et al., 2012) or did not describe the measures used in the included studies (Curtis et al., 2018; Dimitrova et al., 2016). Further, systematic reviews and meta-analyses often did not report on evidence of the psychometric adequacy of the measures used in the included studies with immigrant and/or refugee participants, limiting understanding of the reliability and/or validity adequacy of these tools when used with these culturally diverse samples. To increase confidence in meta-analytic results an important inclusion criterion should be that studies employ measures with evidence of reliability and validity. Further, the internal consistency of the measure for the specific migrant sample should be reported to evaluate the results of the study more comprehensively. In line with this, the meta-analysis included in this thesis as manuscript one will select studies that used a psychometrically strong measure of adjustment or parenting, and data (when available) of the measure's reliability with migrant samples will be extracted.

Incorporating Multiple Informants

In assessment of child behaviour and services, it is generally acknowledged that incorporating multiple perspectives is advantageous, as different informants may provide unique information (Hunsley & Mash, 2018). Achenbach, McConaughy, and Howell's (1987) seminal meta-analysis of agreement across raters in multi-informant, multimethod assessments of child and adolescent psychopathology revealed that correspondence across raters was consistently low to moderate (average Pearson $r = .28$; Achenbach et al., 1987). A robust literature has replicated this finding many times confirming that agreement between raters is imperfect for rating children's behaviour (De Los Reyes & Kazdin, 2005). This trend has also been identified between adolescents and parents in their evaluations of parenting behaviour (e.g., Hou et al., 2020). Despite these well-replicated findings, there is a continuing tendency to view converging reports as a measure of reporting accuracy, treating discrepancies as error or an artifact of measurement (De Los Reyes, Thomas, Goodman, & Kundey, 2013), whereas these discrepancies may reflect real differences. Previous meta-analyses and systematic review of child adjustment often focused only on child/adolescent self-report of adjustment (Blackmore et al., 2020; Bronstein & Montgomery, 2011), which precluded evaluations of younger children who could not reliably self-report, and the perspectives of parents and teachers. Given the known variability across raters, it is important to access information from multiple perspectives. With the notable exception of Stevens and Vollebergh (2008), systematic reviews and meta-analyses of child adjustment rarely distinguish between different informants of child adjustment (e.g., Dimitrova et al., 2016; Kien et al., 2019; Marley & Mauki, 2019). Understanding the well-established differences across informants, it is also important to examine these reports separately. In line with this, I elected to examine informants of adjustment and parenting separately in manuscript one.

Limitations of Past Research

The systematic reviews and meta-analyses examining mental health outcomes among immigrant and refugee children, adolescents, and adults have highlighted many factors that may impact mental health outcomes, however data differentiating between migrant groups is limited and there are multiple concerns with measurement. Data from the large Canadian study conducted by Beiser and colleagues (e.g., Beiser et al., 2010) have mainly been presented for mental health outcomes among immigrant youth, but little data from refugee youth have been published, and some results must be interpreted with caution due to inadequate internal consistency of the measures. Although many researchers have encouraged the use of psychometrically adequate measures of distress (e.g., Falah-Hassani et al., 2015; Stevens & Vollebergh, 2008), untested measures or those with inadequate psychometric evidence continue to be employed, measures are not described, or the reliability or validity of the scores on the measures used in the included studies are not reported (e.g., Curtis et al., 2018; Dimitrova et al., 2016; Fazel et al., 2012). Further, Bronstein and Montgomery (2011), Curtis et al. (2018), Fazel, Reed, and colleagues (2012; 2012), Stevens and Vollebergh (2008), and van Os et al. (2016) all reported difficulty using meta-analysis to accommodate methodological and instrument variability. In addition, some reviews of child adjustment were limited to self-report (Bronstein & Montgomery, 2011; Curtis et al., 2018), others described challenges accounting for variability across informants (Stevens & Vollebergh, 2008), or did not distinguish between different informants of child adjustment (Kien et al., 2019; Marley & Mauki, 2019). Across this body of research, these challenges have yielded mainly systematic reviews of the literature among immigrant or refugee children and adolescents to date rather than quantitative comparisons that accommodate both. Among reviews about migrant parents, many have focused on mothers'

adjustment experiences, particularly in the perinatal period, and little is available that is specific to migrant fathers' adjustment. The use of meta-analysis in the current study will permit quantitative comparisons among diverse families and aims to extend our understanding of adjustment among children and parents in immigrant and refugee families.

The literature reviewed above present an overall picture of immigrant and refugee adults and children at increased risk of mental health difficulty following migration (e.g., Porter & Haslam, 2005). Individuals who have been forced to flee, exposed to violence, and/or traumatic experiences may be particularly at risk for poorer outcomes regardless of immigrant or refugee status (Porter & Haslam, 2005). Having examined adjustment difficulties, I now turn to reviewing supports designed to help families.

Evidence-Based Interventions with Diverse Families

Evidence-based practice (EBP) is defined as the integration of the best available research evidence with clinical judgement, accounting for the needs and characteristics of the individuals seeking treatment (American Psychological Association, 2006, 2008; Canadian Psychological Association Task Force on Evidence-Based Practice of Psychological Treatments, 2012). EBP requires regular syntheses of the research literature to inform clinical practice guidelines (American Psychological Association, 2006, 2008; Canadian Psychological Association Task Force on Evidence-Based Practice of Psychological Treatments, 2012). Replicated research findings are essential to inform prevention and intervention practices. Some critics, however, have questioned the use of evidence-based interventions with culturally diverse samples (e.g., Sue, Zane, Nagayama Hall, & Berger, 2009), a concern initially fueled by underrepresentation of ethnic minority participants in earlier treatment studies (Chambless et al., 1996). This spurred researchers to examine both how the processes involved in effective treatments contribute to

positive outcomes among clients (Norcross, 2002) and to investigate program effectiveness specifically with culturally and ethnically diverse populations (Huey & Polo, 2008; Huey et al., 2014; Miranda et al., 2005). The APA Psychotherapy Division (Division 29) first developed a task force in 1999 with two objectives: to identify empirical support of effectiveness for elements within the therapeutic relationship and methods for adapting interventions to accommodate client characteristics and preferences (Norcross, 2002). Updated task forces followed, cosponsored by the divisions of Psychotherapy (Division 29) and Clinical Psychology (Division 12) in 2011 (Norcross & Wampold, 2011), and the Society for the Advancement of Psychotherapy (Division 29) and Counseling Psychology (Division 17) in 2018 (Norcross & Lambert, 2018; Norcross & Wampold, 2018). These updated results, reported in a series of meta-analyses, highlighted methods of tailoring psychotherapy to reflect clients' characteristics and preferences. In addition, a growing body of research indicates that evidence-based interventions are effective with ethnic minority youth and adults (Huey & Polo, 2008; Huey et al., 2014; Miranda et al., 2005): a review of 300 randomized controlled trials (RCTs) across a variety of problems conducted by Huey and colleagues (2014) found that, on average, evidence-based treatments (EBTs) were equally effective across ethnic groups. Nonetheless, available evidence also suggests that various culture-related barriers to intervention may limit accessibility among ethnic minorities, such as differences in help-seeking beliefs and behaviours (Abdullah & Brown, 2011). Given these barriers it is important to consider cultural factors when developing and disseminating services (Betancourt, Green, Carrillo, & Ananeh-Firempong, 2003; Turner, Sanders, Keown, & Shepherd, 2018).

Service providers must consider how to implement an intervention with fidelity to its evidence base, while integrating the client's preferences, needs, and characteristics in a culturally

appropriate manner. In a review of 10 meta-analyses examining the use of culturally-adapted interventions among youth and adults for a host of challenges (including emotional and behavioural problems, substance use, and family problems), Huey et al. (2014) found that when cultural adaptations of evidence-based interventions were compared to treatment-as-usual or other placebo conditions they yielded more positive outcomes, but there was little evidence that cultural adaptations augmented the effects of nonadapted evidence-based interventions.

Huey et al. (2014) therefore recommended that in serving an ethnic minority client, service providers should first select an evidence-based intervention. They outlined a variety of approaches to accommodate culture, ranging from generating culture-specific variants of an EBT program to individualizing the EBT to match an individual client's needs. In individualizing a program to better fit with a client's needs, Huey et al. (2014) recommended prioritizing ethnocultural factors only if they appear particularly relevant to the development of treatment goals or the clinical issues. This is in line with evidence that treatment outcomes may be enhanced through goal congruence between practitioners and clients, and the integration of metaphors or symbols that resonate with a client's world view (Haslam & Mejia, 2018; Smith, Rodriguez, & Bernal, 2011; Turner, Sanders, et al., 2018). Although the individualization approach affords flexibility to practitioners (without requiring the potentially expensive development of large-scale culture-specific variants), it leaves the practitioner wholly responsible for the program's adaptation, leading to concerns about adherence to the program's evidence base (Kumpfer, Alvarado, Smith, & Bellamy, 2002). It could therefore be helpful for practitioners to have access to explicit recommendations for how to approach adaptations for diverse clients (e.g., Mazzucchelli & Sanders, 2010), a topic to which I return in subsequent sections.

The link between parenting and child adjustment is well established (Barlow & Coren, 2018; Ryan & Ollendick, 2018) as is the link between parental adjustment and parenting (Barlow et al., 2014). Across these bodies of research, child and parental distress are associated with problematic parenting practices. Greater parental adjustment problems are linked to increased adjustment problems among children (van Santvoort et al., 2015), and challenging child behaviour may elicit harsher parenting behaviour (e.g., Slagt et al., 2016; Yan & Ansari, 2016; Zvara, Sheppard, & Cox, 2018). All together, this suggests that parenting behaviours and both parental and child adjustment influence one another and emphasize the central role of parenting in a child's development (Collins, Maccoby, Steinberg, Hetherington, & Bornstein, 2000; Grusec, 2011). Specifically, ineffective or harsh parenting, exposure to family conflict or violence, parental abuse or neglect, and parental psychopathology are among the identified risk factors for a child's development of mental disorder (Clément, Gagné, & Lee, 2014; Gershoff et al., 2018; Kazdin & De Los Reyes, 2008; Weitzman, Rosenthal, & Liu, 2011), whereas supportive or effective parenting and family harmony serve as protective factors (Barrett & Turner, 2004; Biglan, Flay, Embry, & Sandler, 2012; Lamb, 2012; Raby et al., 2015). Given the challenges that families may confront in the face of stress and the fundamental role of parenting in children's development, all parents could benefit from some level of support (Sanders & Prinz, 2018). Parenting programs that can be implemented flexibly and that respect parents' rights to identify their own goals and select strategies that could be helpful are important options that must be available to families if they are struggling.

Evidence-Based Parenting Support

Efficacious and effective parenting programs increase the frequency of positive parenting behaviours (e.g., Kaminski, Valle, Filene, & Boyle, 2008; Sanders, Kirby, Tellegen, & Day,

2014), and improve child behaviour (e.g., Barlow & Coren, 2018; Dretzke et al., 2009; Michelson, Davenport, Dretzke, Barlow, & Day, 2013; Serketich & Dumas, 1996), parental adjustment (Barlow et al., 2014; Sanders et al., 2014), and parental stress (Barlow & Coren, 2018). Evidence-based parenting supports (EBPS) integrate research findings to inform the development of programs aimed at strengthening child development by modifying their family environment (Sanders & Prinz, 2018). Parenting programs with the strongest evidence draw on social learning theory (SLT; Bandura, 1971; Patterson, 1982, 2005), which posits that problematic child behaviours develop and are maintained in coercive cycles. Coercive exchanges occur when an adult inadvertently reinforces a child's misbehaviour by giving in; for example, ending a tantrum by buying the item demanded by the child. This parental reaction unintentionally rewards the child's problematic behaviour, and the child, in turn, rewards the parent by ceasing the aversive misbehaviour. In this way, both children and their parents inadvertently develop maladaptive strategies for communicating needs and negative behaviour is inadvertently rewarded and reinforced (Patterson, 1982, 2005). SLT is supported by a large body of research on the transactional processes within parent-child relationships: direct observations of families in both the laboratory and the home context have shown that in families that are functioning well, children learn prosocial behaviours (such as compromise) in the face of conflict (e.g., Biglan, Flay, Embry, & Sandler, 2012). Children in families that are struggling, however, develop coercive strategies (such as aggression) to resolve conflict (e.g., Smith & Chamberlain, 2010; Stormont, 2002), contributing to the development of coercive cycles of interaction between parents and children.

Bandura (1971) theorized that learning can occur from both direct experience and from the observation of others' experiences. Subsequent research demonstrated that children learn

appropriate behaviours through the observation of important people in their lives (Collins et al., 2000; Taylor & Biglan, 1998). Bandura's work (1977, 1995) examining social information-processing laid the groundwork for understanding that parents' expectations of and explanations for their child's behaviour, as well as their parenting beliefs, act as powerful influences upon their parenting decisions and sense of parental self-efficacy (e.g., Barlow et al., 2014).

According to SLT, ineffective parenting strategies and problematic child behaviours are learned, thus modifying the social environment can contribute to more effective parenting and improvement in child behaviour problems (e.g., Barlow & Coren, 2018; Bodenmann, Cina, Ledermann, & Sanders, 2008).

Parenting programs based upon SLT are designed to promote positive interactions between parents and their children. These programs focus on increasing parents' awareness of the reciprocal interactions between their own and their child's behaviour, emphasizing positive attention in response to desired behaviour, reducing inadvertent rewards for unwanted behaviours, helping parents to set clear limits for undesired behaviour, reducing reliance on demands and coercive or harsh parenting strategies, and increasing parental consistency in response to their child's behaviour (Taylor & Biglan, 1998). Grounded in SLT and family interaction research, the parenting support programs with the strongest evidence base include the Incredible Years (IY) Program (Webster-Stratton, 1998), Parent-Management Training–Oregon Model (PTMO; Forgatch & Patterson, 2010), Parent-Child Interaction Therapy (PCIT; Fernandez & Eyberg, 2009), and the Triple P–Positive Parenting Program (Sanders, 1999, 2012). These programs target behavioural changes (such as parenting behaviours), cognitive processes (how parents explain children's problem behaviour), and developmental principles (having realistic expectations of a child's behaviour, or a child's ability to understand rules and

consequences across developmental stages). In these programs, parents are supported in becoming the agents of change in their family's dynamic and children's behaviour through techniques based on SLT (Bandura, 1971).

Meta-analyses of RCTs indicate parenting programs are helpful with respect to children with specific neurodevelopmental and disruptive behaviour disorders, such as attention-deficit hyperactivity disorder (ADHD; e.g., Pelham & Fabiano, 2008) and oppositional defiant disorder (ODD; e.g., Erford, Paul, Oncken, Kress, & Erford, 2014). Additional meta-analyses indicate the efficacy of parenting programs across combined samples of children with externalizing behaviour problems and/or disruptive behaviour disorders (Eyberg, Nelson, & Boggs, 2008; Maughan, Christianseri, Jensori, Olympia, & Clark, 2005) or conduct problems (Dretzke et al., 2009; Furlong et al., 2012). Furthermore, a benchmarking study found that parenting interventions can successfully and effectively be offered in community settings (Lee, Horvath, & Hunsley, 2013).

Although parenting programs were initially designed for families with children vulnerable to developing behavioural or emotional problems or children who had already developed symptoms (e.g., Carr, 2014; Eyberg et al., 2008; Kazdin, 2005), a growing number of studies have shown that parenting programs can be efficacious in supporting families of children with subclinical levels of disruptive behaviour (e.g., de Graaf, Speetjens, Smit, de Wolff, & Tavecchio, 2008a; de Graaf, Speetjens, Smit, de Wolff, & Tavecchio, 2008b; Nowak & Heinrichs, 2008; World Health Organization, 2009). Furthermore, research indicates that parenting programs are a cost-effective option in the prevention of mental health issues (Mihalopoulos, Vos, Pirkis, & Carter, 2011).

Diversity in Parents' Preferences and a Public Health Approach to Parenting Supports

To better understand Canadian parents' perceptions of and engagement in parenting supports, Lee and colleagues (2014) surveyed 2,340 Canadian parents about their preferences in accessing parenting supports. They found that between 27.1% of parents (of young children) to 33.2% (of school-aged children) had consulted a professional (such as a pediatrician, teacher, nurse, or psychologist) about their child's behaviour in the previous year. Of the 15% of parents who reported participating in a parenting intervention, only a quarter of them had attended a parenting program with an established evidence base. Parents were asked to rate how useful they perceived various methods of accessing services. Their preferences for service delivery were diverse: data suggested that parents perceived web and TV-based services as more useful than phone, radio, or newspaper options. Parents also rated individual services and brief seminars as likely to be useful; services accessed through the workplace were rated more frequently as likely to be helpful than services accessed through religious organizations. These results suggest that no "one size fits all" option was endorsed by the majority (Lee et al., 2014). A range of parenting programs must be flexibly delivered to fit the beliefs, values, and needs of a diverse population.

In line with offering a range of intervention options, a public health approach to parenting support is designed to reduce the prevalence of harmful parenting practices and promote positive parenting at a population level (Sanders, 2008, 2010; Sanders, Burke, Prinz, & Morawska, 2017; Sanders & Prinz, 2018). Based on the assumption that parenting is challenging, that many parents feel unprepared for the parental role, and that all can benefit from some level of support (Prinz & Sanders, 2007; Sanders, 2012), a public health approach is designed to promote parents' competence and confidence in handling current and future parenting challenges (Sanders, 2008). Although several parenting programs based on social learning theory have a

strong evidence base, Triple P is the only parenting program designed to operate within a public health framework. Triple P is designed to meet the needs of families across the spectrum of risk, in supporting the development of positive relationships between parents and children through the reduction of ineffective parenting strategies. Operating within a public health framework means that both targeted and universal interventions are offered so that they complement each other, meeting the needs of families requiring different intensities of intervention rendering it relevant and accessible to all parents in the community (Sanders, 2012; Smith, Brown, Feldgaier, & Lee, 2015) in a cost-effective manner (Foster, Prinz, Sanders, & Shapiro, 2008).

Triple P – Positive Parenting Program

Triple P (Sanders, 1999, 2012) promotes five core principles of positive parenting: (1) providing a safe and engaging environment, (2) a positive learning environment, (3) using assertive discipline, (4) having realistic expectations for a child's behaviour (given the child's developmental stage), and (5) parental self-care (Sanders, 1999). The program consists of five levels of services of increasing intensity so that parents receive the amount of support that they require. Levels one and two focus on providing general information about parenting, level three consists of brief individual consultations, and levels four and five consist of more traditional intervention services. Specifically, level one is a universal media campaign that is designed to acknowledge the challenges of parenting, promote positive parenting practices, and reduce stigma associated with accessing parenting support. Level two is comprised of seminars on parenting issues that are open to any interested parents in the community. Level three is a brief series of individual consultation sessions for parents focused on a specific, common parenting struggle. Level four, presented as either a group or an individual program, consists of an eight-week course. Level five targets vulnerable families that require further support and consists of

an enhanced intervention for families in which parenting is complicated by additional stressors, such as couple conflict or parental mental health issues. Levels three through five involve parents setting specific goals for change within their families and then systematically working to attain goals they have identified (Sanders, 1999, 2012).

Change Process in Triple P

The Triple P program promotes two key principles to guide practitioners in their interactions with parents. The first core principle is the construct of the *minimal sufficiency*. This refers to offering the minimal support needed to facilitate change, and in Triple P it reflects practitioners' use of the minimum intensity of support necessary to encourage parents' self-reflection on their goals, use of the strategies, and progress (Sanders, 1999, 2012; Sanders, Turner, & Metzler, 2019). This construct of minimal sufficiency occurs both in terms of the program level attended (Sanders, 1999, 2012) and within each session, as practitioners use the least prompts or guides needed in supporting parents.

Related to this, the second guiding principle is to enhance parents' engagement in the program and generalization of skills by promoting *self-regulation*. The self-regulation model proposed by Karoly (1993) suggests that responding to cues from others about one's behaviour enables the individual to monitor and regulate his or her own behaviour, thus promoting the capacity to problem-solve independently. Biglan and colleagues (2012) highlighted the importance of a nurturing environment that supports the development of self-regulatory and prosocial skills, and the ability to self-regulate one's behaviour has been linked to positive developmental outcomes (Moffitt et al., 2011). The capacity to self-regulate one's attention, thoughts, feelings, and behaviours emerges early in life and is associated with more positive outcomes; it includes skills in self-efficacy, self-sufficiency, goal setting and self-monitoring,

and personal agency, which then support independent problem-solving and it is central to Triple P (Sanders et al., 2019). Parenting programs that integrate self-regulation encourage parents to use cues from their child's behaviour to regulate their own parenting behaviour (Sanders, 1999; Sanders & Mazzucchelli, 2013; Sanders et al., 2019). The construct of self-regulation permeates every aspect of the Triple P program: practitioners encourage parents to identify their own goals, select the strategies they would like to use to resolve parenting problems, then implement and evaluate the strategies they have chosen. In addition, practitioners prompt parents to identify what they are doing well in using the strategies, as well as what they would like to do differently the next time they are confronted with a similar difficult parenting situation. The ultimate goal of the Triple P parenting program is to promote self-regulation among children (Sanders et al., 2019). To accomplish this, parents are coached in developing their own self-regulation skills to serve as a model more effectively to their children. Practitioners delivering the program, in turn, serve as models to parents by demonstrating self-regulation skills and interacting with parents in a way to promote parents' self-regulatory development.

Flexible Delivery of Triple P

In examining the usefulness of interventions among diverse parents, it is helpful to first examine programs that were developed in one country and later evaluated elsewhere. Research supports this transportability, indicating that parenting programs developed in one country are effective with culturally dissimilar families (Gardner, Montgomery, & Knerr, 2016). Triple P was developed in Australia, has been delivered in 27 countries, translated into over 20 languages (Sanders & Prinz, 2018), and is recognized internationally as an evidence-based family skills training program (United Nations Office on Drugs and Crime, 2010; World Health Organization, 2009). Triple P program materials are offered in different languages (e.g., Errázuriz, Cerfogli,

Moreno, & Soto, 2016; Guo, Morawska, & Sanders, 2016; Haslam & Mejia, 2018) and have been adapted in collaboration with Indigenous communities in Australia and New Zealand (Turner, Hodge, Forster, & McIllduff, 2018). Further, a public health approach necessitates that the program appeal to as many parents as possible in a given community: Triple P has been shown to be acceptable to indigenous parents in Australia (Turner, Richards, & Sanders, 2007), Canada (Houlding, Schmidt, Stern, Jamieson, & Borg, 2012), and New Zealand (Turner, Sanders, et al., 2018), and with parents in Chile (Errázuriz et al., 2016) and South Africa (Wessels & Ward, 2016). Triple P's efficacy and/or effectiveness have been supported in Australia (Sanders, 1999), Mainland China (Guo et al., 2016), Iran (Tehrani-Doost, Shahrivar, Mahmoudi Gharai, & Alaghband-Rad, 2009), Germany (Feldmann, Heinrichs, Hahlweg, & Bertram, 2007), Hong Kong (Leung, Sanders, Ip, & Lau, 2006), Indonesia (Sumargi, Sofronoff, & Morawska, 2015), Panama (Mejia, Calam, & Sanders, 2015), Switzerland (Bodenmann et al., 2008), United Kingdom (Ruane et al., 2019), and the United States (Prinz, Sanders, Shapiro, Whitaker, & Lutzker, 2009), as well as with Japanese families in Australia (Matsumoto, Sofronoff, & Sanders, 2010).

Program developers have worked to guide practitioners in balancing fidelity to a program with its flexible implementation. Eyberg (2005) described three possible ways to adjust delivery of Parent-Child Interaction therapy, not all of which are advised. Practitioners are encouraged to tailor the program by adjusting content or its delivery (e.g., using more concrete or abstract strategies to teach skills) to suit clients' needs. Structural changes to the program's delivery or major content changes, such as adaptations (or changes to a standard protocol, often for a new population) or modifications (when the treatment developer implements universal changes to an existing protocol), however, are only to be made by the developer.

To support practitioners in balancing program adherence and flexible implementation in Triple P, Mazzucchelli and Sanders (2010) offered guidelines for modifications they considered low- and high-risk; low risk adaptations are adherent to the program's evidence base and may boost program effectiveness, whereas high risk adaptations reflect deviations and may compromise its effectiveness. Low-risk content adaptations consist of changes made to better suit the client (such as providing relevant examples); low risk process adaptations could include changing the location of services or increasing the number of sessions to accommodate clients' needs. High-risk content adaptations could include omitting or adding strategies; high risk process adaptations could include omitting role plays. In working with culturally diverse families, the self-regulation framework encourages respect for every parent's right to raise their children according to their own values; parenting strategies are offered as tools for parents to evaluate, and parents are guided in their use to inform their selection for the best fit to meet their family's individual goals (Haslam & Mejia, 2018). This flexible program delivery, based on the practitioner's judgment in collaboration with parents, enables adaptation to the client's individual needs, characteristics, and preferences (Mazzucchelli & Ralph, 2019; Mazzucchelli & Sanders, 2010; Sanders & Mazzucchelli, 2018), supporting the use of Triple P with diverse populations (Haslam & Mejia, 2018). To date, there have been no studies of how practitioners adapt the program to implement it with flexibility and the extent to which practitioners support self-regulation among parents. Both these factors are essential considerations in the evaluation of program adherence to Triple P and require a tool that can support practitioners in implementing the parenting program with both flexibility and adherence.

Measurement of Adherence in Evidence-Based Parenting Programs

In evidence-based practice, it is important to monitor both the implementation of service

and the associated changes in client behaviour (Dozois et al., 2014). An important aspect of implementation monitoring is the assessment of treatment fidelity (Garbacz, Brown, Spee, Polo, & Budd, 2014; Schoenwald, 2011). Fidelity to the evidence base, also referred to as treatment fidelity, implementation fidelity, and treatment integrity, is a multifaceted construct that includes such elements as adherence, competence, and differentiation (Allen, Shelton, Emmons, & Linnan, 2017; Perepletchikova, Treat, & Kazdin, 2007). Adherence is the extent to which the program is delivered in accordance with the core elements of its evidence base (Forgatch et al., 2005), competence refers to the skill with which the practitioner delivers the program, and differentiation is the degree to which an intervention protocol can be consistently distinguished from other protocols (Schoenwald, Garland, Southam-Gerow, Chorpita, & Chapman, 2011). Each facet captures an individual feature that may have a unique impact on therapeutic change (Perepletchikova et al., 2007).

Meta-analytic research has found that programs implemented with high levels of adherence yield greater effect sizes than do those that deviate from their empirical base (Forgatch & DeGarmo, 2011). To accurately determine the efficacy of an intervention, it is essential to first ensure that the program is being implemented with adherence (McArthur, Riosa, & Preyde, 2012); when the degree of program adherence is unknown, efficacious service delivery may be compromised (Fixsen, Blase, Duda, Naoom, & Van Dyke, 2010). Experts recommend that practitioners deliver a program in a way that both retains the elements of its evidence base and adapts to client needs (Kendall et al., 2008). There are multiple ways to assess adherence, including routine self-assessment by practitioners and observational ratings by other people, and it is important to consider adherence to both the program's content and its underlying processes.

Examination of program adherence often yields discrepant results across raters. For

example, Herschell and colleagues (2019) examined clinician, caregiver and expert ratings of adherence in a family intervention and found low rates of agreement across informants, with higher adherence ratings provided by clinicians and caregivers than by expert raters. The researchers attributed this discrepancy to overestimation of adherence by practitioners and caregivers; however, it may be helpful to consider the context from which the raters reported on adherence. In general, external raters tend to rate more strictly than do practitioners and participating caregivers (e.g., Herschell et al., 2019). This may potentially arise because external raters focus exclusively on rating adherence during coding procedures, whereas those participating in the session as either practitioners or clients may be managing the demands of guiding or participation along with recalling or tracking adherence. To maximize the benefits associated with multiple informants of adherence, I incorporated both practitioner self-report and external coder ratings in manuscript two.

Assessment of Content Adherence

The most common measures of adherence in parenting programs consist of self-report checklists (e.g., Leader Session Checklist, Incredible Years; Stern, Alaggia, Watson, & Morton, 2008; MST Consultant Adherence Measure, Multisystemic Therapy; Schoenwald, Sheidow, & Letourneau, 2004; Session-specific checklists from practitioners' manual, Triple P – Positive Parenting Program; Sanders, Markie-Dadds, & Turner, 2000) and observational ratings (e.g., Fidelity Checklist; Breitenstein et al., 2010). Most tools consist of a checklist on which practitioners indicate whether they completed session activities described in a manual. In regular implementation of Triple P, practitioners are encouraged to track their adherence by completing a session checklist after each session. These checklists are first presented in training and are readily available for regular use (Sanders et al., 2000). An example of a session checklist can be

found in [Appendix A](#). Session activities include welcoming parents, outlining the session's goals and agenda, various exercises, concluding review of the session, a reminder of homework tasks, and a preview of the plan for the following session.

Among observational measures, some researchers employ content checklists that align with the program's manual (e.g., Lyon & Budd, 2010). Some approaches supplement a self-report content checklist of session activities with an observational measure. For example, in the Fidelity Checklist (Breitenstein et al., 2010), the practitioner is first asked to indicate whether or not he or she completed each expected action (answers are coded as "yes" or "no") to assess adherence to the program's protocol. An observer completes the same adherence measure.

Assessment of Process Adherence

Among parenting programs, evaluations of program content are far more common than assessment of a program's underlying processes. Notable exceptions consist of observational measures developed for the Parent Management Training – Oregon Model (PTMO; Forgatch & Patterson, 2010) and Triple P (Sanders, 1999, 2012) programs. The process adherence measure for PTMO involves observational coding by a trained rater of a video segment of a PTMO session that focuses on a core program parenting practice. Research using this measure has reported high levels of reliability and validity (Knutson, Forgatch, Rains, & Sigmarsson, 2009). The PTMO program focuses on supporting parents in promoting child adjustment through strengthening parents' use of the five core parenting processes outlined in social learning theory (SLT; Bandura, 1971; Patterson, 1982, 2005). These consist of skill encouragement (or techniques to promote positive behaviours), limit setting (or discipline techniques to discourage negative behaviours), monitoring, tracking behaviour, and positive involvement. The Fidelity of Implementation Rating Scale (FIMP; Knutson et al., 2009) adherence measure is used to rate a

section of a recording focused on an example of either skill development or limit setting, and is rated across five domains that are theoretically consistent with the program's principles. These consist of (i) knowledge of PTMO principles, (ii) session structure, (iii) teaching techniques, (iv) process skills (such as the use of metaphors or connecting content to parents' experiences, managing resistance), and (v) overall development or engaging family in the program's methods (e.g., adjusting to client's context or needs).

As SLT is foundational to both the PTMO and Triple P programs, there are overlapping elements in adherence measurement, and Triple P developers created an observational coding system that also incorporated the SLT core parenting processes: the Adherence Measure for Process Quality in Standard Triple P (AMPQ; Kirby & Sanders, 2014a, 2014b). Instead of examining adherence across multiple constructs (such as the knowledge and teaching domains of the FIMP), this tool provides a global measure of process adherence to the use of the self-regulatory framework by practitioners with parents, and was designed to be used in conjunction with the session checklists evaluating program content. To evaluate both content and process adherence, a trained rater observes a recording of a session and first completes the same session checklist as the practitioner. This provides an observer assessment of adherence to session content to compare with the practitioner's self-report. The rater then completes the global measure of process adherence using the AMPQ to assess the practitioner's adherence to Triple P's underlying processes in a given session. The AMPQ items focus on collaborative interaction with parents and the use of the self-regulation framework to support parents in developing and pursuing their own realistic goals. To date, the reliability of scores on this measure has not yet been reported, and there are currently no self-report measures of process adherence available to practitioners to support multi-informant assessment of both content and process adherence to the

Triple P program.

Limitations of Past Research

As part of evidence-based practices, practitioners are encouraged to tailor their delivery of a program to suit the family's needs (Eyberg, 2005; Huey et al., 2014). However, it is also essential to balance this with adherence to the program's evidence base (Haslam & Mejia, 2018). It is hypothesized that practitioners' promotion of self-regulation will support parents' program engagement and maintenance of gains (Haslam & Mejia, 2018; Turner, Sanders, et al., 2018). The development of measures to assess practitioners' adherence to program processes is at an early stage. Evaluating the degree to which practitioners promote self-regulation among parents (such as in collaboratively setting the session's agenda with parents or guiding them to select their own goals and strategies) is more complex than reporting completion of specific session content (such as whether or not an agenda was used, or if a particular video clip was shown). Evidence-based parenting programs already encourage self-assessments by practitioners of adherence to program content. Assessment of adherence to its underlying processes are rare, however, as are opportunities for practitioners to track the adjustments they make in pursuit of flexible program implementation. This highlights the need for a self-report tool that can assess adherence to both the program's content and underlying processes.

Despite the encouraging findings of the benefits of parenting interventions (e.g., Barlow & Coren, 2018; Barlow et al., 2014), there is growing evidence that many families who could benefit from parenting support do not receive it or drop out of services (Kazdin & Blase, 2011). A traditional therapy model of parenting intervention may reach relatively few families (Sanders, Markie-Dadds, Rinaldis, Firman, & Baig, 2007) and have limited impact upon the prevalence of social and emotional problems among children at a population level (Prinz & Sanders, 2007),

indicating the need to increase the reach of parenting supports (Sanders & Prinz, 2018).

Dissertation Objectives

The research questions of this dissertation are three-fold: first, what is the level of adjustment reported in children and parents (as compared to normative samples) following migration as an immigrant or a refugee? Second, what is known about parenting among immigrant and refugee families who have recently migrated? Third, in supporting practitioners implementing a parenting program in a culturally diverse community, how can we measure the extent to which practitioners implement their services according to the program's evidence base? The meta-analysis (manuscript one) was designed to provide a clearer understanding of child and parental adjustment and parenting among immigrant and refugee families. Manuscript two (published in Practice Innovations) presents preliminary data on two measures of adherence to a parenting intervention, one of which I developed. Manuscript two evaluated the feasibility of these tools in a community implementation operating within a public health approach to inform flexible program delivery to diverse families. Together, these two studies will contribute to our understanding of the needs of diverse families. The overarching aim of these projects is support providing evidence-based parenting supports that both promote adherence to the program's evidence base and practitioners' flexible implementation to best meet the needs of diverse parents.

Table 1*Summary of Reviews of Adjustment Among Immigrant and Refugee Children and Adolescents*

Study	Included studies	Construct(s) examined	Measurement	Issues
Blackmore et al. 2020	Meta-analysis: 8 studies of refugee or asylum-seeking youth ($N = 779$) published 2003-2018	Diagnosis of: • ADHD • Anxiety (GAD, separation anxiety, agoraphobia, OCD, social anxiety, or specific phobia disorder) • Depressive (major or dysthymic) disorder • ODD • PTSD	• Youth self-report • Structured diagnostic interviews • Calculated prevalence rates of diagnoses; compared to world estimates of nonmigrant youth	• Did not report evidence of psychometric adequacy of measures used with refugees • Other raters of child adjustment (parents, teachers) not included • Due to small number of studies, no subgroup analyses for sex for PTSD or depression diagnoses
Bronstein & Montgomery, 2011	Systematic review: 22 studies (14 unique samples) of refugee youth ($N = 3,003$) published 2003-2008	Symptoms of: • Depression • PTSD • General Externalizing and Internalizing problems	• Youth self-report • Standardized measures of adjustment • Five of the 14 unique samples were compared to nonmigrant peers	• Did not report evidence of psychometric adequacy of measures used with refugees • Other raters of child adjustment (parents, teachers) not included • Nine of the 14 unique samples did not include a nonmigrant comparison group
Curtis et al., 2018	Systematic review: 47 studies of combined immigrant and refugee samples (38 studies included mental health: $N = 11,360$) published 2007-2016	Symptoms of: • Anxiety • Depression • PTSD	• Youth self-report • Use of measures with evidence of psychometric adequacy was not an inclusion criterion; measures not described • Migrant samples compared to nonmigrant or to second-generation migrant youth	• Did not report evidence of psychometric adequacy of measures used with immigrants/refugees • Other raters of child adjustment (parents, teachers) not included • Use of measures with evidence of psychometric adequacy was not an inclusion criterion • Did not describe measures used • Limited to studies conducted in Europe
Dimitrova et al., 2016	Meta-analysis: 51 studies of immigrant youth ($N = 224,197$) published 1995-2011	Symptoms of: • General Externalizing and Internalizing problems Level of:	• Youth self-report, parent and teacher reports • Use of measures with evidence of psychometric	• Did not report evidence of psychometric adequacy of measures used with immigrants • Presentation of results did not differentiate

Study	Included studies	Construct(s) examined	Measurement	Issues
		Academic functioning	adequacy was not an inclusion criterion; measures not described Migrant samples compared to nonmigrant or to second-generation migrant youth	between informants of child adjustment - Use of measures with evidence of psychometric adequacy was not an inclusion criterion - Limited to studies conducted in Europe - Did not specify how accounted for multiple informants in meta-analyses
Fazel et al., 2005	Meta-analysis: 5 studies of refugee children ($N = 260$) published 1966-2002	Diagnosis of: PTSD	- Youth self-report & parent report - Structured diagnostic interviews / interview-based psychiatric survey	- Did not report evidence of psychometric adequacy of measures used with refugees - Presentation of results did not differentiate between informants of child adjustment - Limited to studies conducted in western countries
Fazel et al., 2012	Systematic review: 44 studies of refugee and internally displaced youth ($N = 5,776$) published 1980-2010	Symptoms of: - Anxiety - Depression - Externalizing problems - PTSD	- Informants not defined; given age range of included studies, likely both youth self-report and parent reports included, as well as observational measures - Use of measures with evidence of psychometric adequacy was not an inclusion criterion; however, some studies used standardized measures	- Did not report evidence of psychometric adequacy of measures used with refugees - Presentation of results did not differentiate between informants of child adjustment - Use of measures with evidence of psychometric adequacy was not an inclusion criterion
Kien et al., 2019	Systematic review: 47 studies of refugee and asylum-seeking youth ($N = 24,786$) published 1990-2017	Symptoms or Diagnosis of: - Anxiety (GAD, social phobia, panic) disorders - Major depression - PTSD - General Emotional and Behavioural problems	- Youth self-report & proxy-assessments - Standardized measures of adjustment (self-report, proxy-assessment, or structured clinical interview)	- Did not report evidence of psychometric adequacy of measures used with refugees - Presentation of results did not differentiate between informants of child adjustment - Limited to studies conducted in Europe - Few studies ($k = 5$) included a nonmigrant comparison group
Marley & Mauki, 2019	Systematic review: 11 studies of refugee	Resilience or protective factors:	- Informants unspecified - Use of measures with	Authors reported that most included studies did not report on the cross-cultural

Study	Included studies	Construct(s) examined	Measurement	Issues
	youth ($N = 2,959$) published 2000-2015	• Self-efficacy • Self-esteem • Quality of life	evidence of psychometric adequacy was not an inclusion criterion; however, many studies used standardized measures • Comparison groups inconsistently described	validity of the measures used; reliability of measures not reported • Use of measures with evidence of psychometric adequacy was not an inclusion criterion • Presentation of results did not differentiate between informants of resilience constructs • Synthesis of results did not differentiate between constructs of resilience, referring generally to “positive outcomes” • Years in host country not reported
Porter & Haslam, 2005	Meta-analysis: 56 studies (16 youth samples) of refugee and internally displaced adults, adolescents, and children ($N = 22,221$) published 1959-2002	Broadly defined as “mental health” problems	• Youth self-report & parent reports • Use of measures with evidence of psychometric adequacy was not an inclusion criterion; however, some studies used standardized measures • Refugee samples compared to nonrefugee (immigrant) or to nonmigrant comparison groups	• Did not report evidence of psychometric adequacy of measures used with refugees • Presentation of results did not differentiate between informants of child adjustment • Results mainly not presented separately by developmental stage (little information specific to children and adolescents) • Use of measures with evidence of psychometric adequacy was not an inclusion criterion • Comparison groups varied widely across studies • Presentation of results did not differentiate between constructs of child adjustment
Reed et al., 2012	Systematic review: 27 studies of refugee and internally displaced youth ($N = 5,765$) published 1980-2010	Symptoms of: • Anxiety • Depression • Externalizing problems • PTSD	• Informants not defined; given age range of included studies, likely both youth self-report and parent reports included • Use of measures with evidence of psychometric adequacy was not an inclusion criterion; however, some studies used	• Did not report evidence of psychometric adequacy of measures used with refugees • Presentation of results did not differentiate between informants of child adjustment • Use of measures with evidence of psychometric adequacy was not an inclusion criterion

Study	Included studies	Construct(s) examined	Measurement	Issues
			standardized measures	
Stevens & Vollebergh, 2008	Systematic review: 20 studies of labour/economic immigrant youth ($N = 11, 663$) published 1990s-date unspecified	Broadly defined as “mental health” or “adjustment” problems	- Youth self-report, parent, and teacher reports - Standardized measures of adjustment - Immigrant youth compared to nonmigrant peers in same host country	- Authors noted that few studies reported on the reliability or validity of the instruments used with the migrant sample - Authors reported heterogeneity across informants of child adjustment precluded drawing conclusions about degree of adjustment problems among immigrant compared to nonmigrant youth
Tam et al., 2017	Systematic review: 11 longitudinal studies of forcibly displaced (refugee, asylum-seeking, internally displaced) youth ($N = 1,212$) published (date unspecified)-2014	Symptoms of: PTSD	- Informants not defined; given age range of included studies, likely both youth self-report and parent reports included - Use of measures with evidence of psychometric adequacy was not an inclusion criterion; however, some studies used standardized measures	- Authors reported 10 of the included studies met criteria for an adequate outcome measure (reliability coefficient of .75 or greater) - Presentation of results did not differentiate between informants of child adjustment - Authors described the overall quality of included studies as low
van Os et al., 2016	Systematic review: 8 studies of “recently arrived” refugee and asylum-seeking youth ($N = 2,585$) published 1965-2015	Symptoms of: - Anxiety (separation anxiety; other anxiety symptoms unspecified) - Depression - General Internalizing and Externalizing problems - PTSD	- Youth self-report, parent and teacher reports - Use of measures with evidence of psychometric adequacy was not an inclusion criterion; however, most studies used standardized measures	- Did not report evidence of psychometric adequacy of measures used with refugees - Presentation of results did not differentiate between informants of child adjustment

ADHD = Attention-deficit/Hyperactivity Disorder; *GAD* = Generalized Anxiety Disorder; *OCD* = Obsessive Compulsive Disorder; *ODD* = Oppositional Defiant Disorder; *PTSD* = Posttraumatic Stress Disorder

Table 2*Summary of Reviews of Adjustment Among Immigrant and Refugee Adults*

Study	Included studies	Construct(s) examined	Measurement	Issues
Anderson et al., 2017	Meta-analysis: 53 studies of immigrant and refugee mothers ($N = 120,212$; sample size missing $k = 1$) published (database inception)-2015	Symptoms of: • Anxiety (specific disorders unspecified) • Depression (antenatal and postnatal) • PTSD	• Standardized (diagnostic or screening) measures of adjustment • Migrant samples compared to nonmigrant adults, or migrant groups compared to one another	• Did not report evidence of psychometric adequacy of measures used with immigrants & refugees • Data not presented separately for immigrants and refugees • Majority of studies assessed depression (50/53), yielding limited information for other constructs
Falah-Hassani et al., 2015	Meta-analysis: 22 studies of immigrant mothers ($N = 13,749$) published 1950-2014	Symptoms of: • Depression (postpartum)	• Standardized measures • 16 of the included studies compared immigrant to nonmigrants	• Did not report evidence of psychometric adequacy of measures with immigrants • Six of the included studies did not include a nonmigrant comparison group
Fazel et al., 2005	Meta-analysis: 20 studies of refugee adults ($N = 6,743$) published 1966-2002	Diagnosis of: • GAD • Major depression • PTSD • Psychotic disorder	• Structured diagnostic interviews / interview-based psychiatric survey	• Did not report evidence of psychometric adequacy of measures used with refugees • Limited to studies conducted in western countries
Fellmeth et al., 2017	Meta-analysis: 40 studies of immigrant and refugee mothers ($N = 7,895$) published (date unspecified)-2015	Symptoms of: • Anxiety • Depression • PTSD	• Standardized measures of adjustment • 14 studies compared migrants to nonmigrants	• Did not report evidence of psychometric adequacy of measures used with immigrants & refugees • Data not presented separately for immigrants and refugees
Porter & Haslam, 2005	Meta-analysis: 56 studies of refugee and internally displaced adults, adolescents, and children ($N = 22,221$) published 1959-2002	Broadly defined as “mental health”	• Use of measures with evidence of psychometric adequacy was not an inclusion criterion; however, some studies used standardized measures • Refugees compared to immigrant or nonmigrant comparisons	• Use of measures with evidence of psychometric adequacy was not an inclusion criterion • Did not report evidence of psychometric adequacy of measures used with refugees • Presentation of results did not differentiate between constructs of adult adjustment

GAD = Generalized Anxiety Disorder; *PTSD* = Posttraumatic Stress Disorder

Study 1

Adjustment and Parenting Among Immigrant and Refugee Families:
A Systematic Review and Meta-Analysis

Abstract

Several meta-analyses (e.g., Dimitrova, Chasiotis, & van de Vijver, 2016; Porter & Haslam, 2005) have examined adjustment in migrant children, adolescents, and adults. Few include information on parenting or parent adjustment (e.g., van Os, Kalverboer, Zijlstra, Post, & Knorth, 2016). This systematic review and meta-analysis of children, adolescents, and parents in newcomer immigrant and refugee families included studies that used measures of adjustment (overall distress, externalizing and internalizing problems, symptoms of anxiety, depression, and trauma, self-esteem) and parenting (parenting behavior and family functioning) with established psychometric properties. Systematic database searches yielded 18,139 abstracts for screening, review of 4,626 full text articles, and inclusion of 31 independent samples. Results highlight important differences between migrant groups and informants: overall, parents reported greater adjustment problems in youth across both immigrant and refugee groups, compared to normative samples. When migrant groups were examined separately, differences across adjustment variables emerged: apart from self-esteem and trauma, refugee youth generally reported higher levels of problems than normative comparison groups. Immigrant youth reported greater problems than normative samples across all adjustment indices, apart from depression. Parent self-reports of adjustment indicated greater problems across all indices for both immigrant and refugee samples compared to normative data. Compared to normative samples, immigrant and refugee youth and parents generally reported fewer problematic parenting behaviors and more adaptive family functioning. Results highlight the importance of presenting findings for individual informants and migrant groups separately.

Keywords: Immigrant, refugee, adjustment, parenting, families, meta-analysis

Adjustment and Parenting Among Immigrant and Refugee Families

The number of globally displaced persons has increased rapidly over recent years, reaching over 20 million refugees in 2019; among those internationally displaced, approximately 38% are under the age of 18 (United Nations High Commissioner for Refugees, 2020). An increasing number of civilians, including children, are impacted by conflicts and war (e.g., Guha-Sapir, Schlüter, Rodriguez-Llanes, Lillywhite, & Hicks, 2018) that contribute to growing rates of internal and international migration (United Nations High Commissioner for Refugees, 2020). When faced with the experience of forced displacement or migration some immigrants and refugees demonstrate resilience (e.g., Marley & Mauki, 2019), but others may be at increased risk of developing significant adjustment issues (Curtis, Thompson, & Fairbrother, 2018). As the number of displaced persons increases and research examining the experiences of immigrants and refugees evolves, it is essential to continually integrate new information about migration experiences to determine if early findings remain applicable to the experiences of more recent immigrant and refugee groups.

Individuals who have migrated from one country to another are generally grouped as immigrants and refugees. The term immigrant, including economic or labor migrant, is generally used in relation to international migration that is freely pursued. The term refugee, including refugees, asylum seekers, detainees, and (externally or internationally) displaced persons, generally refers to forced international migration (European Migration Network, 2018; Perruchoud & Redpath-Cross, 2011). A growing body of research has examined pre- and post-migration experiences which may impact post-migratory adjustment and inform services designed to facilitate adaptation to host societies.

Child and Adolescent Adjustment

Research examining the experiences of migrant youth has been summarized in meta-analyses (e.g., Blackmore et al., 2020; Dimitrova et al., 2016; Fazel, Wheeler, & Danesh, 2005; Porter & Haslam, 2005) and systematic reviews (e.g., Bronstein & Montgomery, 2011; Curtis et al., 2018; Fazel, Reed, Panter-Brick, & Stein, 2012; Kien et al., 2019; Marley & Mauki, 2019; Reed, Fazel, Jones, Panter-Brick, & Stein, 2012; Stevens & Vollebergh, 2008; Tam, Houlihan, & Melendez-Torres, 2017; van Os et al., 2016). Many of these have focused exclusively on either refugee (Blackmore et al., 2020; Bronstein & Montgomery, 2011; Fazel et al., 2005; Fazel et al., 2012; Kien et al., 2019; Marley & Mauki, 2019; Porter & Haslam, 2005; Reed et al., 2012; Tam et al., 2017; van Os et al., 2016) or immigrant youth (Dimitrova et al., 2016; Stevens & Vollebergh, 2008), with fewer including both (Curtis et al., 2018). These reviews have examined distress in terms of whether participants meet criteria for specific diagnoses (e.g., Blackmore et al., 2020) and at the level of symptoms (e.g., Bronstein & Montgomery, 2011). These include general measures of overall distress and of externalizing and/or internalizing behaviors, as well as examining symptoms of anxiety, depression, and trauma/posttraumatic stress disorder (PTSD); symptoms of attention-deficit/hyperactivity disorder (ADHD) and oppositional-defiant disorder (ODD) are less frequently represented (e.g., Blackmore et al., 2020). We will first present meta-analyses and systematic reviews on refugee samples, then immigrant samples, and for combined samples (summarized in [Table 1](#)).

Across the body of research focused on refugees the rates of mental health problems are generally higher among refugee youth compared to their nonmigrant peers. An early meta-analysis of children, adolescents, and adults (Porter & Haslam, 2005) identified that refugees overall reported moderately poorer adjustment ($d = 0.41$) than did nonmigrant comparison groups. This meta-analysis found that children and adolescents tended to fare better than did

adults post-migration, though still worse than nonmigrant peers. Similarly, Fazel and colleagues' (2005) meta-analysis of five studies that used diagnostic interviews to examine PTSD prevalence in child and adolescent refugees reported an overall prevalence rate of 11%, which is higher than prevalence rates among nonmigrants (e.g., Merikangas et al., 2010). Bronstein and Montgomery (2011) conducted a systematic review of studies that used clinical cut-off scores of measures with evidence of psychometric adequacy to determine prevalence rates in refugee youth. These authors reported higher rates of symptoms of PTSD, depression, and of general psychological distress among refugee youth than among nonmigrant peers.

These reviews included studies with wide ranges of how long refugee participants had been in the host country. Addressing this issue, van Os and colleagues (2016) conducted a systematic review among newly arrived refugee youth. The most common mental health problems reported on arrival to host countries were symptoms of PTSD, depression, and anxiety disorders, impacting up to half of refugee youth. The longer-term picture of PTSD among refugee youth was explored by Tam and colleagues (2017) in a systematic review of longitudinal studies among forcibly displaced children. This review included both internally and internationally displaced youth and reported that high PTSD prevalence rates remained relatively stable over time. Kien and colleagues (2019) found that although estimates varied widely across studies, up to one-third of the refugee youth met criteria for diagnoses of depression or anxiety; for PTSD the proportion was up to half. Similarly, Blackmore and colleagues (2020) reported that rates of post-traumatic stress disorder (PTSD), anxiety, depression, and attention-deficit hyperactivity disorder (ADHD), but not oppositional defiant disorder (ODD), were higher in the refugee samples compared to worldwide prevalence rates.

Reviews specific to immigrant youth have generated mixed results. For example, a

systematic review of studies of mental health among labor (or economic) immigrant youth that used measures with evidence of psychometric adequacy conducted by Stevens and Vollebergh (2008) found variability across informants and operationalization of terms (for example, how “migrant” was defined) precluded clear conclusions. Dimitrova and colleagues’ (2016) subsequent meta-analysis of immigrant youth reported poorer internalizing and externalizing adjustment than in comparison groups. Further, Curtis and colleagues’ (2018) systematic review of both immigrant and refugee youth indicated that symptoms of depression and anxiety were elevated among immigrant youth and symptoms of PTSD were elevated among refugee youth (compared to nonmigrant peers).

Across these studies, authors explored the high levels of variability in the reported rates of adjustment problems by examining potential moderating variables. Meta-analytic moderator analyses indicated that refugee youth who had migrated within two years reported greater distress than did those who had been displaced for longer (Blackmore et al., 2020). In addition, levels of distress varied by geographical region among immigrant youth (Dimitrova et al., 2016). Across systematic reviews, authors noted refugee children who experienced stressful life events of greater frequency, intensity, and duration reported higher rates of symptoms (van Os et al., 2016), as did children who directly experienced violence, had poor financial support or unstable post-migration housing, experienced discrimination, or their parents exhibited poorer parental mental health (Fazel et al., 2012). Finally, some evidence from a systematic review of refugee youth suggested that girls were at a heightened risk for internalizing problems, whereas boys exhibited increased risk for externalizing problems (Reed et al., 2012). Further, the reviews conducted by Fazel, Reed, and colleagues identified protective factors, as refugee youth who reported stronger social support (family cohesion, support from peers, positive school

experiences), more stable resettlement conditions in the host country (Fazel et al., 2012), and those who were repatriated to their country of origin (Reed et al., 2012) appeared to be at a lower risk for mental health problems. Marley and Mauki (2019) also focused on protective factors in their systematic review of refugee youth, describing that – despite experiencing great hardship – youth reported more positive mental health outcomes when they also had greater levels of social support (among peers, family and from ethnically-similar community members), higher self-esteem, higher parental education, greater opportunities to maintain one’s cultural identity, increased school and community connectedness, and access to innovative social care services.

Adult Adjustment

Meta-analyses and systematic reviews (summarized in [Table 2](#)) have synthesized the experiences of adult refugee (Fazel et al., 2005; Porter & Haslam, 2005) or adult immigrant (Falah-Hassani, Shiri, Vigod, & Dennis, 2015), or combined immigrant and refugee samples (Anderson et al., 2017; Fellmeth, Fazel, & Plugge, 2017). The parental status of these adults is rarely specified, except when the focus is on women in the perinatal period (Anderson et al., 2017; Falah-Hassani et al., 2015; Fellmeth et al., 2017). None of these reviews included information specific to migrant fathers’ experiences. Evaluation of PTSD, depression (including in the perinatal period), and anxiety were frequently represented; evaluations of psychotic disorders were frequently included in search criteria, but rarely had sufficient data available for meta-analysis (e.g., Fazel et al., 2005).

Across reviews of adult refugees, both Fazel et al. (2005) and Porter and Haslam (2005) indicated that refugees reported high rates of adjustment problems. Specifically, Fazel and colleagues (2005) reported overall prevalence rates for PTSD of 9%, 5% for major depression, and 4% for generalized anxiety disorder in their meta-analysis. Porter and Haslam (2005)

reported that refugees overall displayed poorer adjustment than nonmigrant comparison groups ($d = 0.41$). In a meta-analysis of postpartum immigrant mothers, Falah-Hassani and colleagues (2015) reported that immigrant women were up to twice as likely as nonmigrant mothers to report postpartum depression. Conversely, in Anderson and colleagues' (2017) meta-analysis of both perinatal immigrant and refugee mothers, the authors reported no evidence for an overall increased risk of antenatal or postnatal depression among migrant women, except when results were presented separately by host country: migrant women in Canada ($k = 9$) displayed increased risk for antenatal and postnatal depression compared to samples in the United States ($k = 5$) and Australia ($k = 3$); authors proposed that differences in nonmigrant comparison groups used in each country may account for this variability in results. Another meta-analysis of mental health among perinatal immigrant and refugee mothers from low- and middle-income countries (Fellmeth et al., 2017) built on this finding and reported higher rates of perinatal depression among migrants compared to nonmigrants. The authors noted that there were too few (or no) studies available to draw meaningful comparisons for perinatal rates of anxiety, PTSD, or psychosis.

As in the syntheses of child adjustment, authors attempted to explain high levels of variability in the reported rates of adjustment problems among adult immigrants and refugees by examining potential moderating variables. Although some studies reported that moderator analyses did not explain the heterogeneity of effect sizes across the included samples (Anderson et al., 2017), other meta-analyses identified a number of factors that appeared to exacerbate or protect adults' mental health: migrants who lived longer in the host country tended to report better outcomes (Falah-Hassani et al., 2015; Porter & Haslam, 2005). Women tended to report poorer outcomes than men (Porter & Haslam, 2005). In the perinatal period, women low in

social support (Falah-Hassani et al., 2015; Fellmeth et al., 2017), with poorer marital adjustment, perceived insufficient household income (Falah-Hassani et al., 2015), and with a previous history of depression (Fellmeth et al., 2017) were more likely to report perinatal depression. In Porter and Haslam's (2005) meta-analysis, region of origin emerged as a potential moderator, as refugees from Europe reported more negative outcomes than refugees from Asia and the Middle East. Further, refugees reported poorer adjustment than immigrants; however, when exposure to violence was accounted for, immigrants who had been exposed to violence reported adjustment that was comparable to refugee samples (Porter & Haslam, 2005). Both Fazel et al. (2005) and Porter and Haslam (2005) identified methodological quality as an important factor, as studies with greater methodological rigor yielded lower estimates of adjustment problems than lower quality studies.

Parenting

It is well established that migration may challenge the adjustment of both adults and children/adolescents (Porter & Haslam, 2005), and that immigrant and refugee youth may be particularly vulnerable soon after migration (Blackmore et al., 2020). Individuals who have been forced to flee, exposed to violence, and/or traumatic experiences may be particularly at risk for poorer outcomes (Porter & Haslam, 2005). As children and families are increasingly exposed to large-scale violence (Guha-Sapir et al., 2018), it is essential to understand the potential impact of stressors associated with migration on children's and parents' adjustment.

The relations between parenting and child adjustment (Hou et al., 2020; Ryan & Ollendick, 2018), and between parental adjustment and parenting (Barlow, Smailagic, Huband, Roloff, & Bennett, 2014) are well-established: higher parental distress is associated with greater problematic parenting practices. In addition, greater parental adjustment problems are connected

to increased child adjustment problems (van Santvoort et al., 2015). The links between child behavior and parenting are reciprocal (Yan & Ansari, 2016; Zvara, Sheppard, & Cox, 2018). As members of immigrant and refugee families may be at increased risk of developing adjustment difficulties which, among parents, may challenge parenting, it is crucial to consider both adjustment and parenting in these families. Further, as evidence-based parenting programs increase the frequency of positive parenting behaviors (Kaminski, Valle, Filene, & Boyle, 2008; Sanders, Kirby, Tellegen, & Day, 2014), and improve both child behavior (Dretzke et al., 2009; Michelson, Davenport, Dretzke, Barlow, & Day, 2013) and parental adjustment (Barlow et al., 2014; Sanders et al., 2014) in culturally diverse families (Haslam & Mejia, 2018), examining these constructs in immigrant and refugee families may provide essential information to inform services for diverse families.

Although it is important to consider the family context of immigrant and refugee families, systematic reviews and meta-analyses of this topic are rare. For example, in their systematic review of newly arrived refugee youth, van Os and colleagues (2016) reported that almost none of the included studies presented information on the child-rearing environment. More recently, Mugadza and colleagues (2019) conducted a systematic review of child-rearing practices among Sub-Saharan migrant families. This primarily qualitative synthesis, however, included both studies of migrant ($k = 27$) families settled mainly in western countries and nonmigrant families ($k = 29$) residing in Sub-Saharan countries. The thematic results did not systematically differentiate between migrant and nonmigrant families, rendering extraction of themes specific to migrant families challenging. There is a significant need to clarify the parenting practices and needs of migrant families.

Examining the Perspectives of Different Raters of Adjustment and Parenting

In evaluating behavior of and services for children, the benefits of examining multiple perspectives are well-established (Hunsley & Mash, 2018). Achenbach, McConaughy, and Howell's (1987) seminal meta-analysis of agreement across raters in multi-informant, multimethod assessments of child and adolescent psychopathology revealed that correspondence across raters was consistently low to moderate (average Pearson $r = .28$; Achenbach et al., 1987). A robust literature has replicated this finding many times confirming that agreement between raters is imperfect for rating children's behavior (De Los Reyes & Kazdin, 2005). This pattern is also evident in discrepancies between adolescents and parents in their evaluations of parenting behavior (e.g., Hou et al., 2020). Despite these well-replicated findings, there is a continuing tendency to view converging reports as a measure of reporting accuracy, treating discrepancies as error or an artifact of measurement (De Los Reyes, Thomas, Goodman, & Kundey, 2013), when these discrepancies may reflect real differences. Previous meta-analyses and systematic reviews of child adjustment often focused only on child/adolescent self-reports (Blackmore et al., 2020; Bronstein & Montgomery, 2011), which precluded evaluations of younger children who could not reliably self-report, and the perspectives of others, such as parents and teachers. Given the known variability across raters, it is important to access information from multiple perspectives. Apart from Stevens and Vollebergh (2008), systematic reviews and meta-analyses of child adjustment often did not distinguish between different informants of child adjustment (e.g., Dimitrova et al., 2016; Kien et al., 2019; Marley & Mauki, 2019). Given the well-established differences across raters, it is important to examine these reports separately.

Measures of Adjustment and Parenting with Evidence of Psychometric Adequacy

To increase confidence in our results, one of our inclusion criteria was for studies to employ measures with demonstrated evidence of reliability and validity. However, psychometric

properties refer to a measure as used with a specific sample. It cannot be assumed that a tool that is psychometrically adequate in the sample for which it was originally developed is also adequate for the participants of the included studies (Hunsley & Mash, 2018). A challenge in research with immigrant and refugee participants is that often a tool with demonstrated psychometric strength has not been validated for the specific ethnocultural group with which it is used.

Translation and back-translation do not guarantee identical meaning across languages. Nor does use of a measure developed with one cultural group, when applied to a second group, guarantee that the second group will interpret the items' meaning in the same way, or describe symptoms or behavior in the same manner. Both concerns may potentially reduce the reliability and validity of scores on the measure. An alternative could be to create a new measure specific to a cultural group, however this may mean large-scale studies to establish the reliability and validity of the measure across various groups, which may not be consistently feasible.

At a minimum, it is important for researchers to use measures with demonstrated psychometric adequacy. It is also essential to report how the measure performed in each sample in which it is used, as it cannot be assumed that reliability and validity evidence found in a normative sample will be replicated in other culturally diverse samples. Although the systematic reviews or meta-analyses conducted by Stevens and Vollebergh (2008), Bronstein and Montgomery (2011), Fellmeth et al. (2017), Falah-Hassani et al. (2015), and Kien et al. (2019) stressed the necessity of using psychometrically sound instruments to measure distress, many reviews continued to include psychometrically weak or untested measures (e.g., Fazel et al., 2012; Marley & Mauki, 2019; Reed et al., 2012) or did not describe the measures used in the included studies (Curtis et al., 2018; Dimitrova et al., 2016). Further, reviews and meta-analyses often did not report on the psychometric adequacy of the measures used in the included studies

with immigrant and/or refugee participants, limiting understanding of the reliability and validity evidence for these tools when used with these culturally diverse samples. To fill this gap, we selected studies that used psychometrically sound measures of adjustment and parenting and extracted data reporting on the measure's internal consistency with immigrant/refugee samples. Studies that reported inadequate reliability (determined using benchmarks outlined by Hunsley & Mash, 2018) for a measure used with an immigrant/refugee sample were excluded. Further, studies that did not report on the reliability of the measure used with the immigrant/refugee sample were excluded, as we could not confirm psychometric adequacy of the tools used with the sample.

Limitations of Past Research

Although the studies reviewed above have helped to illuminate the adjustment experiences of immigrant and refugee children, adolescents, and adults, there are important limitations to these bodies of research: the available meta-analyses often did not describe the reliability or validity evidence for the measures used in the included studies (Dimitrova et al., 2016), did not describe the measures used (Curtis et al., 2018; Dimitrova et al., 2016) or included a small number of studies (Blackmore et al., 2020), limiting the generalizability of their findings. Bronstein and Montgomery (2011), Curtis et al. (2018), Fazel, Reed, and colleagues (2012; 2012), Kien et al. (2019), Stevens and Vollebergh (2008), and van Os et al. (2016) all reported difficulty using meta-analysis to accommodate methodological and instrument variability. This has yielded mainly systematic reviews of the literature among immigrant or refugee children and adolescents to date rather than quantitative comparisons that accommodate both. Further, in presenting themes identified across the included studies narratively, and not quantitatively examining child adjustment, integration and interpretation of findings is challenging. Some

reviews of child adjustment were limited to self-report (Bronstein & Montgomery, 2011; Curtis et al., 2018), others described challenges accounting for variability across raters (Stevens & Vollebergh, 2008), or did not distinguish between different informants of child adjustment (Kien et al., 2019; Marley & Mauki, 2019). Among reviews about migrant parents, many have focused on mothers' adjustment experiences, particularly in the perinatal period and little is available that is specific to migrant fathers' adjustment. Finally, there is a dearth of syntheses of studies examining parenting among migrant families that could help inform supports for these families should they desire them.

Goals of the Meta-Analysis

The aims of this meta-analysis are five-fold. First, we aim to update the state of the literature of immigrant and refugee families by describing the demographic information of the included studies examining adjustment and parenting using measures with adequate psychometric properties. Second, we report on youth adjustment, presenting data separately for immigrant and refugee children and adolescents where possible. Third, we present results of adjustment among immigrant and refugee adults identified as parents, differentiating between immigrant and refugee parents when sufficient data are available. Fourth, we report how immigrant and refugee families describe parenting behavior and parent-child relationships. Finally, we planned to explore potential moderators of adjustment and parenting among immigrant and refugee families should adequate data be available.

Meta-analytic techniques were used to calculate an average estimate of the magnitude of difference in adjustment and parenting between immigrant/refugee youth and parents and comparison groups. Although there have been a series of literature and systematic reviews published on the adjustment of immigrant/refugee youth that have highlighted the needs of these

vulnerable individuals, few have calculated the size of the effect comparing these children to groups of nonmigrants. Further, some of these studies have included psychometrically weak (e.g., Reed et al., 2012), unidentified, or untested measures (e.g., Dimitrova et al., 2016; Fazel et al., 2012). Compared to the wealth of studies examining adjustment among immigrant/refugee adults, there have been fewer studies examining the adjustment of parents who have migrated (and their parenting behavior). It was hypothesized that immigrant/refugee youth and parents would report more adjustment and parenting problems than the comparison groups. The use of meta-analysis in the current studies will permit quantitative comparisons among diverse families across various countries, extending our understanding of adjustment among children and adults, parenting behavior and the parent-child relationship among immigrant and refugee families.

Method

Meta-Analysis Overview

Following a systematic review of the research conducted in English and French with psychometrically adequate assessment tools, this meta-analysis was conducted to synthesize research findings on (i) child and adolescent adjustment and (ii) parental adjustment and parenting among immigrant and refugee families. The review protocol was developed by the research team prior to conducting a systematic review of the literature and refined iteratively as coding progressed. In line with PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, & Group, 2009), [Figure 1](#) provides an overview of the screening, identification, eligibility assessment and inclusion criteria for the meta-analysis.

Operationalization of Terms

In this meta-analysis, we defined “adjustment” as mental health, psychopathology, well-being, adaptation, or psychological resilience. Studies that assessed one or more of these

constructs (among children, adolescents, or parents) was eligible for inclusion. Within adjustment, we differentiated between measures of problems (i.e., distress and symptoms, such as externalizing or internalizing problems) and positive measures (i.e., self-esteem and social support). Parenting was examined in terms of parent self-report of parenting behaviors; (e.g., APQ; Alabama Parenting Questionnaire; Frick, Christian, & Wootton, 1999) or children's reports on their parents' behavior (e.g., CRPBI; Children's Report on Parent Behavior Inventory; Schludermann & Schludermann, 1988). In addition, youth and parent report measures of the parent-child relationship, such as family functioning (e.g., FAD; McMaster Family Assessment Device; Epstein, Baldwin, & Bishop, 1983), were included.

Inclusion and Exclusion Criteria

Studies published from 1994 to January 2020 addressing adjustment and/or parenting among migrant families were examined for inclusion in the systematic review according to several criteria. A 25-year period was selected to ensure a broad representation across publications: as no one region of origin or settlement was the focus of the meta-analysis, and we examined multiple adjustment and parenting constructs among both children/adolescent and adults, we could not anchor our search to a specific change in the literature or to a previous meta-analysis. First, the studies focused on the experiences of immigrants or refugees that met a "newcomer" criterion (residing in their host country for less than 10 years). This was determined by the mean number of years since migration for the sample; when children were born in the host country, their parents had to have migrated within the previous 10 years. Second, as a primary focus of the meta-analysis is to understand potential implications of adjustment on parenting support, studies were limited to children and adolescents who migrated with their parents. Studies were excluded if separation between parent and child exceeded two years. Studies of

unaccompanied minors were excluded. To be included, the mean age of the child/adolescent sample needed to be below 18 years; the upper limit of the age range could extend to age 19 provided the mean age was under 18 years; parents needed to be age 18 and above. Third, the only methodological criterion applied to the selection of studies was the use of a psychometrically adequate (as determined using benchmarks outlined by Hunsley & Mash, 2018) measure of child adjustment, parental adjustment, and/or parenting, and that authors reported on the reliability of the measure as used with the immigrant and/or refugee sample. To include as many studies as possible, there were no limitations placed on the type of research design or sample size. Finally, to conduct a synthesis of the literature that incorporated both immigrant and refugee samples, studies published in any journal in English or French between 1994 and January 27, 2020 were eligible for inclusion. Finally, no limits were placed on geographical locations or cultural groups, as any immigrant or refugee participants were eligible for inclusion, provided the above criteria were satisfied.

Selection of Studies

Consistent with recent reviews in this area (e.g., Dimitrova et al., 2016; Fazel et al., 2012; Gadeberg, Montgomery, Frederiksen, & Norredam, 2017) the following search engines were accessed to identify potential studies: PsycINFO, Medline (Pubmed), EMBase, the Cochrane Review, Cinahl, Scopus, and Web of Science. Database searches for both meta-analyses were conducted simultaneously with keywords (using the Boolean operator * to access multiple endings to a root word) and mapped heading searches (when available) across three levels: the first targeted the migration experience of the sample, the second focused upon the family members, and the third consisted of adjustment and parenting among migrating families (including mental health and well-being, and parenting). Within each level, the search terms

were paired with “OR”; between levels with “AND.” The searches consisted of variants of [“Immigrant” OR “Refugee” OR “Displaced person” OR “Asylum-seeker”] AND [“Parent” OR “Youth” OR “Adolescent” OR “Child” OR “Family”] AND [“Mental Health” OR “Adjust*” OR “Adapt*” OR “Well being” OR “Parental Psychopathology” OR “Parent-Child Relation”]. The search term “Parenting” was not explicitly included as it was incorporated in the “Parent” search term in level two. The list of search terms across databases is available as [Appendix B](#).

A vast literature suggests that studies reporting higher effect sizes or significant results are more likely to be published than studies yielding lower effect sizes or nonsignificant findings (Borenstein, Hedges, Higgins, & Rothstein, 2009). This publication bias may then extend to a biased interpretation of results in systematic reviews or meta-analyses, rendering it an important consideration in determining the scope of search criteria (Bellefontaine & Lee, 2014; Borenstein et al., 2009). Grey literature, or research that is not available through commercial publishers, consists of research produced by such entities as governments, academia, businesses, or industries (Bellefontaine & Lee, 2014; Gokhale, 1997; Paez, 2017). It may consist of ongoing research, unpublished doctoral dissertations, agency or government reports, conference proceedings, and book chapters, and may be available electronically or in print, and its inclusion in reviews and meta-analyses is purported to minimize publication bias (Borenstein et al., 2009; Pappas & Williams, 2011). In determining our search parameters, we debated the advantages and disadvantages of incorporating grey literature into our search criteria. Using an exemplar meta-analysis conducted by Bellefontaine and Lee (2014) examining parenting interventions in both published and unpublished studies, we sought to examine the benefits of including grey literature in our meta-analysis. The authors reported that the effect sizes calculated for published and unpublished research did not differ significantly for child adjustment, parental adjustment, or

parenting (Bellefontaine & Lee, 2014). Given these findings and the breadth of available research examining the experiences of immigrants and refugees, we decided to limit the database searches to peer-reviewed journals. Database searches identified 37,488 records of which 19,349 were excluded as duplicates. The remaining 18,139 abstracts were reviewed by a team of four undergraduate students and the first author, a PhD student. Only abstracts that were unambiguously irrelevant were excluded. Any abstracts that could potentially be eligible were retained; all disagreements were resolved by consensus, erring on the side of retention. The abstract review yielded 4,626 articles to assess for eligibility.

Coding the Studies

Two teams, one in Canada and one in Australia, independently reviewed the full texts of articles for eligibility. The Canadian team consisted of six members: a PhD psychologist, a doctoral student in clinical psychology, and four undergraduate students in psychology. The Australian team included four PhD psychologists. Assessment of articles consisted of two steps, each following a coding protocol: the first to determine the eligibility of the article (the Inclusion protocol; [Appendix C](#)), the second to extract the data from included studies using psychometrically adequate measures (the Data Extraction protocol; [Appendix D](#)).

Development of Coding Protocols

To develop the inclusion and data extraction protocols, both coding teams reviewed 877 full texts from an initial PsycINFO database search. Both protocols were developed iteratively, refined following feedback from both teams and multiple rounds of practice coding, before being applied to the remaining database search results.

To evaluate the eligibility of the available studies for inclusion in the meta-analysis both coding teams reviewed the method section of each article, using the following inclusion criteria:

(i) participants described as immigrants or refugees who resided in their host country for less than 10 years (i.e., “newcomers”); (ii) children/adolescents migrated with at least one parent and the mean age of the child/adolescent sample was below 18 years (and the age range did not extend beyond age 19); parents were 18 years or older; and (iii) the study included at least one measure of child adjustment and/or parental adjustment and/or parenting as an outcome variable. Inter-rater agreement (IRR) was determined by the intra-class correlation (ICC; absolute agreement) between the two teams for inclusion. IRR was .79 and discrepancies were resolved by consensus across the two teams.

The data extraction coding protocol addressed whether the available studies used psychometrically adequate measures and extracted data from included studies. To determine whether measures were psychometrically adequate, we applied the criteria proposed by Hunsley and Mash (2018) to assess the reliability of scores provided by assessment tools. Next, we developed a list of psychometrically adequate measures of child adjustment, adult adjustment, and parenting. We consulted reviews of psychometrically adequate measures for children and adolescents, including evaluation of conduct problems (Frick & McMahon, 2018), depressive symptoms (Dougherty, Klein, & Olino, 2018), self-injurious thoughts and behaviors (Millner & Nock, 2018), symptoms of anxiety disorders (Byrne, Lebowitz, Ollendick, & Silverman, 2018), and PTSD (Keane, Silberbogen, & Weierich, 2008). We then examined reviews of measures of adult adjustment, including depressive symptoms (Persons & Fresco, 2008; Persons, Fresco, & Ernst, 2018), self-injurious thoughts and behaviors (Millner & Nock, 2018), generalized anxiety (Dugas, Charette, & Gervais, 2018), PTSD (Keane et al., 2008; Moshier, Parker-Guilbert, Marx, & Keane, 2018), and substance (Rohsenow, 2018) and alcohol use (Green, Worden, Menges, & McCrady, 2008; Haeny, Boness, McDowell, & Sher, 2018). Additional measures were located

through Beidas and colleagues' (2015) review of free and psychometrically sound measures of youth and adult mental health, and Gadeberg et al.'s (2017) review of tools with evidence of psychometric adequacy to assess trauma and mental health among refugee youth. We then reviewed numerous resources to identify measures of parenting behavior and the parent-child relationship (Cunningham & Scott, 2012; Hamilton & Carr, 2016; Hampson & Beavers, 2012; Hurley, Huscroft-D'Angelo, Trout, Griffith, & Epstein, 2014; Olafson & Connelly, 2012; Sperry, 2012; Yingling, 2012). We also reviewed the measures employed in a subset of potentially eligible articles ($n = 100$), identified 128 distinct measures of adjustment or parenting used in these studies, and evaluated evidence on the psychometric adequacy of these tools, first consulting the reference associated with the measure (a published manual or article) to identify the age range for which the measure was designed, and then examining evidence of the measure's reliability and validity.

Employing the criteria recommended by Hunsley and Mash (2018), internal consistency of .70 or above was considered adequate. Interrater reliability of (a) κ values of .60 or above or (b) Pearson or intraclass correlations of .70 or above were classified as adequate. Adequate test-retest correlations of at least .70 across several days to weeks were also accepted, and/or if norms were available (Hunsley & Mash, 2018). Evidence of validity was acceptable if there was replicated data on content, construct (convergent, divergent, concurrent and/or predictive) validity and/or validity generalization (Hunsley & Mash, 2018). This process yielded 82 measures of child adjustment that have demonstrated psychometric adequacy in the normative sample including reports by youth ($k = 64$), parent ($k = 29$), and "other" reports, such as teacher or clinician ($k = 24$). We identified 78 measures of adult adjustment and 20 measures of parenting (10 measures of parenting behavior, 9 measures of the parent-child relationship, 1

measure of parenting stress). To identify appropriate norms for each measure, first the measure's manual was consulted (if available) and/or peer-reviewed articles that presented the data from the measure's manual. If these were not available, a search of the PsycINFO and PubMed databases was conducted combining the measure's full title, its acronym (combined with "or") and the measure's primary author (combined with "and"). Results were reviewed to locate large and representative samples (that were not primarily immigrants or refugees) to use as a comparison group in the meta-analysis. This process was also pursued for measures developed specifically with refugees, such as the Harvard Trauma Questionnaire (HTQ; Mollica et al., 1992) and the Hopkins Symptoms Checklist-25 (HSCL-25; Mollica, Wyshak, de Marneffe, Khuon, & Lavelle, 1987) to access a comparison sample that was not primarily refugees. A list of the measures is presented in [Appendix E](#).

It is essential in meta-analysis that each study is included in analyses only once to ensure that data are independent (Borenstein et al., 2009). Prior to analyses, multiple publications from the same dataset were identified by reviewing author information and sample characteristics, and the most complete report was selected for the meta-analysis. Further, because psychometric adequacy is a property of a measure in a given sample, it cannot be assumed that reliability and validity found in a normative sample guarantees psychometric adequacy in other culturally diverse samples. Therefore, when presented by the authors of the primary studies, data on the reported reliability of measures among immigrant and refugee samples were extracted to permit its inclusion in the meta-analysis. Articles that did not report the reliability of the measure used with the study's sample, or reported inadequate reliability, were excluded, as the psychometric adequacy of the measure could not be confirmed with the immigrant/refugee sample. Subsequently data (means, standard deviation, and sample size) were extracted for measures of

child adjustment, parental adjustment, and parenting with evidence of psychometric adequacy (for each rater of adjustment or parenting). Consistent with the literature on immigrant and refugee children, adolescents, and adults (e.g., Beiser, 2010; Blackmore et al., 2020; Fazel et al., 2012; Porter & Haslam, 2005), data were extracted on sample characteristics and study methodology to permit examination of potential moderators of adjustment and parenting, including status in host country, years in host country since migration, host country, region of origin, child gender and age, and parent gender and age. A final potential moderator, methodological quality, was also coded as described below.

Methodological Quality

Because studies lower in methodological quality may over-estimate effects (Fazel et al., 2005; Hempel et al., 2011; Porter & Haslam, 2005) we assessed the methodological quality of included studies. We selected a checklist that mapped onto recommended aspects of quality measurement (e.g., Sanderson, Tatt, & Higgins, 2007), using the items that are applicable to nonrandomized controlled trial (RCT) studies: the Checklist for Measuring Quality (CMQ; Downs & Black, 1998), designed to assess the quality of randomized and nonrandomized studies of health care interventions, and has demonstrated adequate reliability and validity (Downs & Black, 1998). Information was drawn from each included study to code the items from the CMQ examining quality of reporting, external validity, study bias, and power, yielding a total score from 0 to 13. Articles were coded by two raters and a subset ($k = 13$) were rated by both raters. Inter-rater agreement (IRR) was determined by the intra-class correlation (ICC; absolute agreement) between the two raters. IRR was .93 and discrepancies were resolved by consensus between raters. Overall, the included studies were rated as relatively high quality ($M = 9.19$; $SD = 0.87$; range = 8 to 11). Full descriptions of coding criteria are presented in [Appendix F](#).

Data Analysis Plan

As the studies included in the meta-analysis were not limited to RCTs or to those with a built-in comparison group, we used the norms of each measure as the comparison group. The means, standard deviations, and sample sizes were entered into the meta-analysis program and paired with the means, standard deviations, and sample sizes of the published norms of the same measure. Some challenges were encountered in locating appropriate norms. Specifically, the Hopkins Symptom Checklist-25 (HSCL-25; Mollica et al., 1987) was developed specifically with refugee populations, and we sought an alternative comparison sample through a systematic search of the PsycINFO and PubMed databases to locate large-scale studies. Glaesmer and colleagues (2014) presented normative data for the HSCL-25 with a large and representative sample in Germany and was the most appropriate option for our purposes for a measure of adult adjustment. Among measures of child trauma, some measures were developed with samples exposed to high levels of violence or disruption (Post-traumatic Stress Symptoms in Children; Ahmad, Sundelin-Wahlsten, Sofi, Qahar, & von Knorring, 2000) or natural disasters (Child PTSD Symptom Scale; Foa, Johnson, Feeny, & Treadwell, 2001). Only the Reactions of Adolescents to Traumatic Stress measure (RATS; Bean, Derluyn, Eurelings-Bontekoe, Broekaert, & Spinhoven, 2006) included a representative sample of nonmigrants from communities without a history of violence or disruption. When data were presented separately for subgroups (e.g., by gender, age, migrant status), the data were entered separately for each subgroup for each available measure of adjustment or parenting. To compare data on newcomer immigrant and refugee families to the norms, an effect size was calculated for each subgroup within each sample, for each available measure of adjustment and parenting, along with 95% confidence intervals, using the Comprehensive Meta-Analysis software program, version 3

professional (Borenstein, Hedges, Higgins, & Rothstein, 2014).

Unit of Analysis

For studies that reported multiple measures (through multiple raters and/or by assessing multiple adjustment or parenting variables), it is important to consider the unit of analysis when selecting the procedure to calculate effect sizes. In the child ($k = 21$) and parental adjustment ($k = 10$) meta-analyses, all studies included only one informant, and several studies (child adjustment: $k = 3$; parental adjustment: $k = 1$) assessed multiple adjustment constructs. All three studies in the parenting meta-analysis incorporated multiple informants and evaluated one parenting construct. To examine data across all studies simultaneously, ensuring that data are independent requires one score for each study or subgroup, and involves calculating one score that averages across all measures of adjustment or parenting (Borenstein et al., 2009), referred to here as a composite score. Although this method enables all studies to be included in a single analysis, composite scores do not permit examination of differences across adjustment or parenting variables, or across raters. It is also possible to conduct a separate meta-analysis for each adjustment and/or parenting construct to examine each individually, provided data for the same participants are not included in the same analysis (Borenstein et al., 2009). This results in fewer studies in each meta-analysis but permits examination of differences across types of adjustment or parenting, and across informants. We selected the latter approach in this meta-analysis, and calculated effect sizes for each measure of adjustment or parenting and conducted separate meta-analyses for each construct. As reports of child/adolescent adjustment and parenting tend to vary across raters (Achenbach et al., 1987; De Los Reyes et al., 2015; Hou et al., 2020), separate meta-analyses were also conducted for each informant of child adjustment and parenting. With this approach, the studies with multiple measures of adjustment and

multiple raters are represented more than once, but never in the same meta-analysis, to ensure data are independent in each analysis. For measures of child and parental adjustment, separate meta-analyses were conducted for positive measures (i.e., measures of self-esteem, such as the Self-Perception Profile for Adolescents; Harter, 2012) and each domain of adjustment problems (i.e., measures of overall problems, such as the total problem score from the Child Behavior Checklist; Achenbach, 1991). Across all measures, higher scores indicate greater problems or more problematic parenting behavior/parent-child relationship; for consistency across measures, this involved reversing the effect direction for two tools that are typically scored such that higher score reflect better functioning (self-esteem; positive parenting behaviors). Across studies, a negative effect size indicates that participants in the immigrant/refugee group rated better adjustment, parenting behavior, or parent-child relationship than did those in the comparison group.

Meta-Analytic Model

We then compared effect sizes across studies using a random-effects model that can more accurately accommodate both homogeneous and heterogeneous studies (Schmidt, Oh, & Hayes, 2009). For each domain of positive and problem measures, we conducted an overall test of heterogeneity of these effect sizes (using a Q statistic) and examined the magnitude of heterogeneity (using I^2 ; Borenstein et al., 2009) to identify potential differences. Because these tests indicated effect sizes were significantly heterogeneous across all indices (as described below), we calculated effect sizes separately for studies that examined immigrants ($k = 15$), refugees ($k = 13$), and samples that combined immigrants and refugees ($k = 3$).

Separate meta-analyses of each adjustment variable are presented in the following sections: results must be interpreted with caution due to the small number of studies included in

several of the analyses. As such, the same study may be represented several times across adjustment variables if multiple assessment tools or multiple raters were used.

Results

Demographic Data

The meta-analysis included 43 articles, describing 31 independent samples with a total of 10,280 participants: 7,581 children or adolescents and 2,699 parents. Among parent samples, 1,851 participants were identified as mothers, 198 as fathers, and 650 as parents. The demographics of the included studies are presented in [Table 3](#) for child adjustment ($k = 21$), in [Table 4](#) for parental adjustment ($k = 10$), and [Table 5](#) for parenting ($k = 3$); 4 studies assessed both child adjustment and parental adjustment or parenting; the remainder assessed only child adjustment ($k = 18$), parental adjustment ($k = 7$), or parenting ($k = 2$). Across the 31 included samples, 15 consisted of immigrants only, 13 included only refugees, and three included a combination of both. The countries of origin represented were diverse: 51 countries were identified by name, spanning the continents (and regions) of Africa (East Africa, Middle Africa, North Africa, West Africa), Asia (East Asia, South-East Asia, South Asia, West Asia), Europe (East Europe, South Europe), and Latin America (Caribbean, Central America, South America). The host countries included the United States ($k = 8$), Australia and Israel ($k = 4$), Canada and Turkey ($k = 3$), Germany, Italy and Jordan ($k = 2$), and Belgium, Lebanon, Netherlands, Spain and Sweden ($k = 1$). Overall, the included studies were rated as relatively high in methodological quality ($M = 9.19$; $SD = 0.87$; range = 8 to 11).

Reports of child adjustment were primarily completed by children/adolescents ($k = 16$), and less frequently by parents ($k = 4$) or teachers ($k = 1$). None of the included studies incorporated multiple raters of child adjustment using psychometrically adequate measures.

Parents reported on their own adjustment across all included studies. All included studies of parenting behavior or the parent-child relationship used multi-informant methodology, and measures were completed by both parents and their children ($k = 3$). Most evaluations of child adjustment assessed overall problems ($k = 9$) or depression ($k = 6$) using a variety of measures. Approximately two-thirds of the evaluations of parental adjustment assessed symptoms of depression ($k = 8$), primarily using the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977) and the Edinburgh Postnatal Depression Scale (EPDS; Cox, Holden, & Sagovsky, 1987). Only three studies assessing parenting met the inclusion criteria with respect to a psychometrically sound measure. Tables 3, 4, and 5 list the measures used in each study; Table 6 lists overall effect sizes; Tables 7, 8, and 9 indicate how often each measure was used.

Measures were organized by construct: for measures of adjustment, measures were identified as assessing (i) overall problems (e.g., Child Behavior Checklist (CBCL) total score; Achenbach, 1991; Symptom Checklist-90-Revised; Derogatis, 1994), (ii) overall levels of externalizing symptoms (e.g., CBCL-externalizing subscale), (iii) overall levels of internalizing symptoms (e.g., CBCL-internalizing subscale), measures of symptoms specific to (iv) anxiety (e.g., Spence Children's Anxiety Scale; Spence, 1997; Hopkins Symptom Checklist-25-anxiety subscale; Mollica et al., 1987), (v) depression (e.g., Children's Depression Inventory; Kovacs, 1985; HSCL-25-depression subscale; Mollica et al., 1987), or (vi) trauma (e.g., Child PTSD Symptom Scale; Foa et al., 2001), as well as a measure of (vii) self-esteem (SPPA; Harter, 2012). Measures of parenting were examined individually, as the measures in the included studies addressed parenting behavior (Alabama Parenting Questionnaire; Frick et al., 1999) and family functioning (McMaster Family Assessment Device; Epstein et al., 1983).

Adjustment Among Immigrant and Refugee Youth

These analyses included 21 studies of child adjustment. Effect sizes and 95% confidence intervals for each adjustment variable are reported in Tables [6](#) (full sample; top section) and [7](#) (presented separately by migrant group—immigrant, refugee or combined); the individual effect size for each included study, for each measure of adjustment, are depicted in Figures [2](#) (youth self-report) and [3](#) (parent and teacher reports). Across the full sample, children/adolescents and parents reported more problems than the normative sample across measures of overall problems ($g = 0.53$, youth self-report; 0.65 , parent report). Evaluations of symptoms of anxiety, depression, trauma and of self-esteem were completed solely by youth self-report. These indicated that immigrant/refugee youth endorsed more symptoms of anxiety, fewer symptoms of depression and trauma, and higher levels of self-esteem compared to normative samples. It is important to note that, apart from anxiety, the 95% confidence intervals of all other youth self-reports of adjustment included zero, indicating that the effect sizes could also fall in the opposite direction. This suggests that although the summary effect indicates differences between migrant samples and normative data, results must be interpreted with caution (Schünemann et al., 2020).

Externalizing and internalizing problems were limited to parent reports of child adjustment. In this single sample, parents reported fewer externalizing ($g = -0.13$), but more internalizing problems ($g = 0.77$) among their children compared to normative comparison groups. The single sample of teacher-reported overall problems indicated fewer problems among migrant youth compared to the normative sample ($g = -0.20$); however, the 95% confidence interval included zero, necessitating caution in the interpretation of this result.

Tests of heterogeneity (using the Q and I^2 statistics) indicated that there was significant variability across all informants and all variables with multiple measures, so we examined reports of adjustment separately by migrant group. Given the significant heterogeneity across

and within variables, we also grouped the analyses by measure to observe differences across tools used to assess child adjustment.

As depicted in [Table 7](#), when multiple assessment tools were available, we reported the summary effect for each adjustment variable (in bold) and indicated (with *) whether heterogeneity tests were significant; when only one measure was available, effects are for that measure. In examining each migrant category independently, differences emerged: overall, refugee children/adolescents more consistently reported or were described by others as experiencing greater adjustment problems than normative samples: the summary effect sizes for each adjustment construct among refugee youth were generally above $g = 0.80$, whereas effect sizes calculated among immigrant-only and combined immigrant/refugee groups generally were lower (although still higher than normative comparisons), suggesting greater problems among refugee youth than among both their immigrant and normative comparison peers. Conversely, the only study assessing self-esteem yielded stronger self-reported self-esteem among refugee youth than the normative sample, and the single study examining trauma among refugee youth indicated lower rates of trauma symptoms compared to normative comparisons. In contrast, immigrant-only and combined samples of immigrant and refugee youth self-reported higher symptoms of trauma compared to normative samples. This apparently counter-intuitive finding may be related to differences in the normative comparison groups' characteristics and will be described in a subsequent section. Further, immigrant-only samples self-reported higher levels of overall problems and symptoms of anxiety, but lower symptoms of depression, than comparison groups.

When examined separately, the pattern of results across parent ratings of each migrant group replicated the pattern observed when all migrant groups were examined simultaneously

(Table 6), indicating that parents generally reported higher overall problems in both immigrant and refugee youth than normative comparisons. Contrary to this, in the single study of combined immigrant and refugee youth, parents reported fewer externalizing, but more internalizing problems, among youth than in comparison groups. Teacher ratings of overall problems were limited to a single study of refugee youth and indicated fewer problems among refugees compared to the normative sample. However, given the small number of studies in each of these analyses, results must be interpreted with caution.

Adjustment Among Immigrant and Refugee Parents

These analyses included 10 studies of parental adjustment. Effect sizes and 95% confidence intervals for each adjustment variable are reported in Tables [6](#) (full sample; middle section) and [8](#) (presented separately by migrant group); individual effect sizes for each study, for each measure of adjustment, are presented in [Figure 4](#). Compared to normative samples, immigrant/refugee parents reported greater adjustment problems across all measures of adjustment, including overall problems, anxiety, and depression. As the Q and I^2 statistics indicated significant variability across studies examining symptoms of depression; we once again conducted separate analyses for each migrant category, grouping results by measure.

As seen in [Table 8](#), when multiple assessment tools were available, we reported the summary effect for each adjustment variable (in bold) and indicated (with *) whether heterogeneity tests were significant; when multiple tools unavailable, effects were reported for each tool specifically. In examining each migrant category independently, the pattern of results was similar to that above, except that although samples consisting of only immigrant or refugee parents continued to report higher levels of depressive symptoms than the normative samples, the one study that combined immigrant and refugee mothers reported lower symptoms of depression

than did the comparison group.

Parenting in Immigrant and Refugee Families

These analyses included three studies examining parenting. Effect sizes and 95% confidence intervals for each parenting measure are reported in Tables [6](#) (full sample; bottom section) and [9](#) (presented separately by migrant group); the individual effect size for each included study are depicted in [Figure 5](#). Both youth and parent reports were limited to one measure each of parenting behavior and family functioning: respectively, the Alabama Parenting Questionnaire (APQ; Frick et al., 1999), which assessed parental monitoring and supervision, use of corporal punishment, positive parenting, parental involvement, and an additional set of discipline strategies (the inconsistent punishment subscale was removed from analyses due to inadequate reliability), and the general functioning subscale of the McMaster Family Assessment Device (FAD; Epstein et al., 1983), which assessed overall health within the family's functioning. Immigrant/refugee youth and parent reports converged on the lower use of problematic parenting behaviors by migrant parents than normative comparisons. However, caution is required as the 95% confidence intervals for each of these reports included zero, indicating that the effect sizes could also fall in the opposite direction. Reports also converged on measures of family functioning, with both youth and their parents reporting better functioning than normative samples. To explore these reports across migrant groups, we once again conducted separate analyses for each category.

As reported in [Table 9](#), we reported effects for each tool individually because of the variability across what the tools were designed to measure. Separate examination of each migrant group revealed the same pattern of results as described above; the only exception was when refugee youth reports were examined specifically, the confidence interval of the effect size

for youth reports of parenting behavior (on the APQ) no longer included zero.

Discussion

This study included a total of 31 independent samples evaluating adjustment and/or parenting among immigrant and/or refugee youth and parents. It included studies published until the end of January 2020. Twenty-five of these studies (17 assessing child adjustment; 8 parental adjustment) were not included in previous meta-analyses or systematic reviews of immigrant or refugee adjustment reviewed earlier. Our searches did not yield any previous meta-analyses and systematic reviews of parenting focused specifically on the experiences of immigrant and refugee families; thus, we did not compare the studies of parenting included in our meta-analysis to previous reviews.

A striking finding is the large number of studies that failed to meet inclusion criteria. Hundreds of studies were excluded because they did not use a measure we identified as being psychometrically adequate ($k = 879$). Others used a measure for which there is evidence of psychometric adequacy, but did not present data in a manner consistent with the norms for the measure ($k = 109$) or scored the measure in a manner that diverged from the manual ($k = 86$). Finally, an additional 30 studies were excluded because either inadequate levels of reliability ($k = 4$) or no reliability data ($k = 26$) were reported. Overall, almost one quarter of the articles assessed for eligibility had problematic measurement issues. This suggests that a large proportion of studies in this area suffer from inadequate measurement of key constructs and raises important concerns about the interpretation of previous findings. Of particular concern the large number of studies excluded due to failing to meet criteria with respect to measurement issues suggests that despite repeated calls to use measures with evidence of psychometric strength (e.g., Anderson et al., 2017; Blackmore et al., 2020; Bronstein & Montgomery, 2011;

Fellmeth et al., 2017; Falah-Hassani et al., 2015; Kien et al., 2019; Stevens & Vollebergh, 2008), the field remains replete with articles that have not addressed this issue. Of the 879 articles excluded for using a quantitative measure of adjustment or parenting without evidence of psychometric adequacy, over half ($k = 535$) were published in the past decade. Further, of studies that used a measure with evidence of psychometric strength but were excluded for not reporting reliability data for their sample, or reported inadequate levels of reliability, the majority ($k = 21$) were published in the past 10 years.

Our meta-analysis focused exclusively on studies that used a measure with evidence of reliability and validity in normative samples and reported adequate reliability in their immigrant/refugee samples. Further, the included studies were rated as methodologically strong. Given the scientific strength of the included studies, we have confidence in our findings. However, it is important to bear in mind not only the effect size, but also its confidence interval. The precision of the confidence intervals of summary effect sizes is impacted by both the sample sizes of the included studies (as smaller samples may display greater variability across ratings) and variability between studies (Borenstein et al., 2009). Summary effect sizes that yielded confidence intervals that included zero provide equivocal evidence of differences between reported adjustment problems and parenting constructs between immigrant/refugee samples and normative samples (Schünemann et al., 2020). Given the high degree of heterogeneity across nearly all summary effect sizes and confidence intervals that include zero for many effects, these results must be interpreted with caution (Borenstein et al., 2009; Schünemann et al., 2020).

This study adds to the existing literature on immigrants and refugees by examining adjustment among children, adolescents, and parents, as well as parenting. The overall results of adjustment replicated the pattern found in previous meta-analyses and systematic reviews that

used psychometrically sound measures (e.g., Bronstein & Montgomery, 2011; Falah-Hassani et al., 2015), and indicated that immigrant/refugee youth and parents report greater adjustment problems compared to normative data, although this varied across migrant groups, informants, and adjustment variables. The small number of studies of parenting measures yielded a pattern that was consistent between informants.

Overall, migrant youth described themselves as having more symptoms of anxiety than normative groups. Migrant youth also reported greater self-esteem, overall problems, and lower symptoms of depression and trauma than normative samples; however, the confidence intervals for these estimated summary effect sizes included zero. Parents of migrant youth rated their children as displaying more overall and internalizing problems, but fewer externalizing problems, than comparison groups. The single study that used teacher ratings yielded fewer overall problems among migrant youth compared to normative data, however, the confidence interval with respect to teacher-reported youth overall problems also included zero. Parent reports of their own adjustment showed more overall problems, anxiety, and depression than in normative samples. Both youth and parents reported fewer problematic parenting behaviors than normative samples; however, confidence intervals for parenting problems included zero for both parent and youth reports. Further, parents and their children reported more positive family functioning than normative comparison groups. Although moderator analyses might explain the variability in the studies across adjustment and parenting constructs, there were insufficient studies to pursue moderator analyses to explain this heterogeneity of effect sizes.

Examination of these variables separately for immigrant, refugee, and combined samples revealed some consistencies and some differences between migrant groups in relation to normative comparison groups. Both immigrant only and refugee only samples reported higher

anxiety than normative groups. Among immigrant youth self-reports, the confidence intervals for all other constructs included zero. Refugee youth also self-reported more depressive symptoms than normative samples; the confidence intervals for the estimates of self-esteem, overall problems, and trauma included zero. A combined sample of immigrant/refugee youth self-reported greater levels of trauma than do normative samples. Parent reports indicated greater overall problems among both immigrant and refugee samples than comparison groups; parents also reported greater internalizing, but lower externalizing, problems in a combined sample of immigrant/refugee youth compared to normative data. Finally, teacher-reports were limited to refugee youth, and indicated fewer overall problems compared to the normative sample; however, the confidence interval included zero.

Taken together, greater self-reported anxiety and parent-reported overall problems among both immigrant and refugee youth (compared to normative samples) emerged as robust findings across migrant groups. The confidence intervals for these estimates did not include zero and the studies were methodologically high quality, and both factors increase confidence in our results. The interpretation of the other adjustment constructs is limited by the availability of a single sample per construct, and/or because the confidence interval of the summary estimate included zero. These factors necessitate caution in the subsequent interpretation of results, as these findings provide equivocal evidence of differences in adjustment between immigrant or refugee youth adjustment and normative comparisons. Potential explanations are explored below, first across all informants of child adjustment among immigrant samples, then refugee, and combined groups.

In the current meta-analysis, results indicated that the only construct for which immigrant youth self-report more problems than do normative groups is in terms of anxiety, replicating past

reviews that evaluated anxiety among immigrant youth (e.g., Curtis et al., 2018). Parent reports among immigrant youth were limited to one study examining overall problems. Results of this study were consistent with past research (e.g., Dimitrova et al., 2016), indicating more problems among immigrant youth than normative comparisons. Confidence intervals for summary effect size estimates of all other self-report constructs included zero, and the summary estimates of overall problems and depression indicated significant variability across studies. Moderator analyses that may have helped explain this heterogeneity were precluded due to inadequate data. Results of self-reported overall problems and trauma symptoms were in the expected direction, indicating greater problems among immigrant youth than normative comparisons. Immigrant youth, however, also self-reported fewer symptoms of depression than do comparison samples. This contradicts the finding of increased depressive symptoms among immigrant youth reported by Curtis and colleagues (2018) who did not limit their analysis to studies that used measures with evidence of reliability and validity, and the measures of the included studies were not described. It is possible that our results differed due to the use of more stringent inclusion criteria for adjustment measures.

Refugee youth self-reports indicated greater symptoms of anxiety and depression than both immigrant and normative comparison samples, replicating findings from past reviews limited to studies using psychometrically sound measures (e.g., Blackmore et al., 2020; Bronstein & Montgomery, 2011; Kien et al., 2019). The current review included single studies that addressed self-reported self-esteem, self-reported trauma, and teacher-reported overall problems, and two studies of self-reported overall problems in refugee youth; for all these studies the confidence interval includes zero. Numerous reviews have reported higher rates of trauma symptoms among refugee youth than comparison groups (e.g., Tam et al., 2017), or greater

mental health problems than nonmigrants (e.g., Porter & Haslam, 2005). These reviews, however, employed less stringent inclusion criteria as studies were not limited to those that used measures with evidence of reliability and validity (Porter & Haslam, 2005; Tam et al., 2017), or measures were not described (Curtis et al., 2018). Further, our findings are discrepant from past reviews that used psychometrically strong measures to assess PTSD (e.g., Bronstein & Montgomery, 2011). However, studies in the review conducted by Bronstein and Montgomery (2011) that compared PTSD symptoms among refugee youth with a nonrefugee comparison group were limited to two studies, one of which focused on unaccompanied minors, the other included both accompanied and unaccompanied youth. Consistent with past literature, it is possible that these unaccompanied youth reported higher levels of PTSD symptoms than accompanied refugee youth (Kien et al., 2019). This may account for the discrepancy with the current meta-analysis, as we limited refugee samples to those accompanied by their parents. In sum, it is possible that the exclusive use of psychometrically strong measures of adjustment and retaining only refugee youth accompanied by their parents may account for differences between the findings in our meta-analysis and previous reviews. Additional potential explanations are explored below.

To further explore the surprising finding in the current meta-analysis that refugee youth reported fewer symptoms of trauma than normative comparisons (whereas immigrant and combined samples reported greater symptoms than comparison groups), it may help to consider the comparison group. The confidence interval with respect to trauma in refugee youth includes zero so does not replicate findings of greater symptoms of PTSD or trauma among refugee youth that also used psychometrically adequate measures (e.g., Blackmore et al., 2019; Bronstein & Montgomery, 2011; Kien et al., 2019), and these results must be interpreted with caution, as they

are limited to one sample. This finding merits further attention, however, as it raises an important question about comparison samples: in attempting to understand this apparent discrepancy, it is possible that the comparison sample available for the measures of child and adolescent trauma may have impacted these results. In this review, the one study that examined trauma symptoms in refugee youth used the Post-traumatic Stress Symptoms in Children (PTSS-C; Ahmad et al., 2000) measure, a tool that has norms based on a nonmigrant sample from a community with a history of violence. The study that used this measure reported lower levels of trauma symptoms among refugee samples compared to the normative sample. Examination of reported rates of trauma symptoms across all immigrant, refugee, and combined samples, however, suggests that the comparison group may be an important consideration: studies that used measures developed with youth recruited from regions without a history of violence (i.e., children exposed to a natural disaster: Child PTSD Symptom Scale; Foa et al., 2001; community sample in a region without a history of disruption: Reactions of Adolescents to Traumatic Stress measure; Bean et al., 2006) reported higher rates trauma among immigrant and combined immigrant/refugee samples. It is reasonable to infer that the comparison group for the PTSS-C was more comparable to the refugee sample included in this meta-analysis than the immigrant-only and combined immigrant/refugee samples that used measures developed with qualitatively different normative groups, and may help to explain this variability in results. Findings from this review are consistent with the interpretation proposed by Porter and Haslam (2005) that comparison groups who may have been exposed to high levels of conflict or violence may report adjustment challenges at rates comparable to refugee samples. This highlights the importance of considering the comparison group used when examining refugee adjustment.

Finally, the results of a single study of externalizing problems in a combined sample of

immigrant and refugee youth (Aroian et al., 2009) yielded a finding that was inconsistent with previous reviews: parent reports indicated fewer problems than normative comparisons, and the confidence interval did not include zero. Contrary to hypotheses, youth were rated as experiencing fewer externalizing (but more internalizing) problems than nonmigrant peers. This finding is inconsistent with a past meta-analysis of externalizing problems among immigrant youth (Dimitrova et al., 2016). This meta-analysis, however, did not describe the measures used to assess externalizing problems and the psychometric strength of the included measures is unknown. In addition, this meta-analysis did not present results separately by raters of adjustment, therefore it is unclear whether this may also have contributed to the discrepancy in results between this meta-analysis and the current study. Another potential explanation may include cultural expectations of behavior: in this study, all families were recruited from a culturally similar community in one city. It is possible that the presentation of externalizing behavior is discouraged within the context of community expectations of behavior, which may have contributed to lower reported rates of externalizing problems (Aroian et al., 2009; Dwairy, 2006).

Across parental adjustment, self-reports among immigrant and refugee parents confirmed the pattern found in earlier syntheses of immigrant and refugee adults (e.g., Fazel et al., 2005), and provide preliminary evidence based on a small number of fathers ($n = 198$) that the findings based on mothers also apply to fathers. When examined for each migrant group, results were fairly consistent with past reviews, however, the confidence interval of the estimate for higher overall problems in one sample of immigrant parents included zero and so does not replicate past reviews (e.g., Porter & Haslam, 2005). Although the majority of the findings indicated greater problems among immigrant and refugee samples, one exception consisted of a single study that

combined immigrant and refugee mothers (Aroian et al., 2009) indicated that migrant mothers reported fewer symptoms of depression than do the normative comparison group. This sample was recruited from a culturally similar community with strong community connections, and exploration of the sample's characteristics may help to understand this discrepancy with past reviews. Results suggested that mothers higher in social support reported significantly fewer depressive symptoms than mothers who reported lower support. Low levels of social support is a well-recognized risk factor in the development of mental health difficulty (Falah-Hassani et al., 2015; Fellmeth et al., 2017), and it is possible that the significant relationship between social support and depressive symptoms exerted a protective effect preventing mothers from developing higher rates of depressive symptoms. These results point to the benefits of assessing potential moderators of adjustment – both risk and protective – to contextualize results and attempt to clarify differences between samples.

Meta-analyses of parenting constructs yielded the same pattern of results when examining all samples together or when migrant groups were analysed separately. This pattern was inconsistent with our hypothesis that immigrant and refugee families would report greater levels of problems across parenting behavior and the parent-child relationship. Immigrant only and combined samples reported more positive family functioning than normative comparison groups. Interestingly, one of these studies also assessed youth adjustment using a psychometrically sound measure (Slonim-Nevo et al., 2009), and found that immigrant youth self-reported fewer overall problems than the normative sample. It is possible that in this sample lower rates of child adjustment problems were related to positive family functioning, consistent with past research examining family functioning and child behavior (e.g., Biglan, Flay, Embry, & Sandler, 2012). Finally, refugee youth reported fewer problematic parenting behaviors than

do normative comparisons (Mellor et al., 2012); the confidence interval for parent reports included zero, providing equivocal evidence of differences between refugee and normative samples. It is also important to note that none of the studies included in the parenting meta-analysis also included validated measures of parental adjustment, signifying a gap the examining the interrelationship between child and parental adjustment and parenting in these diverse families.

Taken together, some of our results based on strong methodological studies were consistent with previous reviews in finding that immigrant and refugee youth reported higher anxiety than normative comparison groups; in addition, our data replicate the findings with respect to depressive symptoms among refugee youth. Further, parent reports indicated greater overall problems in both immigrant and refugee youth compared to normative samples, as well as greater internalizing problems in a combined sample of immigrant/refugee youth as compared to nonmigrant peers. We also found differences in terms of self-reported depression symptoms, as refugee youth reported greater symptoms than normative data, whereas immigrant youth reported fewer symptoms than comparison groups. Further, immigrant only and combined immigrant/refugee samples reported greater trauma symptoms than normative samples, however, refugee youth reported less trauma symptoms than normative data. Parent self-reports of their adjustment were mainly consistent with past reviews, and our investigation of parenting behavior and the parent-child relationship yielded very few studies that used psychometrically sound measures.

All of the included studies used psychometrically strong measures and were rated as high in methodological quality. Some of our results that were inconsistent with past reviews (lower self-reported depression among immigrant youth, higher self-reported self-esteem and lower

trauma symptoms, and lower teacher-reported overall problems among refugee youth) may have differed in the psychometric strength of the measures used to assess these constructs. For instance, past reviews including depression among immigrant youth (e.g., Curtis et al., 2018), and among refugee youth examining self-esteem (Marley & Mauki, 2019), trauma (Curtis et al., 2018; Tam et al., 2017), and overall distress (e.g., Porter & Haslam, 2005) did not restrict their studies to those that used measures with evidence and reliability and validity (Fazel et al., 2012; Marley & Mauki, 2019; Porter & Haslam, 2005; Reed et al., 2012; Tam et al., 2017; van Os et al., 2016), or measures were not described (e.g., Curtis et al., 2018; Dimitrova et al., 2016). It is possible that the inclusion of psychometrically weak or untested measures could account for some of these differences in results.

Limitations, Strengths and Future Research

This systematic review and meta-analysis paints a troubling picture of the literature on adjustment and parenting among immigrant and refugee families. Many potentially eligible studies were excluded on review of the full text because measures were often unidentified, did not have evidence of adequate reliability and validity, or reported inadequate reliability. In addition, many studies were excluded because there was insufficient information in the article to confirm eligibility (e.g., did not report generational status, number of years in host country, or whether participating adults were parents), or because data in the primary studies were collapsed across eligible and ineligible groups (e.g., newcomers and nonnewcomers) and data for the eligible group could not be extracted. Only 31 studies met inclusion criteria indicating a small number of scientifically sound studies with sound measurement of the constructs of interest. The results must be considered preliminary due to small number of studies (i.e., parenting behavior, parent-child relationship; individual measures of adjustment; teacher reports of child

adjustment) that met our inclusion criteria.

Given the significant variability across included studies, it would have been ideal to explore potential explanations for this heterogeneity using moderator analyses. However, the sample of studies did not permit us to conduct moderator analyses with adequate power, limiting the ability to examine variability in effect sizes across the included studies. Authors of past meta-analyses have presented such analyses with fewer available studies. For example, Blackmore and colleagues' (2020) meta-analysis of eight studies of refugee youth presented subgroup analyses to examine variability in prevalence rates of mental disorders. Subgroup analyses were presented for each diagnostic category and included from four to seven studies. Despite having more available studies of immigrant or refugee samples, we decided not to pursue moderator analyses as it is generally not recommended to pursue such analyses with fewer than 10 samples to achieve adequate power to strengthen the interpretation of results (Borenstein et al., 2009).

In addition, the use of normative comparison groups of each measure of adjustment or parenting enabled us to be more inclusive with study methodology (i.e., we were not limited to studies that included a nonmigrant comparison group), but it also raised significant challenges: this required careful confirmation that each included study used the full measure in a manner consistent with its manual or published norms to calculate effect sizes. Many studies were excluded ($k = 86$) because the measure's items and/or scoring had been modified or data were presented in a way that was inconsistent with the published norms. In addition, accessing appropriate norms developed with nonmigrants – particularly for trauma measures – was challenging: for several measures, we were only able to access samples that were exposed to civil war or violence, which may have impacted the results of this meta-analysis. Further, very

few studies were located that assessed parenting with psychometrically adequate measures among immigrant and refugee families, reinforcing a gap identified in previous systematic reviews (van Os et al., 2016). Finally, included studies rarely assessed both child or parental adjustment and parenting, making it impossible to determine whether adjustment accounted for differences in parenting.

This meta-analysis is the first to include both immigrant and refugee youth and their parents. It is also the first to incorporate both adjustment and parenting constructs and adds to the existing literature of these potentially vulnerable families. As international migration continues to grow (United Nations High Commissioner for Refugees, 2020) and research about immigrant and refugee experiences expands, it is essential not only to regularly integrate new data to determine if early findings still apply or identify emerging trends, it is also crucial to continue reviewing findings to inform the development of appropriate mental health and parenting supports that may be helpful. This meta-analysis supports these two aims: first, by examining reports of child adjustment and parenting separately by raters, important differences between raters emerged, reinforcing the importance of considering who is reporting. Although this is consistent with the general literature on child adjustment (Achenbach et al., 1987; De Los Reyes et al., 2015) and parenting behavior (Hou et al., 2020), it is an important reminder to consider the informant when examining the experiences of immigrant and refugee families. Past reviews highlighted challenges in drawing clear conclusions between immigrant and refugee youth and attributed this to variability across raters (e.g., Stevens & Vollebergh, 2008) or did not differentiate between informants (e.g., Kien et al., 2019). Moving forward, presenting results for individual informants may help clarify the adjustment experiences among immigrant and refugee youth, as parents' and teachers' report may provide a different picture than youth self-reports.

Second, in our consideration of adjustment and parenting separately for immigrant, refugee, and combined samples, important differences between migrant groups emerged. As described in Tables [1](#) and [2](#), past systematic reviews and meta-analyses of child adjustment have tended to focus exclusively on either immigrant (e.g., Dimitrova et al., 2016) or refugee (e.g., Fazel et al., 2012) youth. A large systematic review with both immigrant and refugee youth (Curtis et al., 2018) described their adjustment experiences separately; however, lack of quantitative comparisons limited our understanding of the strength of the differences between groups. Similarly, reviews and meta-analyses of adult adjustment incorporating both immigrant and refugee samples (Anderson et al., 2017; Fellmeth et al., 2017) did not present results separately for each migrant group, or included only immigrant (e.g., Falah-Hassani et al., 2015) or refugee adults (e.g., Fazel et al., 2005). Including both immigrant and refugee samples, and systematically presenting results separately in future reviews could help clarify the variability in findings identified in past reviews.

Third, by focusing our meta-analysis on studies that used measures with evidence of psychometric adequacy, we aimed to increase our confidence in results. Many studies were excluded, however, because they did not report on the reliability of measures used with immigrant and refugee samples. It is crucial that reliability data be reported for each sample as psychometric adequacy achieved with the normative sample is not a guarantee of adequacy in subsequent samples (Hunsley & Mash, 2018). In reviewing the characteristics of studies excluded for not reporting reliability data ($k = 26$) or those that reported inadequate reliability within their sample ($k = 4$), we found that more refugee ($k = 17$) than immigrant ($k = 10$) or combined ($k = 3$) samples were excluded. A variety of regions of origin and host countries were represented across these excluded studies, however, most of these regions and countries

continued to be represented in the final meta-analysis. Thus, although our overall sample of included studies was reduced and there are samples which may not be as robustly represented, no particular groups appear to have been systematically excluded. Finally, in using the published norms of measures of adjustment and parenting meant that we were not limited to studies that included their own comparison groups, enabling greater inclusivity of available studies of immigrant and refugee families.

Future meta-analyses will benefit if authors of primary studies consistently present data on sample characteristics and basic data of the measures used, presenting data separately for age and/or gender groups to facilitate their inclusion in future meta-analyses. Further, in studies of adult immigrants and refugees, it would be helpful to report the proportion of the sample who are parents (and if possible to present basic data on this subsample separately) to help inform our understanding of this population, given that parental adjustment, child adjustment, and parenting are interconnected (Barlow et al., 2014; Ryan & Ollendick, 2018). Second, given the variability across informants of child adjustment identified in past reviews (e.g., Stevens & Vollebergh, 2008) and in the current meta-analysis, it is essential to clearly state in primary studies and reviews the informant for each measure, and the specific measures that are used. Future research would also benefit from selecting measures with evidence of psychometric adequacy and to report on the psychometric data within their samples to increase confidence in the methodological rigor of the study and its results.

The results of this meta-analysis contribute not only to an updated review of available literature on immigrant and refugee children, adolescents, and adults, it also included numerous studies that had not previously been captured in recent reviews. In addition, this meta-analysis incorporated a family perspective by examining adjustment and parenting from the perspective

of both children and their parents. Overall, the findings indicate that immigrant and refugee families are more likely to experience heightened distress compared to nonmigrants, reaffirming the need to ensure that services for families address their needs.

References

* *Indicates study was included in the meta-analysis*

Abidin, R. R. (2012). *Parenting Stress Index* (4th ed.). Psychological Assessment Resources.

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4–18 and 1991 profile*.

University of Vermont.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232. <https://doi.org/10.1037/0033-2909>

Ahmad, A., Sundelin-Wahlsten, V., Sofi, M. A., Qahar, J. A., & von Knorring, A.-L. (2000). Reliability and validity of a child-specific crosscultural instrument for assessing posttraumatic stress disorder. *European Child & Adolescent Psychiatry*, 9, 285-294.

<https://doi.org/10.1007/s007870070032>

* Ahmad, A., von Knorring, A. L., & Sundelin-Wahlsten, V. (2008). Traumatic experiences and post-traumatic stress disorder in Kurdistanian children and their parents in homeland and exile: An epidemiological approach. *Nordic Journal of Psychiatry*, 62, 457-463.

<https://doi.org/10.1080/08039480801984305>

* Alhasanat-Khalil, D., Fry-McComish, J., Dayton, C., Benkert, R., Yarandi, H., & Giurgescu, C. (2018). Acculturative stress and lack of social support predict postpartum depression among U.S. immigrant women of Arabic descent. *Archives of Psychiatric Nursing*, 32, 530-535.

<https://doi.org/10.1016/j.apnu.2018.02.005>

Anderson, F. M., Hatch, S. L., Comacchio, C., & Howard, L. M. (2017). Prevalence and risk of mental disorders in the perinatal period among migrant women: A systematic review and meta-analysis. *Archives of Women's Mental Health*, 20, 449-462.

<https://doi.org/10.1007/s00737-017-0723-z>

- * Aroian, K., Hough, E. S., Templin, T. N., Kulwicki, A., Ramaswamy, V., & Katz, A. (2009). A model of mother-child adjustment in Arab Muslim immigrants to the US. *Social Science & Medicine*, 69, 1377-1386. <https://doi.org/10.1016/j.socscimed.2009.08.027>
- Barlow, J., Smailagic, N., Huband, N., Roloff, V., & Bennett, C. (2014). Group-based parent training programmes for improving parental psychosocial health. *Cochrane Database of Systematic Reviews*, CD002020. <https://doi.org/10.1002/14651858.CD002020.pub4>
- Bean, T., Derluyn, I., Eurelings-Bontekoe, E., Broekaert, E., & Spinhoven, P. (2006). Validation of the multiple language versions of the Reactions of Adolescents to Traumatic Stress questionnaire. *Journal of Traumatic Stress*, 19, 241-255. <https://doi.org/10.1002/jts.20093>
- Bean, T., Derluyn, I., Eurelings-Bontekoe, E., Broekaert, E., & Spinhoven, P. (2007). Validation of the multiple language versions of the Hopkins Symptom Checklist-37 for refugee adolescents. *Adolescence*, 42, 51-71.
- * Bean, T., Derluyn, I., Eurelings-Bontekoe, E., Broekaert, E., & Spinhoven, P. (2007). Comparing psychological distress, traumatic stress reactions, and experiences of unaccompanied refugee minors with experiences of adolescents accompanied by parents. *The Journal of Nervous and Mental Disease*, 195, 288-297. <https://doi.org/10.1097/01.nmd.0000243751.49499.93>
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory*, 2nd Edition. PsychCorp.
- * Beeber, L. S., Holditch-Davis, D., Perreira, K., Schwartz, T. A., Lewis, V., Blanchard, H., Canuso, R., & Goldman, B. D. (2010). Short-term in-home intervention reduces depressive symptoms in Early Head Start Latina mothers of infants and toddlers. *Research in Nursing*

& *Health*, 33, 60-76. <https://doi.org/10.1002/nur.20363>

Beidas, R. S., Stewart, R. E., Walsh, L., Lucas, S., Downey, M. M., Jackson, K., Fernandez, T., & Mandell, D. S. (2015). Free, brief, and validated: Standardized instruments for low-resource mental health settings. *Cognitive and Behavioral Practice*, 22, 5-19.

<https://doi.org/10.1016/j.cbpra.2014.02.002>

Beiser, M. (2010). The mental health of immigrant and refugee children in Canada: A description and selected findings from the New Canadian Children and Youth Study (NCCYS). *Canadian Issues, Summer*, 103–107.

Bellefontaine, S. P., & Lee, C. M. (2014). Between black and white: Examining grey literature in meta-analyses of psychological research. *Journal of Child and Family Studies*, 23, 1378–1388. <http://dx.doi.org/10.1007/s10826-013-9795-1>

Blackmore, R., Gray, K. M., Boyle, J. A., Fazel, M., Ranasinha, S., Fitzgerald, G., Misso, M., & Gibson-Helm, M. (2020). Systematic review and meta-analysis: The prevalence of mental illness in child and adolescent refugees and asylum seekers. *Journal of the American Academy of Child and Adolescent Psychiatry*, 59, 705-714.

<https://doi.org/10.1016/j.jaac.2019.11.011>

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to Meta-Analysis*. John Wiley & Sons Ltd.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2014). *Comprehensive Meta-Analysis Version 3*. Biostat. <https://www.meta-analysis.com/>

Bronstein, I., & Montgomery, P. (2011). Psychological distress in refugee children: A systematic review. *Clinical Child and Family Psychology Review*, 14, 44-56.

<https://doi.org/10.1007/s10567-010-0081-0>

- Byrne, S. P., Lebowitz, E. R., Ollendick, T. H., & Silverman, W. K. (2018). Anxiety disorders in children and adolescents. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (2nd ed., pp. 217-241). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190492243.003.0011>
- * Cho, Y. B., & Haslam, N. (2010). Suicidal ideation and distress among immigrant adolescents: the role of acculturation, life stress, and social support. *Journal of Youth and Adolescence*, 39, 370-379. <https://doi.org/10.1007/s10964-009-9415-y>
- * Cleary, S. D., Snead, R., Dietz-Chavez, D., Rivera, I., & Edberg, M. C. (2018). Immigrant trauma and mental health outcomes among Latino youth. *Journal of Immigrant and Minority Health*, 20, 1053-1059. <https://doi.org/10.1007/s10903-017-0673-6>
- * Costigan, C. L., & Koryzma, C. M. (2011). Acculturation and adjustment among immigrant Chinese parents: Mediating role of parenting efficacy. *Journal of Counseling Psychology*, 58, 183-196. <https://doi.org/10.1037/a0021696>
- Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry*, 150, 782-786. <https://doi.org/10.1192/bjp.150.6.782>
- Cunningham, A., & Scott, J. (2012). Child and adolescent assessment strategy and inventories. In L. Sperry (Ed.), *Family assessment: Contemporary and cutting-edge strategies* (2nd ed., pp. 163-202). Taylor & Francis.
- Cunningham, S. A., Ruben, J. D., & Narayan, K. M. (2008). Health of foreign-born people in the United States: A review. *Health Place*, 14, 623-635. <https://doi.org/10.1016/j.healthplace.2007.12.002>
- Curtis, P., Thompson, J., & Fairbrother, H. (2018). Migrant children within Europe: A systematic

review of children's perspectives on their health experiences. *Public Health*, 158, 71-85.

<https://doi.org/10.1016/j.puhe.2018.01.038>

- * Cyril, S., Halliday, J., Green, J., & Renzaho, A. M. (2016). Relationship between body mass index and family functioning, family communication, family type and parenting style among African migrant parents and children in Victoria, Australia: A parent-child dyad study. *BMC Public Health*, 15, 707. <https://doi.org/10.1186/s12889-016-3394-1>

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, 141, 858–900.

<https://doi.org/10.1037/a0038498>

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, 131, 483-509.

<https://doi.org/10.1037/0033-2909.131.4.483>

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, 9, 123-149. <https://doi.org/10.1146/annurev-clinpsy-050212-185617>

De Maio, F. G., & Kemp, E. (2010). The deterioration of health status among immigrants to Canada. *Global Public Health*, 5, 462-478. <https://doi.org/10.1080/17441690902942480>

Derluyn, I., Mels, C., & Broekaert, E. (2009). Mental health problems in separated refugee adolescents. *Journal of Adolescent Health*, 44, 291-297.

<https://doi.org/10.1016/j.jadohealth.2008.07.016>

- Derogatis, L. R. (1993). *Brief Symptom Inventory: Administration, scoring, and procedures manual*. PsychCorp.
- Derogatis, L. R. (1994). *Symptom Checklist-90-R: Administration, scoring, and procedures manual*. PsychCorp.
- * Dimitrova, R., & Chasiotis, A. (2012). Are immigrant children in Italy better adjusted than mainstream Italian children? In C. Garcia Coll (Ed.), *The impact of immigration on children's development: Contributions to human development* (Vol. 24, pp. 35-48). Karger. <https://doi.org/10.1159/000331024>
- Dimitrova, R., Chasiotis, A., & van de Vijver, F. (2016). Adjustment outcomes of immigrant children and youth in Europe: A meta-analysis. *European Psychologist*, 21, 150–162. <https://doi.org/10.1027/1016-9040/a000246>
- Dougherty, L. R., Klein, D. N., & Olino, T. M. (2018). Depression in children and adolescents. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (2nd ed., pp. 99-130). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190492243.003.0006>
- Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*, 52, 377–384. <https://doi.org/10.1136/jech.52.6.377>
- Dretzke, J., Davenport, C., Frew, E., Barlow, J., Stewart-Brown, S., Bayliss, S., Taylor, R. S., Sandercock, J., & Hyde, C. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: A systematic review of randomised controlled trials. *Child and Adolescent Psychiatry and Mental Health*, 3, 7. <https://doi.org/10.1186/1753-2000-3-7>

- Dugas, M. J., Charette, C. A., & Gervais, N. J. (2018). Generalized anxiety disorder. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (pp. 293-310). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190492243.003.0014>
- Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster Family Assessment Device. *Journal of Marital and Family Therapy*, 9, 171-180.
<https://doi.org/10.1111/j.1752-0606.1983.tb01497.x>
- * Eruyar, S., Maltby, J., & Vostanis, P. (2018). Mental health problems of Syrian refugee children: The role of parental factors. *European Child & Adolescent Psychiatry*, 27, 401-409. <https://doi.org/10.1007/s00787-017-1101-0>
- European Migration Network. (2018). *Asylum and migration glossary: A tool for better comparability produced by the European Migration Network (May 2018)*.
https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/networks/european_migration_network/docs/interactive_glossary_6.0_final_version.pdf
- Falah-Hassani, K., Shiri, R., Vigod, S., & Dennis, C. L. (2015). Prevalence of postpartum depression among immigrant women: A systematic review and meta-analysis. *Journal of Psychiatric Research*, 70, 67-82. <https://doi.org/10.1016/j.jpsychires.2015.08.010>
- Fazel, M., Reed, R. V., Panter-Brick, C., & Stein, A. (2012). Mental health of displaced and refugee children resettled in high-income countries: Risk and protective factors. *The Lancet*, 379, 266-282. [https://doi.org/10.1016/s0140-6736\(11\)60051-2](https://doi.org/10.1016/s0140-6736(11)60051-2)
- Fazel, M., Wheeler, J., & Danesh, J. (2005). Prevalence of serious mental disorder in 7000 refugees resettled in western countries: A systematic review. *The Lancet*, 365, 1309-1314.
[https://doi.org/10.1016/s0140-6736\(05\)61027-6](https://doi.org/10.1016/s0140-6736(05)61027-6)
- Fellmeth, G., Fazel, M., & Plugge, E. (2017). Migration and perinatal mental health in women

- from low- and middle-income countries: A systematic review and meta-analysis. *British Journal of Obstetrics and Gynaecology*, 124, 742-752. <https://doi.org/10.1111/1471-0528.14184>
- Fendrich, M., Weissman, M. M., & Warner, V. (1990). Screening for depressive disorder in children and adolescents: Validating the Center for Epidemiologic Studies Depression Scale for Children. *American Journal of Epidemiology*, 131, 538-551. <https://doi.org/10.1093/oxfordjournals.aje.a115529>
- Foa, E. B., Johnson, K. M., Feeny, N. C., & Treadwell, K. R. H. (2001). The Child PTSD Symptom Scale: A preliminary examination of its psychometric properties. *Journal of Clinical Child Psychology*, 30, 376-384. https://doi.org/10.1207/S15374424JCCP3003_9
- Frick, P. J., Christian, R. E., & Wootton, J. M. (1999). Age trends in the association between parenting practices and conduct problems. *Behavior Modification*, 23, 106-128. <https://doi.org/10.1177/0145445599231005>
- Frick, P. J., & McMahon, R. J. (2018). Child and adolescent conduct problems. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (pp. 71-96). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190492243.003.0005>
- Gadeberg, A. K., Montgomery, E., Frederiksen, H. W., & Norredam, M. (2017). Assessing trauma and mental health in refugee children and youth: A systematic review of validated screening and measurement tools. *The European Journal of Public Health*, 27, 439-446. <https://doi.org/10.1093/eurpub/ckx034>
- * Giordano, F., Cipolla, A., Ragnoli, F., & Brajda Bruno, F. (2019). Transit migration and trauma: The detrimental effect of interpersonal trauma on Syrian children in transit in Italy. *Psychological Injury and Law*, 12, 76-87. <https://doi.org/10.1007/s12207-019-09345-x>

- Glaesmer, H., Braehler, E., Grande, G., Hinz, A., Petermann, F., & Romppel, M. (2014). The German version of the Hopkins Symptoms Checklist-25 (HSCL-25): Factorial structure, psychometric properties, and population-based norms. *Comprehensive Psychiatry*, 55, 396–403. <https://doi.org/10.1016/j.comppsy.2013.08.020>
- Gokhale, P. A. (1997, November 13–14). *Grey literature varieties—Definitional problems* [Conference proceedings]. Third International Conference on Grey Literature: Perspectives on the Design and Transfer of Scientific and Technical Information. Luxembourg, LU. <http://hdl.handle.net/10068/697961>
- Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40, 1337–1345. <https://doi.org/10.1097/00004583-200111000-00015>
- Green, K., Worden, B., Menges, D., & McCrady, B. (2008). Alcohol use disorders. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (pp. 339 - 369). Oxford University Press. <https://doi.org/10.1093/med:psych/9780195310641.001.0001>
- Guha-Sapir, D., Schlüter, B., Rodriguez-Llanes, J. M., Lillywhite, L., & Hicks, M. H.-R. (2018). Patterns of civilian and child deaths due to war-related violence in Syria: A comparative analysis from the Violation Documentation Center dataset, 2011–16. *The Lancet Global Health*, 6, e103-e110. [https://doi.org/10.1016/s2214-109x\(17\)30469-2](https://doi.org/10.1016/s2214-109x(17)30469-2)
- Haeny, A. M., Boness, C. L., McDowell, Y. E., & Sher, K. J. (2018). Alcohol use disorder. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (2nd ed., pp. 381-411). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190492243.003.0018>
- * Hamdan-Mansour, A. M., Abdel Razeq, N. M., AbdulHaq, B., Arabiat, D., & Khalil, A. A. (2017). Displaced Syrian children's reported physical and mental wellbeing. *Child and*

- Adolescent Mental Health*, 22, 186-193. <https://doi.org/10.1111/camh.12237>
- Hamilton, E., & Carr, A. (2016). Systematic review of self-report family assessment measures. *Family Process*, 55, 16-30. <https://doi.org/10.1111/famp.12200>
- Hampson, R. B., & Beavers, W. R. (2012). Observational assessment. In L. Sperry (Ed.), *Family assessment: Contemporary and cutting-edge strategies* (2nd ed., pp. 83–114). Taylor & Francis.
- Harter, S. (2012). *Self-Perception Profile for Adolescents: Manual and questionnaires*. University of Denver.
- Haslam, D. M., & Mejia, A. (2018). Accommodating race and ethnicity in parenting interventions. In M. R. Sanders & T. G. Mazzucchelli (Eds.), *The power of positive parenting: Transforming the lives of children, parents, and communities using the Triple P system* (pp. 332–343). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190629069.003.0030>
- Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9, 426-445. <https://doi.org/10.1037/1082-989X.9.4.426>
- Hempel, S., Suttorp, M. J., Miles, J. N. V., Wang, Z., Maglione, M., Morton, S., Johnsen, B., Valentine, D., & Shekelle, P. G. (2011). *Empirical evidence of associations between trial quality and effect size*. <http://effectivehealthcare.ahrq.gov>
- Hou, Y., Benner, A. D., Kim, S. Y., Chen, S., Spitz, S., Shi, Y., & Beretvas, T. (2020). Discordance in parents' and adolescents' reports of parenting: A meta-analysis and qualitative review. *American Psychologist*, 75, 329-348. <https://doi.org/10.1037/amp0000463>
- Hunsley, J., & Mash, E. J. (2018). 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* (2nd ed., pp. 3-16). Oxford University Press.

<https://doi.org/10.1093/med-psych/9780190492243.003.0001>

Hurley, K. D., Huscroft-D'Angelo, J., Trout, A., Griffith, A., & Epstein, M. (2014). Assessing parenting skills and attitudes: A review of the psychometrics of parenting measures.

Journal of Child and Family Studies, 23, 812-823. <https://doi.org/10.1007/s10826-013-9733-2>

* Jaycox, L. H., Stein, B. D., Kataoka, S. H., Wong, M., Fink, A., Escudero, P., & Zaragoza, C. (2002). Violence exposure, posttraumatic stress disorder, and depressive symptoms among recent immigrant schoolchildren. *Journal of the American Academy of Child & Adolescent Psychiatry*, 41, 1104-1110. <https://doi.org/10.1097/01.CHI.0000020261.43550.30>

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. <https://doi.org/10.1007/s10802-007-9201-9>

Keane, T. M., Silberbogen, A. K., & Weierich, M. R. (2008). Post-traumatic stress disorder. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (pp. 293-318). Oxford University Press. <https://doi.org/10.1093/med:psych/9780195310641.003.0014>

Kien, C., Sommer, I., Faustmann, A., Gibson, L., Schneider, M., Krczal, E., Jank, R., Klerings, I., Szelag, M., Kerschner, B., Brattstrom, P., & Gartlehner, G. (2019). Prevalence of mental disorders in young refugees and asylum seekers in European Countries: A systematic review. *European Child & Adolescent Psychiatry*, 28, 1295-1310.

<https://doi.org/10.1007/s00787-018-1215-z>

* Kirchner, T., Patiño, C., & Forns, M. (2011). Children left behind during immigration:

Repercussions on the mental health of Latin-American mothers and fathers. *Anuario de Psicología*, 41, 107-122.

- * Kovacev, L., & Shute, R. (2004). Acculturation and social support in relation to psychosocial adjustment of adolescent refugees resettled in Australia. *International Journal of Behavioral Development*, 28, 259-267. <https://doi.org/10.1080/01650250344000497>

Kovacs, M. (1985). The Children's Depression Inventory (CDI). *Psychopharmacology Bulletin*, 21, 995-998.

Kroenke, K. S., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16, 606-613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>

- * Lecompte, V., & Rousseau, C. (2018). Determinants of child attachment in the years postpartum in a high-risk sample of immigrant women. *Journal of Immigrant and Minority Health*, 20, 1166-1172. <https://doi.org/10.1007/s10903-017-0662-9>

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behavior Research and Therapy*, 33, 335-343. [https://doi.org/10.1016/0005-7967\(94\)00075-u](https://doi.org/10.1016/0005-7967(94)00075-u)

Marley, C., & Mauki, B. (2019). Resilience and protective factors among refugee children post-migration to high-income countries: A systematic review. *The European Journal of Public Health*, 29, 706-713. <https://doi.org/10.1093/eurpub/cky232>

- * Mellor, D., Renzaho, A., Swinburn, B., Green, J., & Richardson, B. (2012). Aspects of parenting and family functioning associated with obesity in adolescent refugees and migrants from African backgrounds living in Australia. *Australian and New Zealand*

- Journal of Public Health*, 36, 317-324. <https://doi.org/10.1111/j.1753-6405.2012.00894.x>
- Merikangas, K. R., He, J.-P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication–Adolescent supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, 49, 980 - 989. <https://doi.org/10.1016/j.jaac.2010.05.017>
- Michelson, D., Davenport, C., Dretzke, J., Barlow, J., & Day, C. (2013). Do evidence-based interventions work when tested in the "real world?" A systematic review and meta-analysis of parent management training for the treatment of child disruptive behavior. *Clinical Child and Family Psychology Review*, 16, 18-34. <https://doi.org/10.1007/s10567-013-0128-0>
- Millner, A. J., & Nock, M. K. (2018). Self-injurious thoughts and behaviors. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (2nd ed., pp. 193-214). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190492243.003.0010>
- * Mohammad, K. I., Abu Awad, D., Creedy, D. K., & Gamble, J. (2018). Postpartum depression symptoms among Syrian refugee women living in Jordan. *Research in Nursing & Health*, 41, 519-524. <https://doi.org/10.1002/nur.21919>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Journal of Clinical Epidemiology*, 62, 1006-1012. <https://doi.org/10.1016/j.jclinepi.2009.06.005>
- Mollica, R. F., Brooks, R., Tor, S., Lopes-Cardozo, B., & Silove, D. (2014). The enduring mental health impact of mass violence: A community comparison study of Cambodian civilians living in Cambodia and Thailand. *International Journal of Social Psychiatry*, 60,

6-20. <https://doi.org/10.1177/0020764012471597>

- Mollica, R. F., Caspi-Yavin, Y., Bollini, P., Truong, T., Tor, S., & Lavelle, J. (1992). The Harvard Trauma Questionnaire. Validating a cross-cultural instrument for measuring torture, trauma, and posttraumatic stress disorder in Indochinese refugees. *Journal of Nervous and Mental Disease*, 180, 111-116. <https://doi.org/10.1097/00005053-199202000-00008>
- Mollica, R. F., Wyshak, G., de Marneffe, D., Khuon, F., & Lavelle, J. (1987). Indochinese versions of the Hopkins Symptom Checklist-25: A screening instrument for the psychiatric care of refugees. *American Journal of Psychiatry*, 144, 497-500. <https://doi.org/10.1176/ajp.144.4.497>
- Moshier, S. J., Parker-Guilbert, K. S., Marx, B. P., & Keane, T. M. (2018). Post-traumatic stress disorder in adults. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (2nd ed., pp. 329-356). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190492243.003.0016>
- Mugadza, H. T., Mujeyi, B., Stout, B., Wali, N., & Renzaho, A. M. N. (2019). Childrearing practices among sub-saharan African migrants in Australia: A systematic review. *Journal of Child and Family Studies*, 28, 2927-2941. <https://doi.org/10.1007/s10826-019-01463-z>
- * Muller, L. R. F., Buter, K. P., Rosner, R., & Unterhitzenberger, J. (2019). Mental health and associated stress factors in accompanied and unaccompanied refugee minors resettled in Germany: A cross-sectional study. *Child and Adolescent Psychiatry and Mental Health*, 13, 8. <https://doi.org/10.1186/s13034-019-0268-1>
- * Nakash, O., Nagar, M., & Lurie, I. (2016). The association between postnatal depression, acculturation and mother-infant bond among Eritrean asylum seekers in Israel. *Journal of*

Immigrant and Minority Health, 18, 1232-1236. <https://doi.org/10.1007/s10903-016-0348-8>

Newbold, K. B. (2005). Self-rated health within the Canadian immigrant population: Risk and the healthy immigrant effect. *Social Science & Medicine*, 60, 1359-1370.

<https://doi.org/10.1016/j.socscimed.2004.06.048>

Ng, E., Wilkins, R., Gendron, F., & Berthelot, J.-M. (2005). Healthy today, healthy tomorrow? Dynamics of immigrants' health in Canada: Evidence from the National Population Health Survey. *Statistics Canada, Catalogue 82-618-MWE2005002*.

Olafson, E., & Connelly, L. (2012). Child abuse assessment strategy and inventories. In L. Sperry (Ed.), *Family assessment: Contemporary and cutting-edge strategies* (2nd ed., pp. 265–308). Taylor & Francis.

Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, 10, 233–240. <https://10.1111/jebm.12266>

Pappas, C., & Williams, I. (2011). Grey literature: Its emerging importance. *Journal of Hospital Librarianship*, 11, 228–234. <https://10.1080/15323269.2011.587100>

Perruchoud, R., & Redpath-Cross, J. (2011). Glossary on migration, 2nd edition. *International Migration Law*, 25. International Organization for Migration.

Persons, J. B., & Fresco, D. M. (2008). Adult depression. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (pp. 96-120). Oxford University Press.

<https://doi.org/10.1093/med:psych/9780195310641.003.0005>

Persons, J. B., Fresco, D. M., & Ernst, J. S. (2018). Adult depression. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (2nd ed., pp. 131-151). Oxford University Press.

<https://doi.org/10.1093/med-psych/9780190492243.003.0007>

- Porter, M., & Haslam, N. (2005). Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: A meta-analysis. *Journal of the American Medical Association*, 294, 602 - 612. <https://doi.org/10.1001/jama.294.5.602>
- * Quinlan, R., Schweitzer, R. D., Khawaja, N., & Griffin, J. (2016). Evaluation of a school-based creative arts therapy program for adolescents from refugee backgrounds. *The Arts in Psychotherapy*, 47, 72-78. <https://doi.org/10.1016/j.aip.2015.09.006>
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385-401.
<https://doi.org/10.1177/014662167700100306>
- Reed, R. V., Fazel, M., Jones, L., Panter-Brick, C., & Stein, A. (2012). Mental health of displaced and refugee children resettled in low-income and middle-income countries: Risk and protective factors. *The Lancet*, 379, 250-265. [https://doi.org/10.1016/s0140-6736\(11\)60050-0](https://doi.org/10.1016/s0140-6736(11)60050-0)
- Rohsenow, D. J. (2018). Substance use disorders. In J. Hunsley & E. J. Mash (Eds.), *A guide to assessments that work* (2nd ed., pp. 359-380). Oxford University Press.
<https://doi.org/10.1093/med-psych/9780190492243.003.0017>
- Ryan, S. M., & Ollendick, T. H. (2018). The interaction between child behavioral inhibition and parenting behaviors: Effects on internalizing and externalizing symptomology. *Clinical Child and Family Psychology Review*, 21, 320-339. <https://doi.org/10.1007/s10567-018-0254-9>
- Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014). The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review*, 34, 337-357.

<https://doi.org/10.1016/j.cpr.2014.04.003>

Sanderson, S., Tatt, I. D., & Higgins, J. P. (2007). Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: A systematic review and annotated bibliography. *International Journal of Epidemiology*, 36, 666-676.

<https://doi.org/10.1093/ije/dym018>

Schludermann, S. & Schludermann, E. (1988). *CRPBI: Children's Report on Parent Behavior Inventory*. University of Manitoba.

Schmidt, F. L., Oh, I. S., & Hayes, T. L. (2009). Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. *British Journal of Mathematical and Statistical Psychology*, 62, 97-128.

<https://doi.org/10.1348/000711007X255327>

Schünemann, H. J., Vist, G. E., Higgins, J. P. T., Santesso, N., Deeks, J. J., Glasziou, P., Akl, E. A., Guyatt, G. H. (2020). Interpreting results and drawing conclusions. In J. P. T Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Ed), *Cochrane Handbook for Systematic Reviews of Interventions version 6.1* (updated September 2020). Available from <https://training.cochrane.org/handbook/current/chapter-15>

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.

[https://doi.org/10.1016/S0005-7894\(96\)80013-X](https://doi.org/10.1016/S0005-7894(96)80013-X)

* Short, K. H., & Johnston, C. (1997). Stress, maternal distress, and children's adjustment following immigration: The buffering role of social support. *Journal of Consulting and Clinical Psychology*, 65, 494–503. <https://doi.org/10.1037/0022-006X.65.3.494>

* Shoshani, A., Nakash, O., Zubida, H., & Harper, R. A. (2016). School engagement,

acculturation, and mental health among migrant adolescents in Israel. *School Psychology Quarterly*, 31, 181-197. <https://doi.org/10.1037/spq0000133>

- * Sim, A., Bowes, L., & Gardner, F. (2018). Modeling the effects of war exposure and daily stressors on maternal mental health, parenting, and child psychosocial adjustment: A cross-sectional study with Syrian refugees in Lebanon. *Global Mental Health*, 5, e40.

<https://doi.org/10.1017/gmh.2018.33>

- * Slonim-Nevo, V., Mirsky, J., Rubinstein, L., & Nauck, B. (2009). The impact of familial and environmental factors on the adjustment of immigrants: A longitudinal study. *Journal of Family Issues*, 30, 92-123. <https://doi.org/10.1177/0192513x08324575>

- * Slonim-Nevo, V., & Sharaga, Y. (2000). Psychological and social adjustment of Russian-born and Israeli-born Jewish adolescents. *Child and Adolescent Social Work Journal*, 17, 455-475. <https://doi.org/10.1023/A:1026483828067>

Spence, S. H. (1997). Structure of anxiety symptoms among children: A confirmatory factor-analytic study. *Journal of Abnormal Psychology*, 106, 280-297.

<https://doi.org/10.1037/0021-843x.106.2.280>

Sperry, L. (2012). Ongoing assessment of couples and families. In L. Sperry (Ed.), *Family assessment: Contemporary and cutting-edge strategies* (2nd ed., pp. 115–136). Taylor & Francis.

Spielberger, C. D. (1973). *Manual for the State-Trait Anxiety Inventory for Children*. Consulting Psychologists Press.

Stafford, M., Newbold, B. K., & Ross, N. A. (2011). Psychological distress among immigrants and visible minorities in Canada: A contextual analysis. *International Journal of Social Psychiatry*, 57, 428-441. <https://doi.org/10.1177/0020764010365407>

- * Stein, G. L., & Polo, A. J. (2014). Parent–child cultural value gaps and depressive symptoms among Mexican American youth. *Journal of Child and Family Studies*, 23, 189-199.
<https://doi.org/10.1007/s10826-013-9724-3>
- Steinberg, A. M., Brymer, M. J., Decker, K. B., & Pynoos, R. S. (2004). The University of California at Los Angeles Post-traumatic Stress Disorder Reaction Index. *Current Psychiatry Reports*, 9, 96-100. <https://doi.org/10.1007/s11920-004-0048-2>
- Stevens, G. W., & Vollebergh, W. A. (2008). Mental health in migrant children. *Journal of Child Psychology and Psychiatry*, 49, 276-294. <https://doi.org/10.1111/j.1469-7610.2007.01848.x>
- Tam, S. Y., Houlihan, S., & Melendez-Torres, G. J. (2017). A systematic review of longitudinal risk and protective factors and correlates for posttraumatic stress and its natural history in forcibly displaced children. *Trauma Violence Abuse*, 18, 377-395.
<https://doi.org/10.1177/1524838015622437>
- * Tummala-Narra, P., & Claudius, M. (2013). Perceived discrimination and depressive symptoms among immigrant-origin adolescents. *Cultural Diversity and Ethnic Minority Psychology*, 19, 257-269. <https://doi.org/10.1037/a0032960>
- United Nations High Commissioner for Refugees. (2020). *Global trends: Forced displacement in 2019*. <https://www.unhcr.org/globaltrends2019/>
- van Os, E. C., Kalverboer, M. E., Zijlstra, A. E., Post, W. J., & Knorth, E. J. (2016). Knowledge of the unknown child: A systematic review of the elements of the best interests of the child assessment for recently arrived refugee children. *Clinical Child and Family Psychology Review*, 19, 185-203. <https://doi.org/10.1007/s10567-016-0209-y>
- van Santvoort, F., Hosman, C. M., Janssens, J. M., van Doesum, K. T., Reupert, A., & van Loon,

- L. M. (2015). The impact of various parental mental disorders on children's diagnoses: A systematic review. *Clinical Child and Family Psychology Review*, 18, 281-299.
<https://doi.org/10.1007/s10567-015-0191-9>
- Vang, Z. M., Sigouin, J., Flenon, A., & Gagnon, A. (2015). The healthy immigrant effect in Canada: A systematic review. *Population change and lifecourse strategic knowledge cluster discussion paper series/Un réseau stratégique de connaissances changements de population et parcours de vie document de travail*, 3. <http://ir.lib.uwo.ca/pclc/vol3/iss1/4>
- Vang, Z. M., Sigouin, J., Flenon, A., & Gagnon, A. (2017). Are immigrants healthier than native-born Canadians? A systematic review of the healthy immigrant effect in Canada. *Ethnicity & Health*, 22, 209-241. <https://doi.org/10.1080/13557858.2016.1246518>
- Yan, N., & Ansari, A. (2016). Child adjustment and parent functioning: Considering the role of child-driven effects. *Journal of Family Psychology*, 30, 297-308.
<https://doi.org/10.1037/fam0000180>
- * Yayan, E. H., & Duken, M. E. (2019). Determination of psychosocial conditions of refugee children living in society. *Perspectives in Psychiatric Care*, 55, 644-651.
<https://doi.org/10.1111/ppc.12387>
- * Yayan, E. H., Duken, M. E., Ozdemir, A. A., & Celebioglu, A. (2020). Mental health problems of Syrian refugee children: Post-traumatic stress, depression and anxiety. *Journal of Pediatric Nursing*, 51, e27-e32. <https://doi.org/10.1016/j.pedn.2019.06.012>
- Yingling, L. C. (2012). Parent-child and family assessment strategy and inventories. In L. Sperry (Ed.), *Family assessment: Contemporary and cutting-edge strategies* (2nd ed., pp. 203–230). Taylor & Francis.
- Zvara, B. J., Sheppard, K. W., & Cox, M. (2018). Bidirectional effects between parenting

sensitivity and child behavior: A cross-lagged analysis across middle childhood and adolescence. *Journal of Family Psychology*, 32, 484-495.

<https://doi.org/10.1037/fam0000372>

Table 1*Literature Reviews of Child Adjustment*

Study	Results	Methodological considerations
Blackmore et al., 2020 (Meta-analysis of refugee and asylum-seeking youth: 8 studies of accompanied youth)	Main results: • Overall, rates of PTSD (22.71%), anxiety (15.77%), depression (13.81%), and ADHD (8.6%) diagnoses were higher in refugee samples than worldwide prevalence rates; rates were lower for ODD (1.69%) Moderators of adjustment: • <u>Years in host country</u>: Prevalence of all diagnoses (except ADHD) were higher among those displaced less than 2 years • <u>Visa Status</u>: Higher prevalence of PTSD and ADHD were associated with <u>insecure visa status</u>; higher depression and anxiety were both associated with <u>refugee visa status</u> • <u>Language</u>: Lower prevalence of PTSD was associated in studies that conducted diagnostic interviews in the <u>native language</u> of refugee child/adolescent; this was associated with higher prevalence rates of both depression and anxiety disorders • <u>Residence</u>: Lower PTSD was associated with <u>current residence status</u> (i.e., those residing in the local community fared better than those living in refugee facility/reception center); living in the community was associated with higher prevalence of depression and anxiety	• Small number of studies limited generalizability of findings • Subgroup analyses for sex could not be calculated because of lack of reported data for: PTSD, depression • Inclusion criteria: Studies needed to use a diagnostic interview to determine diagnosis
Bronstein & Montgomery, 2011 (Systematic review of refugee youth: 22 studies)	Main results: • Prevalence rates (proportion of youth scoring above clinical cut-off of measure) varied across studies: PTSD (ranged from 19 – 54%), depression (3 – 30%) and general internalizing and externalizing behaviors were high among refugee youth Group differences or correlates of adjustment: • <u>Age</u>: Older refugee youth reported significantly higher levels of PTSD symptoms and externalizing and internalizing behavior than younger refugees • <u>Gender</u>: Girls reported higher levels of internalizing behavior, overall emotional distress, and depression. Boys reported higher total difficulties and conduct problems. Results for PTSD were mixed • <u>Country of origin</u>: Youth from Eritrea, Ethiopia, and Guinea reported greater symptoms of internalizing and externalizing behavior; Ethiopian youth of the Oromo community reported higher PTSD scores; Chinese youth reported overall lower levels of distress • Specific <u>pre-migratory adverse events</u> (e.g., parental separation, personal injury, violent death of a family member) were correlated with increased scores of PTSD. Cumulative adverse events predicted or were correlated with increased scores on measures of internalizing and externalizing behavior, PTSD, and depression	• Limited to self-report data • Data not presented separately for accompanied and unaccompanied refugee minors • Heterogeneity between populations and methods precluded meta-analysis • Not all included studies involved a nonmigrant comparison group • Inclusion criteria: Studies needed to use a measure with evidence of psychometric adequacy

Study	Results	Methodological considerations
Curtis et al., 2018 (Systematic review of migrant youth: 47 studies; 38 specific to mental health)	• Specific <u>post-migration stressors</u> (e.g., unstable asylum status or failed claims, low levels of support, discrimination, more restricted living arrangements, lower language level, poorer financial situation), were associated with increased levels of PTSD, depression, or internalizing behavior across studies Main results: Included immigrant, refugee, and asylum-seeking youth • PTSD symptoms were higher among refugee youth and depression and anxiety symptoms were higher among immigrant youth compared to nonmigrant peers: ◦ PTSD prevalence rates among refugee youth varied widely, from 20 to 84% of children reported clinically significant levels of symptoms Group differences or correlates of adjustment: • Among refugee youth: ◦ Unaccompanied youth were at a higher risk of PTSD than accompanied youth ◦ The <u>number of stressful events</u> positively correlated with PTSD symptoms ◦ Effect of <u>age</u> on symptoms was unclear, as age differences were often unreported ◦ Effect of <u>gender</u> among refugees was hard to determine, as boys tended to be overrepresented. Available results varied from no gender difference to higher prevalence among girls for symptoms of PTSD, stress, anxiety, and depression ◦ <u>Newly arrived</u> refugee youth reported high levels of anxiety and depression ◦ Cumulative pre- and post-migration stressors (e.g., insecure status in host country, frequent changes in housing, school, and social connections) may exacerbate symptoms and contribute to deterioration over time • Among immigrant youth: ◦ Stress, anxiety, and depression symptoms appear to vary according to level of psychological adaptation ◦ The potential relationship between age or gender and mental health outcomes was mixed across studies	• Limited to studies conducted in Europe • Heterogeneity across measures precluded meta-analysis • Limited to self-report data • Did not present data comparing immigrant to refugee youth • Includes studies of “return migrants” who migrated back to country of origin • Did not describe measures used (use of measures with evidence of psychometric adequacy was not an inclusion criterion)
Dimitrova et al., 2016 (Meta-analysis of immigrant youth: 51 studies)	Main results: Examined academic, externalizing, and internalizing adjustment among first-generation immigrant youth • Overall, immigrant youth reported somewhat poorer adjustment across internalizing and externalizing problems, and slightly better academic adjustment than comparison groups Moderators of adjustment: • <u>Host region</u>: immigrant youth in Northern Europe reported significantly more externalizing problems than youth in Southern Europe; those in Western Europe reported significantly more internalizing problems than those in Northern or Southern Europe • <u>Developmental stage</u>: adolescent (compared to child) samples reported significantly more externalizing problems, whereas child samples reported greater internalizing problems.	• Limited to studies conducted in Europe • Comparison group could be either nonmigrant or 2nd generation immigrant samples • Did not describe measures used (use of measures with evidence of psychometric adequacy was not an inclusion criterion) • Did not specify how they

Study	Results	Methodological considerations
	• <u>Gender</u>: samples with fewer girls reported increased academic problems • <u>Socio-economic status (SES)</u>: Samples with lower SES reported higher academic adjustment problems • <u>Residence</u>: Immigrants with long-term residence in the host country reported fewer externalizing problems compared to those more recently arrived • <u>Family reunification</u> positively related to lower internalizing and academic problems	accounted for multiple informants' ratings in the meta-analyses
Fazel et al., 2005 (Meta-analysis of refugee youth: 5 studies)	Main results: Refugee children • Five studies assessing children identified: all evaluated PTSD and yielded an overall prevalence of 11% (7–17%); no studies of depression, anxiety, or other disorders among children identified Moderators of adjustment: • None presented (tests of heterogeneity not significant)	• Limited to studies conducted in western countries • Limited to studies that recruited from the general population (excluded clinic-referred samples) • Excluded studies using self-report questionnaires • Inclusion criteria: Studies needed to use an interview-based psychiatric survey
Fazel et al., 2012 (Systematic review of refugee and internally displaced youth: 44 studies)	Main results: • <u>Increased risk</u> of reporting symptoms of internalizing disorders or trauma if they were <u>directly exposed to violence</u> (prior to and following migration), were girls, family had poor financial support or unstable post-migration housing, experienced discrimination, or their parents exhibited poorer parental mental health • <u>Decreased risk</u> when they reported stronger <u>social support</u> (family cohesion, support from peers, positive school experiences) and more <u>stable resettlement</u> conditions in the host country	• Methodological and clinical heterogeneity precluded meta-analysis • Synthesis of results did not differentiate between informants • Use of measures with evidence of psychometric adequacy was not an inclusion criterion
Kien et al., 2019 (Systematic review of refugees and asylum-seeking youth: 47 studies)	Main results: • Prevalence rates (presented as interquartile ranges) varied widely across studies of PTSD (19.0 to 52.7%), major depression (10.3 to 32.8%), anxiety disorders (8.7 to 31.6%), and emotional and behavioral problems (19.8 to 35.0%) Group differences or correlates of adjustment: • Rates of disorder or problems higher among <u>unaccompanied</u> minors than accompanied minors • <u>Assessment method</u>: studies using clinical interviews to diagnose PTSD or depression yielded lower prevalence rates than studies relying on self-report or proxy-questionnaires;	• Limited to studies conducted in Europe • Few studies ($k = 5$) included a nonmigrant comparison group • High and unexplained heterogeneity across outcomes impeded meta-analysis • Data not presented separately for informants of adjustment

Study	Results	Methodological considerations
	not reported for anxiety disorders or emotional/behavioral problems (relied on self-report/proxy-questionnaires)	• Inclusion criteria: Studies needed to use a measure with evidence of psychometric adequacy
Marley & Mauki, 2019 (Systematic review of refugee youth: 11 studies)	Main results: • Despite adversity, refugee youth who reported <u>higher levels</u> of these <u>protective</u> factors also reported more positive outcomes: peer social support, family unity and parental/familial support, social support from ethnically similar community members, self-esteem, parental education, maintaining one's cultural identity, school connectedness/feeling safe, community connectedness, and innovative social care services • One study evaluating emotional distress among unaccompanied refugee minors indicated that older age was associated with poorer mental health outcomes	• Synthesis of results did not differentiate between informants • Synthesis of results did not differentiate between constructs of adjustment, referring to "positive outcomes" generally • Use of measures with evidence of psychometric adequacy was not an inclusion criterion • Evaluation of methodological quality rated most (8/11) of the included studies as above average in quality
Porter & Haslam, 2005 (Meta-analysis of refugees and internally displaced adults and youth: 56 studies: 16 included child/adolescent samples)	Main results: • Overall, refugees' (including children, adolescents, and adults) mental health was moderately poorer ($d = 0.41$) than nonrefugee comparison groups Moderators of adjustment: • Child and adolescent refugees displayed better adjustment than adults	• Comparison groups varied widely across studies • Results mainly not presented separately by age or developmental stage • Authors calculated an average effect size when multiple measures of adjustment were included in a single sample • Use of measures with evidence of psychometric adequacy was not an inclusion criterion
Reed et al., 2012 (Systematic review of refugee youth: 27 studies)	Main results: • <u>Gender</u>: Girls were at a heightened risk for internalizing problems, whereas boys exhibited increased risk for externalizing problems • Pre-migration exposure to violence, resettlement in a refugee camp, and being internally displaced emerged as more <u>general risk factors</u> for mental health problems	• Methodological and clinical heterogeneity precluded meta-analysis • Synthesis of results did not differentiate between

Study	Results	Methodological considerations
Stevens & Vollebergh, 2008 (Systematic review among labour/economic migrant youth: 20 studies)	• <u>Repatriation</u> to one's country of origin emerged as protective factor among forcibly displaced youth Main results: • The results across included studies varied widely, and did not unequivocally identify an increased risk of mental health problems among migrant youth compared to nonmigrant peers: in some samples, migrant youth were reported to have better adjustment than nonmigrants, other studies revealed no differences, or better adjustment among nonmigrants • However, authors note that few studies accounted for different informants of child adjustment, or when these reports differed when evaluating the same children. Authors proposed that these differences could be explained by (1) true differences in child/adolescent behavior across contexts, (2) differences across informants in what constitutes "problem" behavior and meets the threshold for reporting it, or (3) biases that may be inherent in self-, parent, or teacher reports of problems	informants • Use of measures with evidence of psychometric adequacy was not an inclusion criterion • Limited to labour migrants • Authors noted that few studies reported on the reliability or validity of the instruments used with the migrant sample. • Inclusion criteria: Studies needed to use a measure with evidence of psychometric adequacy
Tam et al., 2017 (Systematic review of forcibly displaced youth: 11 studies)	Main results: • Eleven longitudinal studies (of seven independent samples) identified PTSD prevalence rates as staying relatively stable over time, though the number of longer-term longitudinal studies was limited: rates ranged from 20 to 48.7% (baseline), 10 to 48.3% (after 1 year; $k = 7$), 18 to 48% (2 to 3 years; $k = 2$), 8 to 38% (6 years; $k = 2$), and 35% (12 years; $k = 1$) Group differences or correlates of adjustment: • Greater symptoms of post-traumatic stress were reported by youth who reported greater exposure to potentially traumatic stressors and/or stressors associated with daily life/resettlement, who were older, and those who also reported greater symptoms of depression and/or anxiety • Findings examining differences in (1) time since displacement and (2) between girls and boys were either inconsistent or no association was found. However, girls were the minority (15.5 to 35%) in all included studies	• Samples included internally and externally displaced youth, unaccompanied and accompanied refugee youth • Heterogeneity precluded meta-analysis • Authors described the overall quality of studies as low • Synthesis of results did not differentiate between informants • Use of measures with evidence of psychometric adequacy was not an inclusion criterion
van Os et al., 2016 (Systematic review of mental health problems among recently arrived accompanied refugee)	Main results: • The most common mental health problems reported by refugee youth on arrival were symptoms of PTSD, depression and anxiety disorders, impacting up to half of unaccompanied minors and accompanied minors, though this varied across studies Group differences or correlates of adjustment: • Among accompanied youth	• Heterogeneity across instruments and definitions of refugee precluded meta-analysis • Synthesis of results did not differentiate between informants

Study	Results	Methodological considerations
youth: 8 studies)	○ The number of adverse life events and (if separated from parents) the duration of separation was associated with increased risk of PTSD symptoms ○ Greater exposure to violence was associated with increased internalizing, externalizing, and PTSD symptoms ○ The effect of age was mixed: older youth appeared to have increased risk of PTSD symptoms in two studies, but not in the others	● Use of measures with evidence of psychometric adequacy was not an inclusion criterion

Table 2*Literature Reviews of Adult Adjustment*

Study	Results	Methodological considerations
Anderson et al., 2017 (Meta-analysis of immigrant and refugee mothers: 53 studies)	Main results: • No evidence for overall increased risk of antenatal or postnatal depression among migrant women in included studies, except when results were presented separately by host country: migrant women in Canada displayed increased risk for antenatal and postnatal depression compared to samples in the United States and Australia • Authors suggested characteristics of the included samples may have impacted these results, but available data were insufficient to conduct additional analyses • Majority of studies (50 of 53) investigated depression alone; two assessed depression in combination with (1) anxiety) and (2) PTSD; one assessed any diagnosis Moderators of adjustment: • Sensitivity analyses stratified by measure and host country did not explain heterogeneity of effect sizes when examining risk for antenatal or postpartum depression (lack of data availability precluded further analyses)	• Due to high heterogeneity, authors did not report pooled estimate, and conducted sensitivity analyses that excluded all studies rated at a high risk of bias • Inclusion criteria: Studies needed to use a measure with evidence of psychometric adequacy
Falah-Hassani et al., 2015 (Meta-analysis of immigrant mothers: 22 studies)	Main results: • Overall, immigrant women were 1.5 to 2 times more likely to report postpartum depressive symptoms than nonmigrant women Moderators of adjustment: • Shorter time in host country, lower social support, poorer marital adjustment and perceived insufficient household income were associated with greater symptoms	• Inclusion criteria: Studies needed to use a measure with evidence of psychometric adequacy
Fazel et al., 2005 (Meta-analysis of refugee adults: 20 studies)	Main results: Refugee adults • Presented meta-analytic results specifically for larger studies ($N = 200+$), which yielded estimates of PTSD of 9% (99% CI 8–10%), and 5% (4–6%) for major depression, with evidence of a high psychiatric comorbidity between the two. Prevalence rates of generalized anxiety disorder were estimated at 4% (3–6%) ○ Far fewer studies available for psychotic disorder ($k = 2$) Moderators of adjustment: • <u>Methodological quality</u>: Larger surveys with greater methodological rigor yielded lower prevalence rates than studies lower in methodological quality, although heterogeneity was maintained when examining findings from larger studies alone	• High heterogeneity across samples • Low number of studies used in meta-analysis itself reduces generalizability • Use of measures with evidence of psychometric adequacy was not an inclusion criterion

Study	Results	Methodological considerations
Fellmeth et al., 2017 (Meta-analysis of immigrant and refugee mothers: 40 studies)	Main results: Results collapsed across immigrant and refugee mothers • Migrants from low- and middle-income countries (LMIC) reported higher rates of depression than nonmigrants in high-income countries (HIC) and LMIC, though the pooled mean scores on measures of depression were below the clinical cut-offs. • Results from individual studies suggest that rates of anxiety and PTSD were also high, but data were insufficient to calculate pooled estimates Moderators of adjustment: • Women with a previous history of depression and low social support appeared to be more likely to report perinatal depression	• Data not presented separately for immigrants and refugees • High heterogeneity across studies • Inclusion criteria: Studies needed to use a measure with evidence of psychometric adequacy
Porter & Haslam, 2005 (Meta-analysis of refugees and internally displaced adults and youth: 56 studies)	Main results: • Overall, refugees' mental health was moderately poorer ($d = 0.41$) than nonrefugee comparison groups Moderators of adjustment: • More <u>years in host country</u> associated with more positive outcomes (though many studies included participants with relatively short duration in host) • <u>Region of origin</u>: refugees from Europe reported more negative outcomes than refugees from Asia and Middle East • <u>Status in host country</u>: Refugees reported poorer adjustment than immigrants • However, when <u>exposure to violence</u> was accounted for, immigrants who had been exposed reported adjustment that was comparable to refugees • Studies higher in <u>methodological quality</u> yielded smaller effect sizes • <u>Gender</u>: women reported poorer outcomes than men	• Comparison groups varied widely across studies (e.g., some comparison groups had likely been exposed to high levels of violence due to ongoing conflict, other comparison groups were from more peaceful regions) • Use of measures with evidence of psychometric adequacy was not an inclusion criterion

Table 3*Characteristics of Child Adjustment Studies*

Study	Status	Country of origin	Host country (years in host)	Sample: <i>N</i> , <i>M</i> (<i>SD</i>) age & age range ^a	Informant	Construct (measure); reliability in sample	Methodological Quality Rating
Ahmad et al., 2008	Refugee	Kurdistan regions of Turkey, Iran, Iraq, Syria, Lebanon	Sweden (7.9)	<i>N</i> = 111; <i>M</i> (<i>SD</i>): <i>NR</i> , 6–18	Youth	Trauma (PTSS-C); α = .85	10
Aroian et al., 2009	Immigrant & refugee	Iraq; Lebanon; Others	United States (8.4)	<i>N</i> = 635; 13.8 (1.2), 11–15	Parent	EXT; INT (CBCL); α = .90–.92	8
Bean et al., 2007	Immigrant & Refugee	111 countries	Belgium, Netherlands (“Recent”)	<i>N</i> = 1,130; 15.3 (1.8), 13–18	Youth	Trauma (RATS); α = .73 ^b	9
Cho & Haslam, 2010	Immigrant	South Korea	United States (2.3)	<i>N</i> = 70; 16.6 (1.1), range: <i>NR</i>	Youth	Depression (BDI-II); α = .87	8
Cleary et al., 2018	Immigrant	El Salvador; Guatemala; Honduras	United States (1.3)	<i>N</i> = 101; 15.4 (1.7), 12–17	Youth	Anxiety (SCAS); Depression (PHQ-9); Trauma (CPSS); α = .84–.94	8
Dimitrova & Chasiotis, 2012	Immigrant	Serbia	Italy (< 10.0)	<i>N</i> = 124; 9.2 (1.5), 7–12	Youth	Depression (CDI); α = .71	9
Erucar et al., 2018	Refugee	Syria	Turkey (2.1)	<i>N</i> = 263; 11.6 (1.9), range: <i>NR</i>	Parent	Overall problems (SDQ-P); α = .71	10
Giordano et al., 2019	Asylum-seeker	Syria; Palestine	Italy (< 3.0)	<i>N</i> = 271; 10.0 (2.1), 6–14	Parent	Overall problems (SDQ-P); α = .73	9
Hamdan-Mansour et al., 2017	Refugee	Syria	Jordan (< 6.0)	<i>N</i> = 250; 11.4 (3.1), 6–18	Youth	Overall problems (SDQ-S); α = .81	10

Study	Status	Country of origin	Host country (years in host)	Sample: <i>N, M (SD)</i> age & age range ^a	Informant	Construct (measure); reliability in sample	Methodological Quality Rating
<i>Table 3 continued...</i>							
Jaycox et al., 2002	Immigrant	Mexico, El Salvador, Guatemala, Korea, Russia; Armenia	United States (< 3.0)	<i>N</i> = 1,004; 11.4 (1.8), 8–15	Youth	Trauma (CPSS); α = .89	8
Kovacev & Shute, 2004	Refugee	Former Yugoslavia	Australia (“New”)	<i>N</i> = 83; 15.3 (1.5), 12–19	Youth	Self-esteem (SPPA) ^c ; α = .71	8
Muller et al., 2019	Refugee	Afghanistan, Syria, Iraq	Germany (1.8)	<i>N</i> = 30; 15.1 (2.0), range: <i>NR</i>	Youth	Overall problems (HSC-37-A); α = .88	9
Quinlan et al., 2016	Refugee	Middle East, East Asia, Africa	Australia (< 1.0)	<i>N</i> = 42; INT: 15.8 (1.3), COM: 15.0 (1.7), range: <i>NR</i>	Teacher	Overall problems (SDQ-T); α = .76	9
Short & Johnston, 1997	Immigrant	Hong Kong	Canada (1.4)	<i>N</i> = 97; Girls: 8.6 (1.7), Boys: 8.0 (1.8), 6–11 yrs	Parent	Overall problems (CBCL); α = .95	9
Shoshani et al., 2016	Immigrant	Asia, Philippines, Eastern Europe, South America, Africa, Middle East	Israel (7.8)	<i>N</i> = 128; 14.9 (2.0), age range: <i>NR</i>	Youth	Overall problems (BSI); α = .74	11
Slonim-Nevo et al., 2009	Immigrant	Former Soviet Union	Germany; Israel (6–8 months)	<i>N</i> = 195; Germany: 14.2 (2.0); Israel: 16.0 (1.2); 12–18	Youth	Overall problems (BSI); α = .94	10
Slonim-Nevo & Sharaga 2000	Immigrant	Former Soviet Union	Israel (“New”)	<i>N</i> = 75; 15.0 (<i>NR</i>), 14–17	Youth	Overall problems (BSI); α = .96	10
Stein & Polo, 2014	Immigrant	Mexico, Central or South America	United States (7.7)	<i>N</i> = 77; 13.1 (0.7), 11–15	Youth	Depression (CES-DC); α = .79	10

Study	Status	Country of origin	Host country (years in host)	Sample: <i>N</i> , <i>M</i> (<i>SD</i>) age & age range ^a	Informant	Construct (measure); reliability in sample	Methodological Quality Rating
<i>Table 3 continued...</i>							
Tummala-Narra & Claudius, 2013	Immigrant	Caribbean, South & Central America, Asia, Middle East, Africa	United States (6.3)	<i>N</i> = 49; 15.4 (<i>NR</i>), 13–19	Youth	Depression (CES-DC); α = .85	8
Yayan & Duken, 2019	Refugee	Syria	Turkey (< 7.0)	<i>N</i> = 738; 11.9 (3.1), 7–18	Youth	Anxiety (STAIC); Depression (CDI); α = .74–.78	9
Yayan et al., 2020	Refugee	Syria	Turkey (< 7.0)	<i>N</i> = 1,115; 10.5 (2.8), 9–15	Youth	Anxiety (STAIC); α = .78	9

Sample column: *COM* = comparison group; *INT* = intervention group; Construct (measure) column: *EXT* = Externalizing symptoms; *INT* = Internalizing symptoms; *Overall* = measures of overall problems *NR*: Not reported

BDI-II: Beck Depression Inventory, 2nd edition; BSI: Brief Symptom Inventory; CBCL: Child Behavior Checklist; CDI: Child Depression Inventory; CES-DC: Center for Epidemiological Studies Depression Scale for Children; CPSS: Child PTSD Symptom Scale; HSC-37-A: Hopkins Symptom Checklist-37 for Adolescents; PHQ-9: Patient Health Questionnaire–Depression scale; PTSS-C: Post-traumatic Stress Symptoms in Children; RATS: Reactions of Adolescents to Traumatic Stress; SCAS: Spence Children's Anxiety Scale; SDQ: Strengths & Difficulties Questionnaire (-S: child self-report; -P: parent report; -T: teacher report), total problem score; SPPA: Self-Perception Profile for Adolescents; STAIC: State-Trait Anxiety Inventory for Children

a: Mean (SD) & age range reported in years, unless indicated as months

b: Internal consistency reported in Derluyn, Mels, & Broekaert (2009) with an overlapping sample

c: The effect direction for this measure was reversed so that higher scores reflect worse functioning

Table 4*Characteristics of Parental Adjustment Studies*

Study	Status	Country of origin	Host country (years in host)	Sample: <i>N</i> , <i>M</i> (<i>SD</i>) age & age range ^a Parents Children		Measured child adjust?	Construct (measure); reliability in sample	Methodological Quality Rating
Alhasanat-Khalil et al., 2018	Immigrant	Yemen, Lebanon, Iraq, others	United States (7.8)	<i>N</i> = 115 mothers; 29.3 (5.3), 19–40	<i>NR</i>	No	PP-Depression (EPDS); α = .79	9
Aroian et al., 2009	Immigrant & refugee	Iraq; Lebanon; Other Arab countries	United States (8.4)	<i>N</i> = 635 mothers; 44.5 (8.4), range: <i>NR</i>	<i>N</i> = 635; 13.8 (1.2), 11–15	Yes	Depression (CES-D); α = .91	8
Beeber et al., 2010	Immigrant	Latina (mainly Mexico)	United States (5.3)	<i>N</i> = 80 mothers; INT: 26.2 (6.1); COM: 26.5 (5.8), range: <i>NR</i>	<i>N</i> = 80; INT: 1.3 (1.0), COM: 1.4 (1.0), 0.1–3.0	No	Depression (CES-D); α = .90	10
Costigan & Koryzma, 2011	Immigrant	People's Republic of China, Taiwan, Hong Kong, Singapore; Other	Canada (7.0)	<i>N</i> = 177; fathers: 44.7 (4.6), mothers: 41.7 (4.0), range: <i>NR</i>	<i>N</i> = 177; 11.8 (1.8), 10–14	No	Depression (CES-D); α = .83–.88	9
Kirchner et al., 2011	Immigrant	Colombia, Bolivia, Ecuador, Peru, Honduras, Paraguay	Spain (2.4–4.6)	<i>N</i> = 105; fathers: 36.9 (8.9), mothers: 34.9 (8.3), range: <i>NR</i>	<i>NR</i>	No	Overall problems (SCL-90-R); α = .97	10
Lecompte & Rousseau, 2018	Immigrant	Bangladesh, Comoros, Ghana, Haiti, India, Morocco, Mexico, Pakistan, Sri Lanka, Turkey, Tunisia	Canada (8.0)	<i>N</i> = 37 mothers; <i>M</i> & range: <i>NR</i> ; 86% between ages 20–40	<i>N</i> = 37; 37.0 (19) months, 0.42–5 yrs	No	Anxiety, Depression (HSCL-25); α = .79–.88	9

Study	Status	Country of origin	Host country (years in host)	Sample: <i>N</i> , <i>M</i> (<i>SD</i>) age & age range ^a Parents Children		Measured child adjust?	Construct (measure); reliability in sample	Methodological Quality Rating
<i>Table 4 continued...</i>								
Mohammad et al., 2018	Refugee	Syria	Jordan (<7.0)	<i>N</i> = 365 mothers; <i>M</i> : <i>NR</i> , 18–45	<i>NR</i>	No	PP-Depression (EPDS); α = .76	11
Nakash et al., 2016	Asylum-Seeker	Eritrea	Israel (<6.0)	<i>N</i> = 38 mothers; <i>M</i> : <i>NR</i> , 20–37	<i>NR</i>	No	PP-Depression (EPDS); α = .72	9
Short & Johnston, 1997	Immigrant	Hong Kong	Canada (1.3–1.4)	<i>N</i> = 97; mothers of boys: 37.0 (4.0), of girls: 37.7 (4.1), range: <i>NR</i>	<i>N</i> = 97; Boys: 8.0 (1.8), Girls: 8.6 (1.7), 6–11	Yes	Depression (CES-D); α = .89	9
Sim et al., 2018	Refugee	Syria	Lebanon (<3.0)	<i>N</i> = 291 mothers; 31.8 (8.2), range: <i>NR</i>	<i>N</i> = 291; 7.4 (3.2), 2–12	No	Overall problems (DASS-21); α = .91	9

COM = comparison group; *INT* = intervention group; *NR*: Not reported; *PP-Depression* = Measure of depression used in postpartum period

All measures completed by parents as self-report; CES-D: Center for Epidemiological Studies Depression Scale; DASS-21: Depression Anxiety Stress

Scales: EPDS: Edinburgh Postnatal Depression Scale; HSCL-25: Hopkins Symptom Checklist-25; SCL-90-R: Symptom Checklist-90-Revised

a: Mean (*SD*) & age range reported in years, unless indicated as weeks or months

Table 5*Characteristics of Parenting Studies*

Study	Status	Country of origin	Host country (years ^a in host)	Sample: <i>N</i> , <i>M</i> (<i>SD</i>) age & age range ^a Parents Children		Measured child adjust?	Measure (informant); reliability in sample	Methodological Quality Rating
Cyril et al., 2016	Immigrant & refugee	Africa	Australia (7.3)	<i>N</i> = 269 parents; 39.6 (7.9), range: <i>NR</i>	<i>N</i> = 273; 14.7 (1.9), 12–17	No	FAD (Parent & Youth); α = .78–.82	8
Mellor et al., 2012	Refugee	Sudan, Somalia, DRC, Burundi, Djibouti, Kenya, Tanzania	Australia (7.2)	<i>N</i> = 99 parents; 40.9 (7.0), range: <i>NR</i>	<i>N</i> = 100; 15.0 (1.7), 13–17	No	APQ (Parent & Youth) ^b ; α = .71–.87	10
Slonim-Nevo et al., 2009	Immigrant	Former Soviet Union	Germany, Israel (6–8 months)	<i>N</i> = 391 mothers & fathers; Germany: 39.2 (5.5)–41.1 (5.7); Israel: 41.5 (5.2)–43.1 (7.1), range: <i>NR</i>	<i>N</i> = 195; Germany: 14.2 (2.0); Israel: 16.0 (1.2), 12–18	Yes	FAD (Mothers, Fathers & Youth); α = .72	10

NR: Not reported; *Parents* = indicates both fathers and mothers participated

APQ: Alabama Parenting Questionnaire; FAD: McMaster Family Assessment Device (general functioning subscale)

a: Mean (SD) & age range reported in years, unless indicated as months

b: The effect direction for two subscales of this measure were reversed so that higher scores reflect worse functioning

Table 6*Overall Effect Sizes of Child Adjustment, Parental Adjustment, and Parenting across Informants*

Variable	Overall Hedges' g [CI]	Range of g	Heterogeneity Tests
Child adjustment: positive measures			
Self-esteem ($k = 1$)	-0.11 [-0.37 to 0.16], $p = .44$	-0.11	--
Child adjustment: problems			
Youth self-report			
Overall problems ($k = 5$)	0.53 [-0.45 to 1.51], $p = .29$	-0.68 to 1.68	$Q(4) = 573.70, p < .01; I^2 = 99.30$
Anxiety ($k = 3$)	0.87 [0.50 to 1.24], $p < .01$	0.36 to 1.15	$Q(2) = 47.61, p < .01; I^2 = 95.80$
Depression ($k = 6$)	-0.12 [-1.20 to 0.95], $p = .82$	-1.70 to 1.77	$Q(5) = 834.65, p < .01; I^2 = 99.40$
Trauma ($k = 4$)	-0.18 [-1.08 to 0.71], $p = .69$	-1.55 to 0.55	$Q(3) = 231.27, p < .01; I^2 = 98.70$
Parent report			
Overall problems ($k = 3$)	0.65 [0.39 to 0.92], $p < .01$	0.29 to 0.82	$Q(2) = 19.87, p < .01; I^2 = 89.93$
Externalizing ($k = 1$)	-0.13 [-0.23 to -0.03], $p = .01$	-0.13	--
Internalizing ($k = 1$)	0.77 [0.67 to 0.88], $p < .01$	0.77	--
Teacher report			
Overall problems ($k = 1$)	-0.20 [-0.51 to 0.10], $p = .19$	-0.20	--
Parental adjustment: problems			
Parent self-report			
Overall problems ($k = 2$)	2.14 [2.00 to 2.28], $p < .01$	2.13 to 2.15	$Q(1) = 0.02, p = .89; I^2 = 0.00$
Anxiety ($k = 1$)	1.36 [1.03 to 1.68], $p < .01$	1.36	--
Depression ($k = 8$)	0.96 [0.001 to 1.917], $p = .050$	-1.11 to 2.22	$Q(7) = 1890.90, p < .01; I^2 = 99.63$
Parenting: problems			
Youth report			
APQ ($k = 1$)	-0.28 [-0.60 to 0.04], $p = .09$	-0.28	--
FAD ($k = 2$)	-0.38 [-0.60 to -0.17], $p < .01$	-0.49 to -0.27	$Q(1) = 4.00, p < .05; I^2 = 75.01$

Variable	Overall Hedges' g [CI]	Range of g	Heterogeneity Tests
<i>Table 6 continued...</i>			
Parent report			
APQ ($k = 1$)	-0.09 [-0.45 to 0.27], $p = .64$	-0.09	--
FAD ($k = 2$)	-0.55 [-0.78 to -0.32], $p < .01$	-0.66 to -0.43	$Q(1) = 6.04, p = .01; I^2 = 83.45$

Note: A positive effect size reflects greater problems

APQ: Alabama Parenting Questionnaire; FAD: McMaster Family Assessment Device (general functioning subscale)

Table 7*Effect Sizes of Adjustment in Immigrant, Refugee, and Combined Samples of Children/Adolescents*

Variable	Immigrant only: Overall Hedges' <i>g</i> [<i>CI</i>]	Refugee only: Overall Hedges' <i>g</i> [<i>CI</i>]	Combined: Overall Hedges' <i>g</i> [<i>CI</i>]
Youth report: Positive Measures, Self-esteem			
SPPA	--	-0.11 [-0.37 to 0.16] (<i>k</i> = 1)	--
Youth report: Problems			
Overall Problems	--	0.94 [-0.54 to 2.42] * (<i>k</i> = 2)	--
BSI (GSI)	0.26 [-0.78 to 1.29] * (<i>k</i> = 3)	--	--
HSC-37-A	--	0.18 [-0.19 to 0.54] (<i>k</i> = 1)	--
SDQ-S	--	1.68 [1.55 to 1.82] (<i>k</i> = 1)	--
Anxiety			
SCAS	0.36 [0.16 to 0.56] (<i>k</i> = 1)	--	--
STAIC	--	1.10 [1.02 to 1.18] (<i>k</i> = 2)	--
Depression			
	-0.50 [-1.64 to 0.64] * (<i>k</i> = 5)	--	--
BDI-II	-0.60 [-3.14 to 1.95] (<i>k</i> = 1)	--	--
CDI	0.04 [-2.50 to 2.58] (<i>k</i> = 1)	1.77 [1.67 to 1.87] (<i>k</i> = 1)	--
CES-DC	-0.77 [-2.58 to 1.03] * (<i>k</i> = 2)	--	--
PHQ-9	-0.42 [-2.97 to 2.12] (<i>k</i> = 1)	--	--
Trauma			
	--	--	--
CPSS	0.14 [-0.23 to 0.50] (<i>k</i> = 2)	--	--
PTSS-C	--	-1.55 [-4.05 to 0.95] (<i>k</i> = 1)	--
RATS	--	--	0.55 [0.47 to 0.64] (<i>k</i> = 1)
Parent report: Problems			
Overall problems	--	--	--
CBCL Total score	0.29 [0.08 to 0.50] (<i>k</i> = 1)	--	--

Variable	Immigrant only: Overall Hedges' <i>g</i> [<i>CI</i>]	Refugee only: Overall Hedges' <i>g</i> [<i>CI</i>]	Combined: Overall Hedges' <i>g</i> [<i>CI</i>]
<i>Table 7 continued...</i>			
SDQ-P	--	0.81 [0.72 to 0.89] (<i>k</i> = 2)	--
Externalizing	--	--	--
CBCL Externalizing scale	--	--	-0.13 [-0.23 to -0.03] (<i>k</i> = 1)
Internalizing	--	--	--
CBCL Internalizing scale	--	--	0.77 [0.67 to 0.88] (<i>k</i> = 1)
Teacher report: Problems			
Overall problems	--	--	--
SDQ-T	--	-0.20 [-0.51 to 0.10] (<i>k</i> = 1)	--

Note: A positive effect size reflects greater problems; * = tests of heterogeneity (*Q* statistic) < .05

BDI-II: Beck Depression Inventory, 2nd edition; BSI: Brief Symptom Inventory; CBCL: Child Behavior Checklist; CDI: Child Depression Inventory; CES-DC: Center for Epidemiological Studies Depression Scale for Children; CPSS: Child PTSD Symptom Scale; HSC-37-A: Hopkins Symptom Checklist-37 for Adolescents; PHQ-9: Patient Health Questionnaire–Depression scale; PTSS-C: Post-traumatic Stress Symptoms in Children; RATS: Reactions of Adolescents to Traumatic Stress; SCAS: Spence Children's Anxiety Scale; SDQ: Strengths & Difficulties Questionnaire (-S: child self-report; -P: parent report; -T: teacher report), total problem score; SPPA: Self-Perception Profile for Adolescents; STAIC: State-Trait Anxiety Inventory for Children

Table 8*Effect Sizes of Adjustment in Immigrant, Refugee, and Combined Samples of Parents*

Variable	Immigrant only: Overall Hedges' <i>g</i> [<i>CI</i>]	Refugee only: Overall Hedges' <i>g</i> [<i>CI</i>]	Combined: Overall Hedges' <i>g</i> [<i>CI</i>]
Parent self-report:			
Problems			
Overall problems	--	--	--
DASS-21	--	2.13 [1.95 to 2.31] (<i>k</i> = 1)	--
SCL-90-R	2.15 [-5.36 to 9.67] (<i>k</i> = 1)	--	--
Anxiety	--	--	--
HSCL-25 Anxiety scale	1.36 [1.03 to 1.68] (<i>k</i> = 1)	--	--
Depression	0.97 [0.13 to 1.81] * (<i>k</i> = 5)	--	--
CES-D	1.16 [0.11 to 2.22] * (<i>k</i> = 3)	--	-1.11 [-1.21 to -1.01] (<i>k</i> = 1)
EPDS	0.12 [-1.71 to 1.95] (<i>k</i> = 1)	1.94 [1.44 to 2.45] * (<i>k</i> = 2)	--
HSCL-25 Depression scale	1.27 [-0.58 to 3.12] (<i>k</i> = 1)	--	--

Note: A positive effect size reflects greater problems; * = tests of heterogeneity (*Q* statistic) < .05

CES-D: Center for Epidemiological Studies Depression Scale; DASS-21: Depression Anxiety Stress Scales; EPDS: Edinburgh Postnatal Depression Scale; HSCL-25: Hopkins Symptom Checklist-25; SCL-90-R: Symptom Checklist-90-Revised

Table 9*Effect Sizes of Parenting in Immigrant, Refugee, and Combined Samples of Families*

Variable		Immigrant only: Overall Hedges' <i>g</i> [<i>CI</i>]	Refugee only: Overall Hedges' <i>g</i> [<i>CI</i>]	Combined: Overall Hedges' <i>g</i> [<i>CI</i>]
Youth Self-report				--
	APQ	--	-0.28 [-0.46 to -0.10] (<i>k</i> = 1)	--
	FAD	-0.49 [-0.65 to -0.34] (<i>k</i> = 1)	--	-0.27 [-0.42 to -0.13] (<i>k</i> = 1)
Parent Report				
	APQ	--	-0.09 [-0.29 to 0.12] (<i>k</i> = 1)	--
	FAD	-0.66 [-0.77 to -0.55] (<i>k</i> = 1)	--	-0.43 [-0.58 to -0.28] (<i>k</i> = 1)

Note: A positive effect size reflects greater problems

APQ: Alabama Parenting Questionnaire; FAD: McMaster Family Assessment Device (general functioning subscale)

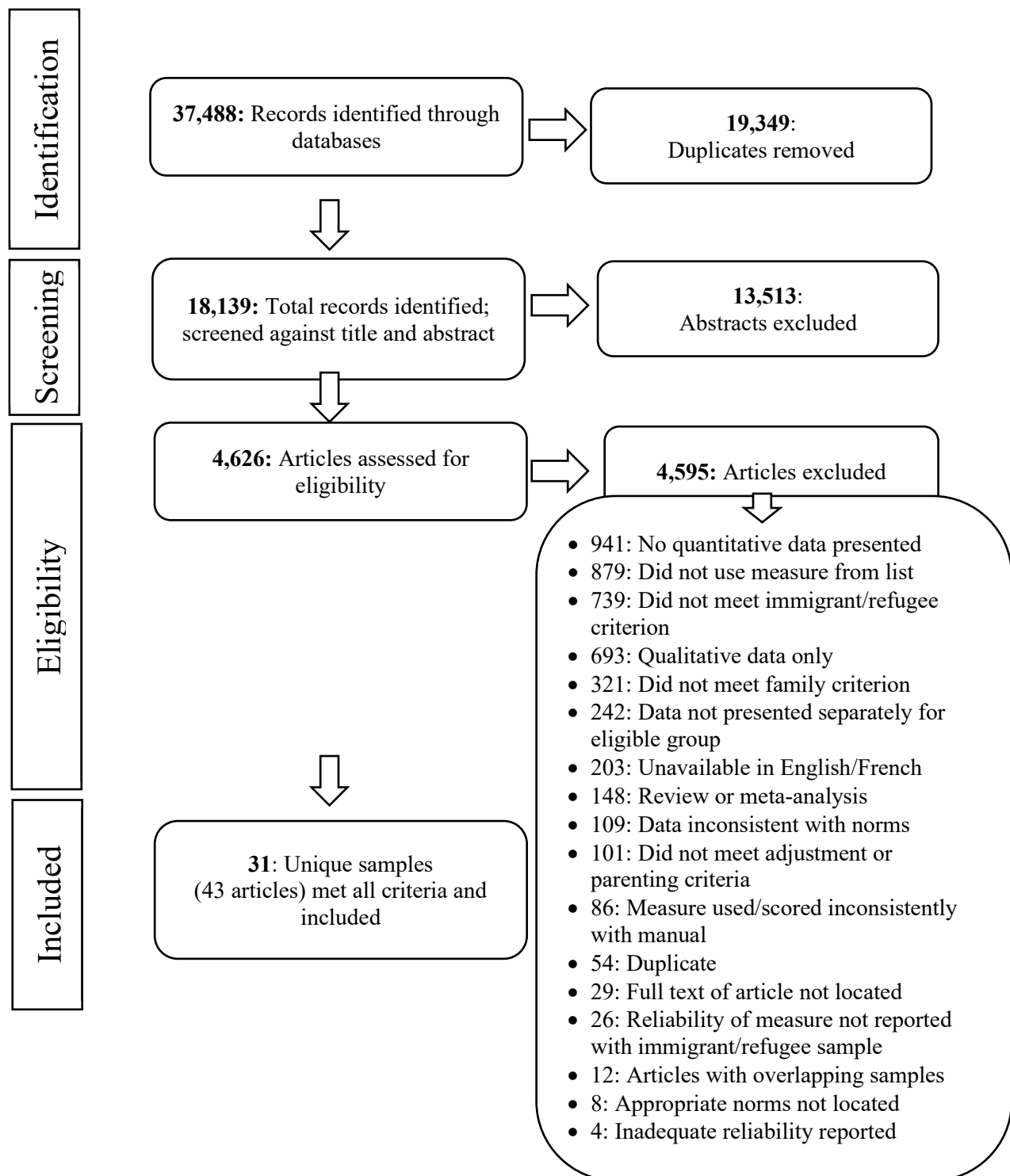
Figure 1*Flowchart of Included and Excluded Studies*

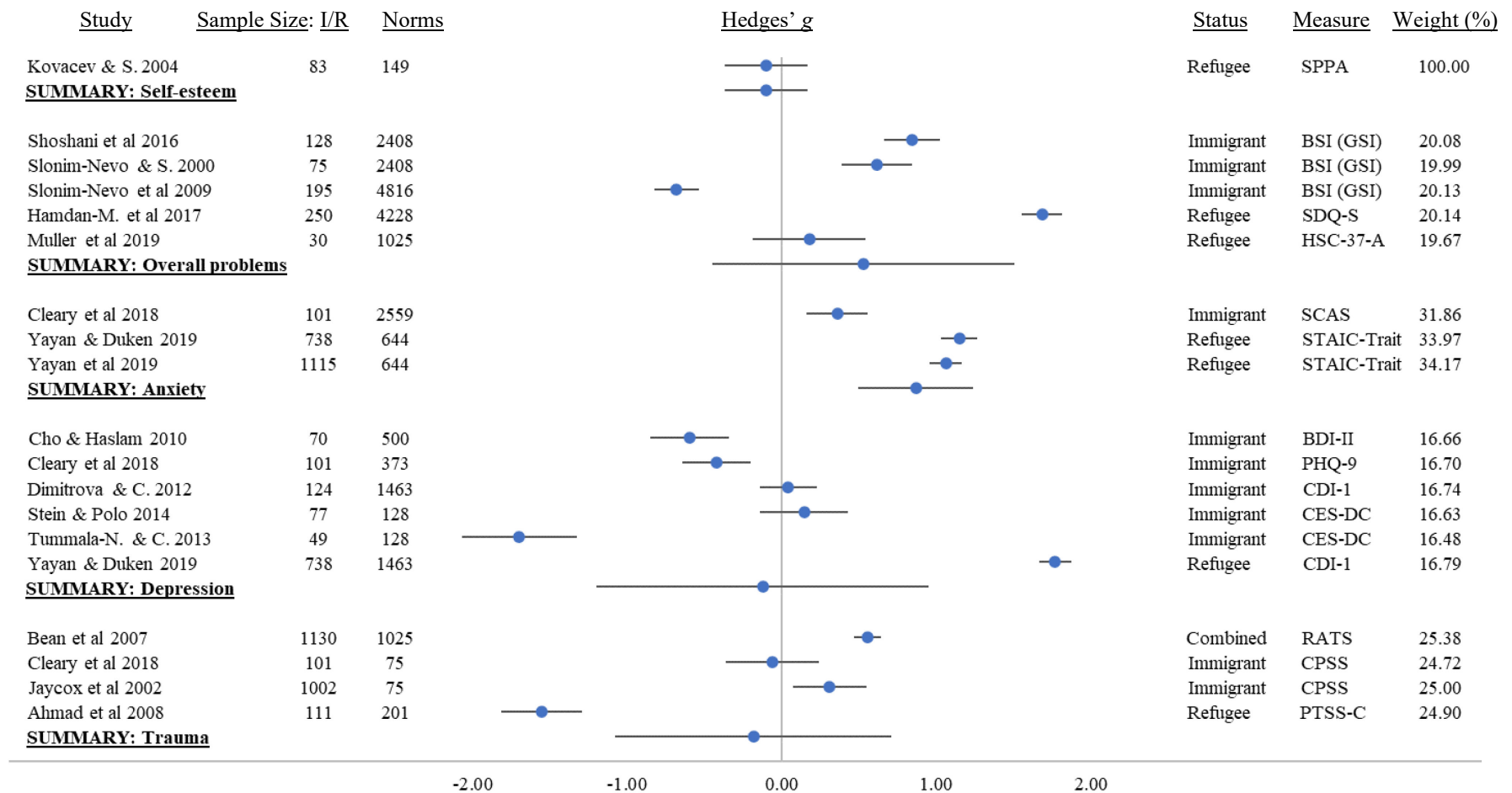
Figure 2*Individual & Summary Effect Sizes: Child Adjustment (Youth Self-Reports)*

Figure 3

Individual & Summary Effect Sizes: Child Adjustment (Parent and Teacher Reports)

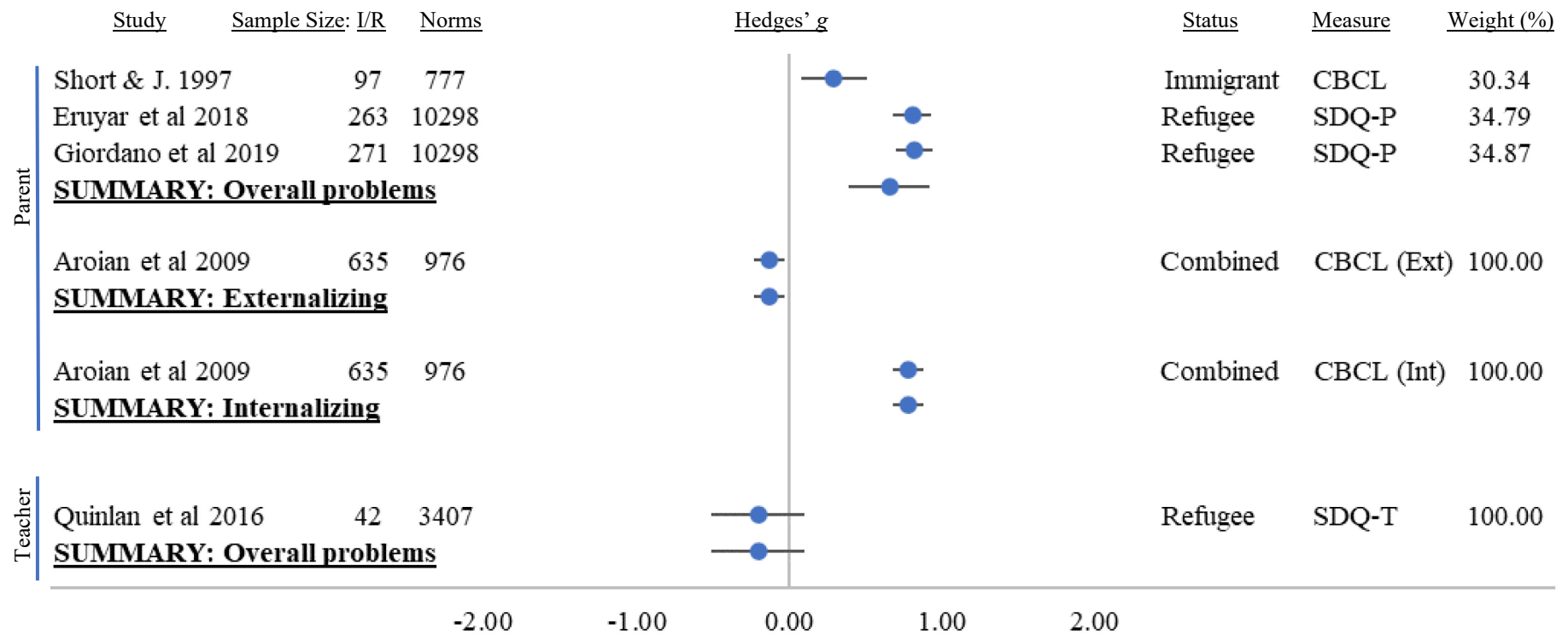


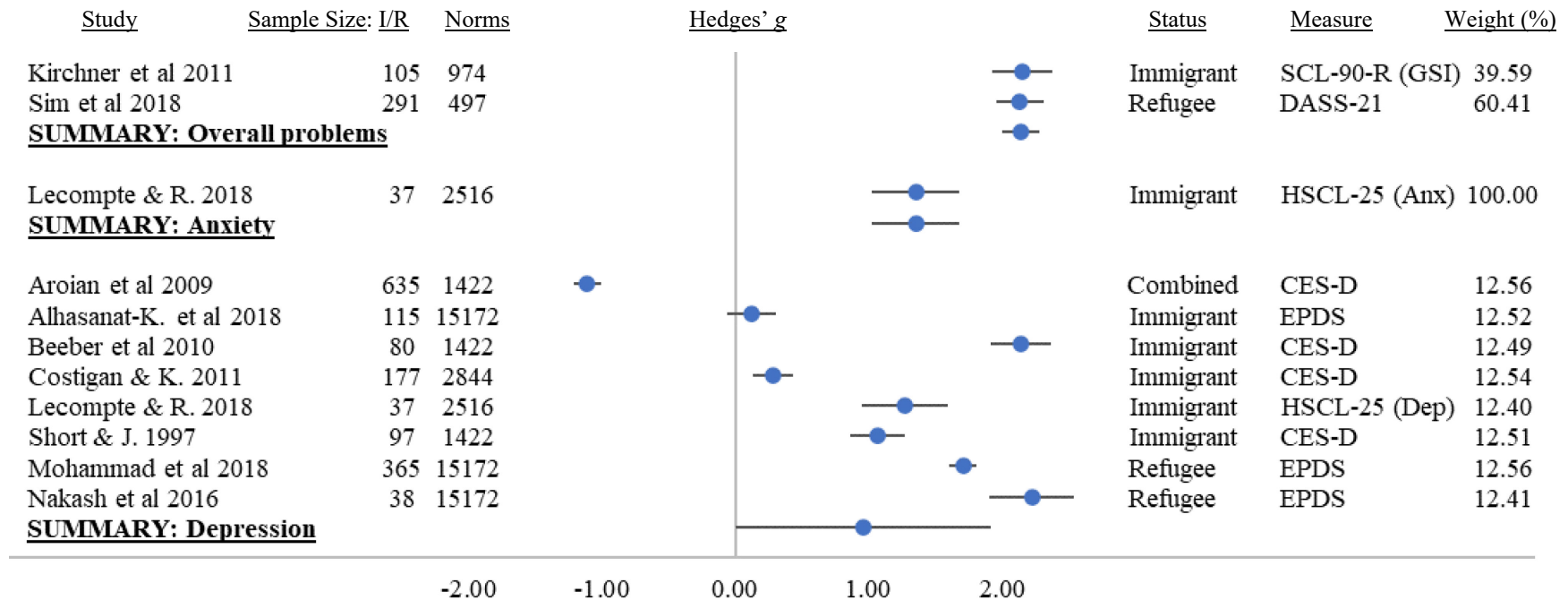
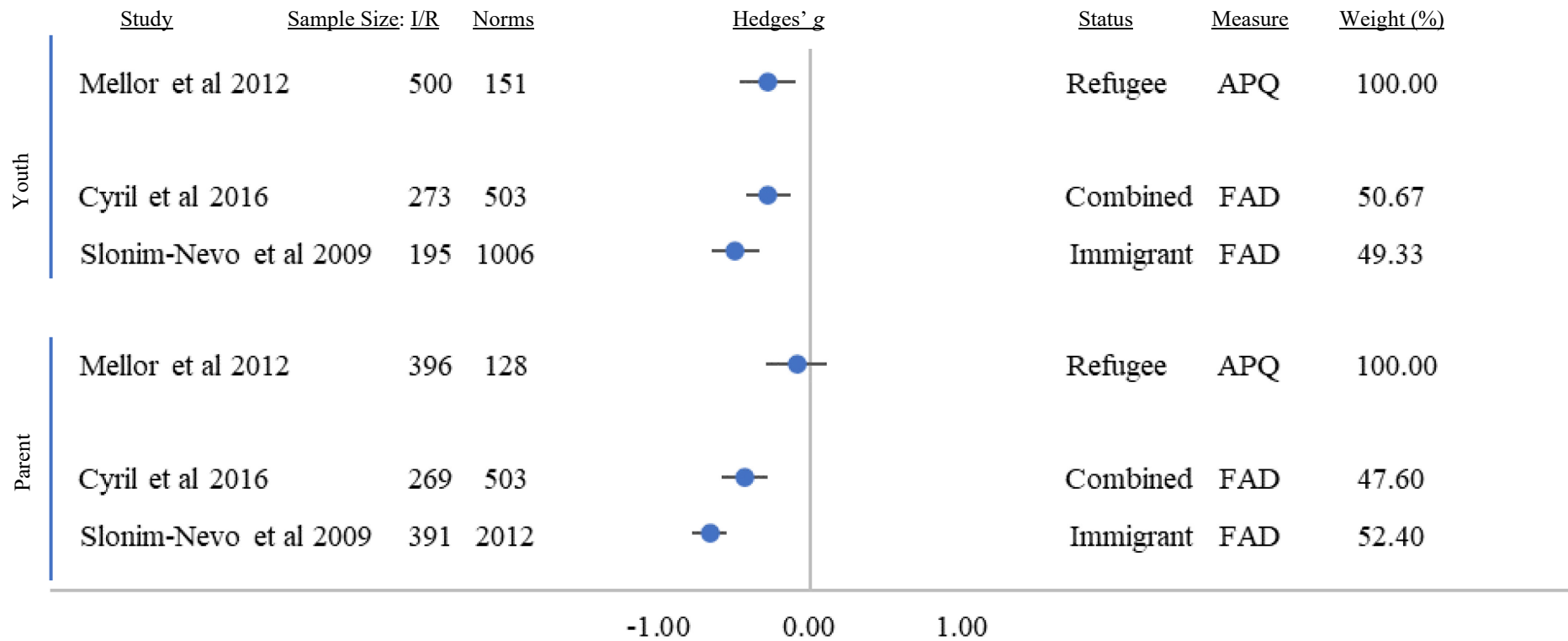
Figure 4*Individual & Summary Effect Sizes: Parental Adjustment*

Figure 5*Individual & Summary Effect Sizes: Parenting*

Study 2

Multimethod Adherence Measurement in an Evidence-Based Parenting Program

Dana M. Sheshko and Catherine M. Lee, University of Ottawa

Marie-Hélène Gagné, Université Laval

20 January 2020

Prepublication copy: Accepted to Practice Innovations on 26 January 2020

(DOI: 10.1037/pri0000110)

© 2020, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article will be available, upon publication, via its

DOI: 10.1037/pri0000110

AUTHOR NOTE

Dana M. Sheshko, School of Psychology, University of Ottawa; Catherine M. Lee, School of Psychology, University of Ottawa; Marie-Hélène Gagné, École de Psychologie, Université Laval

This research was supported by a partnership grant from the Social Sciences and Humanities Research Council of Canada. Catherine M. Lee is a consultant to Triple P Canada. Correspondence regarding this article can be addressed to Dana Sheshko, School of Psychology, University of Ottawa, Ottawa, ON K1N 6N5

Abstract

We present preliminary multimethod, multi-informant data on a new measure of adherence (Practitioner Session Reflection Tool) that invited practitioners to report on their adherence and content modifications to an evidence-based parenting program. Fifty-two practitioners from 12 community agencies submitted 361 self-reports; a subset of 23 practitioners submitted 63 audio-recordings to permit additional ratings by a coding team. Results support the feasibility of using the self-report tool in community settings and preliminary evidence of reliability and convergent validity. The comparison of ratings across practitioner self-report and external coders yielded low correlations between informants. This may highlight the different contexts from which informants identify and report on adherence, rather than as a measure of reporting accuracy. We review the implications of these results for adherence measurement using multiple informants, study limitations, clinical implications, and future directions for research.

Keywords: adherence, parenting, measure development, multiple informants

Clinical Impact Statement:

We present data on a new self-report adherence measure that may be a useful tool for practitioners to reflect on their delivery of a parenting program and promotion of parents' self-regulation.

Multimethod Adherence Measurement in an Evidence-Based Parenting Program

Effective parenting is linked to positive emotional and behavioral adjustment among children (Biglan, Flay, Embry, & Sandler, 2012), whereas ineffective or harsh parenting practices are associated with negative outcomes in child development (Clément, Gagné, & Lee, 2014; Gerschoff et al., 2018). There is abundant evidence that programs based on social learning theory are efficacious in supporting and enhancing positive parenting (Barlow, Parsons, & Stewart-Brown, 2005; Rothenberg, Weinstein, Dandes, & Jent, 2019). This study presents preliminary data on a new practitioner self-report measure of adherence in the delivery of a parenting intervention (Sheshko, Lee, & Gagné, 2015). To evaluate the usefulness of this tool in a community setting, we examine three issues. First, we assessed practitioners' completion rates. Second, we examined the internal consistency and interrater reliability of the new self-report tool and an observational measure designed to assess the same variables (Kirby & Sanders, 2014a, 2014b). Finally, for a subset of participants, we compared self-report adherence data to coder ratings based on the observational measure.

Evidence-based parenting interventions are designed to increase the frequency of positive interactions among parents and children, promote parental consistency, and enhance parents' skills to support children's emotional development (Scott & Gardner, 2015). The Triple P – Positive Parenting Program is designed to offer a range of supports that vary in their intensity (Sanders, 1999, 2012). The Group Triple P program, consisting of four group sessions, three individual sessions (typically conducted by phone), and an eighth group session, is offered to parents of children with behavioral challenges. A core concept in the delivery of Triple P is self-regulation: the capacity to self-regulate one's attention, thoughts, feelings, and behaviors emerges early in life and is associated with more positive outcomes; it includes skills in self-

efficacy, self-sufficiency, goal setting and self-monitoring, and personal agency, which then support independent problem-solving (Sanders, Turner, & Metzler, 2019). Self-regulation involves responding to cues from others about one's behavior to enable the individual to monitor and regulate his or her own behavior (Karoely, 1993). Triple P encourages parents to use cues from their child's behavior (e.g., noticing signs of escalating child behavior) and awareness of their own reactions (e.g., becoming frustrated) to regulate their parenting behavior (Sanders, 1999; Sanders & Mazzucchelli, 2013). Triple P promotes self-regulation in parents, with the aim that there will be cascading positive effects as a parent's self-regulation is strengthened, so does their child's capacity to self-regulate (Sanders et al., 2019). In each session, practitioners invite parents to take an active role in setting their own goals for change, selecting the strategies they wish to try, and gathering data to assess their usefulness.

Measurement of Adherence in Evidence-Based Programs

Following the demonstration of program efficacy, evaluating effectiveness requires assessment of its implementation (Dozois et al., 2014). Effective implementation of efficacious programs requires monitoring both the implementation of service and the associated changes in client behavior (Dozois et al., 2014). A key issue is the extent to which a program is delivered in a manner that is consistent with its evidence base (Forgatch, Patterson, & DeGarmo, 2005). Fidelity to the evidence base, also referred to as treatment fidelity, implementation fidelity, and treatment integrity, is a multi-faceted construct that includes such elements as adherence, competence, and differentiation (Allen, Shelton, Emmons, & Linnan, 2017; Perepletchikova, Treat, & Kazdin, 2007). Adherence is the extent to which the program is delivered in accordance with the core elements of its evidence base (Forgatch, Patterson, & DeGarmo, 2005), competence refers to the skill with which the practitioner delivers the program, and

differentiation is the degree to which an intervention protocol can be consistently distinguished from other protocols (Schoenwald, Garland, Southam-Gerow, Chorpita, & Chapman, 2011). Each facet captures an individual feature that may have a unique impact on therapeutic change (Perepletchikova et al., 2007). Meta-analytic research has found that programs implemented with higher levels of adherence yield greater effect sizes than do those that deviate from their empirical base (Forgatch & DeGarmo, 2011). The focus of this study is on adherence.

There are multiple ways to assess adherence, including routine self-assessment by practitioners and ratings by other people. The Triple P program encourages practitioners to complete a session-specific checklist after each session (Sanders, Markie-Dadds, & Turner, 2000) indicating whether they completed each activity described in the manual for that session. Session checklists are designed to promote reflective practice. They do not address adjustments to the content's presentation and do not assess adherence to the program's underlying principles (e.g., self-regulation in Triple P). Observational measures of adherence to parenting programs include content checklists consistent with the program manual (e.g., Parent-Child Interaction Therapy; Lyon & Budd, 2010 and Triple P). Others also include a measure of process adherence (e.g., Parent Management Training – Oregon model; Knutson, Forgatch, Rains, & Sigmarsson, 2009); these code underlying processes present across sessions regardless of content and focus on how the program is delivered. Despite growing acknowledgement of the importance of measuring adherence (e.g., Meyers, Durlak, & Wandersman, 2012), the psychometric strength of its measures is rarely reported (Schoenwald & Garland, 2013).

Flexible Delivery and Adherence in Triple P

To optimize outcomes in the delivery of an evidence-based program practitioners are encouraged to deliver the program in a way that is both adherent and sensitive to client needs

(Kendall, Gosch, Furr, & Sood, 2008). Eyberg (2005) considered three ways that delivery of Parent-Child Interaction therapy could be adjusted. Practitioners are routinely encouraged to tailor the program by adjusting content or its delivery (e.g., teaching skills in a more concrete or abstract fashion) to suit clients' needs. Structural changes to the program's delivery or major content changes, such as adaptations (or changes to a standard protocol, often for a new population) or modifications (when the treatment developer implements universal changes to an existing protocol), are made by the developer, and tend not to be encouraged by practitioners. To guide practitioners in their delivery of Triple P, the program's developer distinguished between practitioner adaptations with a low or high risk of diminishing program efficacy (Mazzucchelli & Sanders, 2010; Sanders, 2012). Low risk adaptations may enhance program efficacy as they address client needs, whereas high risk adaptations are likely to reduce program efficacy. Low risk adaptations to program content include using pertinent examples; useful process adaptations include modifying program length to adjust to client needs. High risk content adaptations include removing or replacing program material; high risk process adaptations include a practitioner advising a parent to use or avoid particular parenting strategies (instead of promoting self-regulation through discussing the pros and cons of a particular strategy, the rationale behind its use, and exploring instances where it may be more useful).

Research by Lau and colleagues (2017) examining self-reported adaptations to five evidence-based programs implemented in children's mental health services (including Triple P) asked practitioners to indicate how frequently they made changes; researchers then categorized these as either augmenting (i.e., modifying how session content is presented, integrating supplemental material, and extending the length of the program) or reducing/reordering (i.e., removing or rearranging the order of components and condensing the program) adaptations. In

this context, augmenting adaptations are similar to Eyberg's (2005) definition of tailoring and echo Mazzucchelli and Sanders' (2010) recommendations for low risk variations. Reducing or reordering adaptations are comparable to Mazzucchelli and Sanders' (2010) description of high-risk variations and go beyond acceptable levels of tailoring (Eyberg, 2005). Across all five programs practitioners more frequently reported augmenting than reducing adaptations (Lau et al., 2017). Triple P provides practitioners with a tool to consider what information is presented (i.e., content adherence). A tool that allows them to reflect on their adherence to how they deliver it may be very beneficial in ensuring that the program is implemented both as effectively and flexibly as possible (Forehand, Dorsey, Jones, Long, & McMahon, 2010).

Multiple Assessment Perspectives

The benefits of examining multiple perspectives are well-established in services for children (Hunsley & Mash, 2018). A robust literature shows that agreement between raters is imperfect for rating children's behavior (e.g., De Los Reyes & Kazdin, 2005) and therapeutic alliance (e.g., McLeod, Southam-Gerow & Kendall, 2017). In a seminal meta-analysis conducted by Achenbach, McConaughy, and Howell (1987) of agreement across raters in multi-informant, multimethod assessments of child and adolescent psychopathology, correspondence across raters was consistently low to moderate (average Pearson $r = 0.28$; Achenbach et al., 1987; De Los Reyes et al., 2015). Despite this robust finding, researchers persist in viewing converging reports as a measure of reporting accuracy, treating discrepancies as error or an artifact of measurement (De Los Reyes, Thomas, Goodman, & Kundey, 2013).

In examining reports of program adherence by different raters, discrepancy is also common. For example, therapist and observational ratings of adherence to interventions for adolescent behavior problems were only weakly correlated (Hogue, Dauber, Lichvar, Bonek, &

Henderson, 2013). Herschell, Quetsche, and Kolko (2019) examined concordance between different raters in a family intervention and found low rates of agreement, with clinicians and caregivers reporting higher adherence than did expert raters. In both cases, researchers attributed these discrepancies to overestimation by clinicians and caregivers.

Measurement experts have proposed that discrepancies may reflect unique contributions influenced by the different perspectives of the raters (De Los Reyes, Henry, Tolan, & Wakschlag, 2009; Hunsley & Mash, 2018). In the field of child psychopathology, De Los Reyes and Kazdin (2005) proposed that discrepancies among raters may be, in part, related to the perspective from which the rater perceives the behavior. For example, when informants are asked to rate a behavior (e.g., parents reporting on a child's disruptive behavior), they may access information from memory that confirms its presence, potentially at the expense of disconfirming information. Further, De Los Reyes and colleagues (2017) proposed guidelines for interpreting results from multiple raters not only when they converge, but also when they diverge or may compensate for one another, as relying on convergence as a measure of adherence may neglect important information that may come from diverging or compensating results. From an adherence perspective, it may be helpful to consider the context from which practitioners, clients, supervisors, and/or observational raters report on program adherence.

We present preliminary adherence data in the delivery of Group Triple P, including practitioner self-reports and ratings on a subset of the data by a coding team. In line with previous research (e.g., Herschell et al., 2019), we anticipated that practitioners would report higher content and process adherence than would be rated by the coding team.

Method

Participant Characteristics

This study was part of a project evaluating the implementation of Triple P in Québec, Canada (Delawarde-Saïas, Gagné, Brunson, & Drapeau, 2018; Gagné, Thomas, Delawarde, Daly, & Brunson, 2017). Participants in this study ($N = 52$) were front line practitioners from 12 agencies trained and accredited in Group Triple P for parents of children under age 12. Across agencies, one to 18 practitioners delivered Group Triple P (per agency: $M = 4.25$; $SD = 5.02$; agency information unknown for one practitioner). A subset of practitioners ($n = 23$) consented to audio-record sessions to permit coding of their adherence ([Table 1](#)). Across the seven agencies represented in this subset, one to nine practitioners delivered Group Triple P (per agency: $M = 3.29$; $SD = 2.66$). Chi square and t-test analyses indicated that practitioners who submitted recordings did not differ demographically from those that did not. Demographic data for the full sample and subsample are presented in [Table 1](#).

Sampling Procedures

Self-report measurement of adherence. Throughout 2014 and 2015, 68 practitioners were trained in Group Triple P. They were followed for two years (January 2015 to December 2016) and invited to complete the self-report measure after each group session they delivered. Fifty-two practitioners completed self-reports of adherence for group sessions yielding a 76.5% participation rate among trained practitioners. A total of 75 groups were registered in the project; as each includes five group sessions, the total number of potential submissions of the PSRT is 375. Participants submitted 361 valid entries with full or partial adherence data (96.3% response rate) and are nested within several levels: each group program included five submissions of the self-report, some practitioners ran multiple groups, and many of the agencies had multiple practitioners delivering Group Triple P.

Coding ratings of adherence. Audio-recording of sessions for adherence coding required

approval by agency managers, practitioners, and parents. One agency did not consent to record sessions; lack of parental consent ($n = 12$ groups; 60 sessions), discomfort among practitioners or their agency ($n = 2$ groups; 10 sessions), and recording equipment failure or practitioner error ($n = 12$ sessions) account for the remaining unrecorded sessions. Twenty-three practitioners from seven agencies submitted 66 audio-recordings yielding a 48.0% participation rate in submitting audio-recordings across practitioners and accounted for 17.6% of all completed group sessions. Of these recordings, 63 were usable (in which the full session was recorded, and audio quality was sufficient for coding). The three submitted sessions with recording problems were excluded. All procedures were conducted in French; measures were translated into French for this study, and all procedures were approved by ethical review boards in Québec (the Integrated University Center for Health and Social Services – CIUSSS; #MP-CJQ-IU-13-017) and the University of Ottawa (#H03-15-08).

Measures

Self-report measurement of adherence. The Practitioner Session Reflection Tool (PSRT; Sheshko, Lee, & Gagné, 2015) is a 12 to 17-item tool that takes approximately 10 minutes to complete. Part one of the PSRT assessed adherence to program content; part two focuses on adherence to the self-regulation framework as a measure of process adherence. The number of items varies across sessions as each group session consists of a different number of session activities. Further, not all items on the process portion (part two) are applicable to the early group sessions (e.g., promoting generalization of strategies cannot occur before all strategies are introduced, thus item 8 is not applicable until later group sessions). A copy of the PSRT for the first group session is included as [Appendix A](#).¹ Copies of the PSRT for each

¹ Included as [Appendix G](#) in dissertation

session are available from the first author.

In part one of the PSRT, practitioners indicate whether they conducted each session component as described in the program manual, then they describe any changes by checking all that applied from: I (a) added elements, (b) removed elements, (c) adapted activity, (d) replaced an element with another, or (e) dropped activity, and then select the reason for this change (e.g., perceived irrelevance to the clients, insufficient time in session; categories were drawn from Asgary-Eden and Lee, 2011). At the end of the measure, the practitioner can describe in a text field the nature of the modifications. The PSRT yields frequency of reported changes to session content and reasons for these changes.

Part two of the PSRT consisted of eight items. Participants were asked to indicate on a three-point scale (from “not at all” scored zero to “a lot” scored two points) the extent to which s/he used the self-regulation framework with parents during the session. An additional option, “not applicable in this session,” was available. Example items included “I invited parents to develop their own parenting goals” and “I invited parents to select the strategies that they wanted to employ,” drawn from Sanders and Mazzucchelli (2013). The items applicable to each session were averaged to enable comparison across sessions and the average across sessions was calculated.

Coder ratings of adherence. Two measures of adherence were rated by a coding team: the Coder Session Review Tool (CSRT) and the Adherence Measure for Process Quality in Standard Triple P (AMPQ; Kirby & Sanders, 2014a, 2014b). All audio recordings were de-identified by research staff before files were uploaded to a secure shared drive for the coding team. The CSRT paralleled the measure completed by practitioners after each session. The AMPQ, a coding system created by the developer of Triple P, addressed processes common to

all sessions. Items of the AMPQ focused on collaborative interaction with parents and the use of the self-regulation framework to support parents in developing and pursuing their own realistic goals. Fifteen items were rated at a global level on a scale from “Not Present” (scored 1) to “Fully Present” (scored 4), yielding a total process adherence score (range: 15–60). Higher scores indicated higher use of the self-regulation framework. We calculated a percentage score that indicated practitioners’ use of opportunities to promote self-regulation. To our knowledge, this is the first published study to use this measure and prior psychometric data are not available.

Coding team. A three-person coding team consisting of the first author, a doctoral student in clinical psychology, and a senior undergraduate in psychology completed 30 hours of training in the coding system. Didactic training delivered by an accredited Triple P trainer (the second author) provided an overview of the Triple P program and the self-regulation model. The team then reviewed the CSRT and AMPQ, practiced and discussed coding as a group, then independently practiced coding. Once the team agreed at a rate of 80.0% across codes, coders rated recordings individually. Of these recordings, 20.0% were coded by at least one additional rater to calculate inter-rater reliability. The coding team met regularly to prevent coder drift, review questions and resolve discrepancies. The coder completed the CSRT as s/he listened to the audio-recording to assess adherence to session content. After listening to the recording, the coder completed measures of process adherence, as outlined in [Appendix B](#).²

Results

Feasibility of PSRT in Community Implementation

Across the two years of the project, practitioners conducted 75 groups and completed 370 PSRTs, yielding self-report adherence data for 98.7% of group sessions. Of these, nine

² Included as [Appendix H](#) in dissertation

submitted PSRTs did not include any adherence data; 361 included full or partial data and represented 96.0% of group sessions.

Reliability of Adherence Measures

Internal consistency. To evaluate the reliability of the adherence measures, we calculated the internal consistency of part two of the PSRT (self-regulation checklist) and the AMPQ using Cronbach's alpha. As a content checklist that changes every session, part one of the PSRT/CSRT did not lend itself to measures of internal consistency. Reliability analyses demonstrated "good" (i.e., $\alpha = .80 - .89$) levels of internal consistency (Hunsley & Mash, 2018) for practitioner self-reports (PRST; $n = 246$; $\alpha = .80$) and coder ratings (CRST; $n = 59$; $\alpha = .85$ and AMPQ; $n = 63$; $\alpha = .81$).

Interrater reliability. Intraclass correlations (ICCs) examining the absolute agreement across raters achieved good levels of inter-rater reliability (Hunsley & Mash, 2018): for the CRST content checklist (.89) and self-regulation rating scale (.81), and .88 across the rating system of the AMPQ.

Convergent Validity of the Coder Session Review Tool

The same rater completed part two of the CSRT and AMPQ for a given session, thus Spearman's correlation coefficient was used to account for this nonindependence of data. A Spearman's rho ($r_s = .72$; $N = 59$; $p = .01$) of the correlation between coder ratings of the CSRT and AMPQ demonstrated adequate convergent validity (Hunsley & Mash, 2018), indicating that they are tapping into a similar construct.

Multiple Informants of Adherence

Having established both the internal consistency and interrater reliability, the next step was to examine the correlation between the self-report and coder ratings of adherence. The

intraclass correlation (ICC; two-way mixed model, absolute agreement) between self-report and coder ratings of content adherence (PSRT and CSRT part one) was .40. This indicated a poor level of agreement (Hunsley & Mash, 2018) on whether session content was conducted as it was described in the manual or if it was modified. The correlation between coders and self-reports on part two of the PSRT and CSRT (two-way mixed model, absolute agreement ICC = .08) indicated no relation between ratings on practitioners' promotion of self-regulation.

Self-Report Measurement of Adherence

Content adherence (PSRT part one). The 361 SRTs yielded a total of 2862 entries for content adherence (as each PSRT included multiple session activities). The frequency of session activities reported “completed as described in the manual” and “modified” were tabulated to descriptively examine adherence. Overall, practitioners reported ([Table 2](#), top left panel) that they implemented the program’s content as described in the manual 79.2% of the time. The components most frequently reported as completed according to the manual were: (a) teaching parents how to progressively phase out use of the program in the last group session (98.5%); (b) setting the agenda at the beginning of each session (94.4%); (c) reviewing the content of the previous session (92.7%); and (d) reviewing parents’ progress throughout the program in session eight (91.3%). Practitioners reported modifications most often when they (a) established group rules (28.8%); (b) conducted the final evaluation (27.5%); and (c) guided parents through the second (27.2%) and third (25.7%) set of strategies in a given session.

The self-reported modifications (and reasons for modifying program content) are presented in [Table 2](#); the frequency of each modification type presented in the top right and the reasons for modifications in the bottom panel. Practitioners most frequently reported they adapted (43.5% of modifications) and removed (32.3%) content. Practitioners reported that they

dropped a session component entirely (15.8%) or added elements to a component (8.2%) less frequently, and rarely reported replacing an element of a session component (0.2%).

Practitioners primarily reported that they modified content because they did not have enough time in session (48.7%). Practitioners also frequently endorsed the “other” category (40.4%); this included descriptions of moving material to a different session, completing the activity as a large group discussion (instead of practicing in pairs or smaller groups), or because the parent had not completed the tasks at home necessary for the activity.

Process adherence (PSRT part two). Practitioners consistently reported a high level of process adherence to self-regulation: across individual items and sessions, practitioners rated that they promoted self-regulation in 80.0 to 95.0% of the available opportunities.

Coder Ratings of Adherence

Content adherence (CSRT part one). The 63 audio-recordings coded by the team yielded 505 entries of content adherence. Overall, the coding team rated practitioners as adhering to the manual approximately half of the time (49.9%); the most common modifications rated by coders were removing elements (51.5%), adapting elements (32.7%), and eliminating a session component entirely (24.4%). Coders rarely or never rated modifications as additions (0.5%) or content as replaced (0.0%). Coders frequently rated components at the beginning (e.g., program introduction, establishing group rules, and agenda setting) and end of a session (e.g., third or fourth set of strategies, final evaluation, and session wrap-up) as modified. Overall, coders rated content as modified more frequently than did practitioners (50.1% compared to 15.4%); practitioners more frequently reported adaptations (43.5% of reported modifications compared to 32.7%), whereas coders more often rated practitioners as removing elements (51.5% compared to 32.3%) or dropping an activity entirely (24.4% to 15.8%).

Process adherence (CSRT part two). Across the 63 sessions, coders rated process adherence to the self-regulation checklist as variable throughout the program, and rated practitioners as promoting self-regulation among parents in 30.0 to 55.0% of the available opportunities. Process adherence was rated highest in the sessions introducing positive parenting practices (#1: 55.0%) and wrapping up the group intervention (#8: 45.0%); lowest for the session focused on managing misbehavior (#3: 30.0%). Across items, coders most frequently rated practitioners as inviting parents to monitor their child's behavior (50.0%) and selecting the parenting strategies they wished to use (45.0%). Coders less frequently noted practitioners inviting parents to generalize strategies to different situations (25.0%).

Process adherence (AMPQ). Using the global measure of process adherence, coders recorded that practitioners used 58% of the available opportunities to engage in the skills overall, with little variability across sessions. Coders rarely rated skills as either "Not present" or "Fully present," and most consistently rated techniques as "Minimally" or "Moderately present." This suggests that practitioners were rated as regularly engaging in techniques to promote self-regulation, but that there were also opportunities that they did not use.

Discussion

In this study, 75% of trained practitioners delivered Triple P, which is comparable to the rates reported by Asgary-Eden and Lee (2011). The results of this study provide preliminary data for a new self-report measure of adherence to a parenting program that addresses both content and process adherence. As the sessions were recorded with audio (and not video), the coding team was unable to capture nonverbal communication that may have lent itself to rating process adherence (e.g., encouraging nods; gestures inviting parents to share). Further, as the qualitative field available to describe modifications was at the end of the measure (not after each session

activity), some modification descriptions were unclear due to insufficient information. It is also important to note that adherence ratings were not independent: practitioners submitted PSRTs for five group sessions, and many practitioners conducted groups with multiple sets of parents.

Further, some agencies had numerous practitioners conducting groups (and others had only one).

Despite these limitations, the results of this study contribute to the development of low-cost and low-burden tools to measure adherence and support implementation in community settings. First, as a high proportion of eligible practitioners completed the PSRT, this suggests that it can feasibly be used by practitioners in a community setting. Further, practitioners' self-report using the self-regulation checklist (PSRT part 2) was internally consistent.

Several considerations may help interpret the discrepancy between self-report and coder ratings. First, the PSRT was designed to support practitioners in tracking (i) adherence to program content and (ii) the degree to which they promoted self-regulation. The low correlation between self-reports and coder ratings is consistent with past research in adherence (e.g., Hogue et al., 2013; McLeod et al., 2017) demonstrating weak correlations between reports of the same circumstances. Rather than interpreting differences across raters as indicators of accuracy, it may be more helpful to consider them as distinct sources of information. Across multiple domains or problems, self-report and external raters yield discrepant results. For example, in parenting research, self-reported parenting behaviors are only weakly correlated ($r = .17$) with observer ratings (Hendriks, Van der Giessen, Stams, & Overbeek, 2018), reflecting a low level of agreement between raters (Hunsley & Mash, 2018). The correspondence between raters of adult depression is moderate (e.g., Carter, Frampton, Mulder, Luty, & Joyce, 2010; Cuijpers, Li, Hofmann, & Andersson, 2010) and each report uniquely contributes to predicting treatment outcome (e.g., Tada et al., 2014). In these fields, the benefits of self-monitoring are not

disregarded due to imperfect correlation with others; instead, discussion centres on the potential benefits each report may contribute (e.g., Uher et al., 2012).

In adherence research, coders tend to rate more strictly than do practitioners and past efforts have mainly explained discrepancy as practitioners' tendency to over-estimate adherence (e.g., Herschell et al., 2019). However, another explanation can be drawn from the field of child psychopathology. In the Attribution Bias Context model, De Los Reyes and Kazdin (2005) proposed that different informants attribute the causes of behavior to the context or environment in unique ways, and some informants' perspectives may be more heavily influenced by contextual factors. In this study, the external raters completed the CSRT while listening to the session; practitioners completed the PSRT afterwards. In completing the CSRT, coders had access to the whole session and focused specifically on assessing adherence. Practitioners, on the other hand, had to rely on their memory of the session as they managed the multiple demands of covering content and managing a group. It is possible that coders rated whether practitioners were seizing every opportunity to promote self-regulation and referenced the manual throughout coding to rate content modifications. Practitioners may have applied the principle of minimal sufficiency, reporting adherence as a reflection of covering the gist of a session component or an aspect of self-regulation.

Clinical Implications

The PSRT was designed to invite practitioners to track modifications to inform selection of low-risk adaptations and reflect on promotion of self-regulation. In Triple P, low-risk, minor adaptations are encouraged (such as modifying examples to clarify how strategies may best fit with a family's needs, or increasing the number of sessions to permit parents more time if required) and represent flexible implementation that incorporate the needs, characteristics, and

preferences of clients (Mazzucchelli & Sanders, 2010; Sanders, 2012). The PSRT expanded on existing adherence tools to support practitioners in tracking modifications that could help in intervention planning, implementation evaluation, and supervision. For example, tracking both content and process adherence could support practitioners and supervisors in identifying sessions (or activities) that are frequently modified (particularly by eliminating content or when a session becomes more directive than self-regulatory). In addition, in this study coders rated whether practitioners seized all available opportunities to promote self-regulation. As outsiders, they identified even more opportunities to engage in self-regulation than did practitioners.

Practitioners, however, possibly operating within the construct of minimal sufficiency, may have reflected on their sessions and identified sufficient opportunities to promote self-regulation to keep the processes firmly in mind. Further, the practitioners in this study were all relatively new to Triple P. It is possible that as their experience expands, they will increasingly identify even more opportunities to promote self-regulation. This raises interesting supervision opportunities: for example, as new practitioners are trained and begin to implement the program, perhaps a seasoned practitioner within the agency could periodically observe and code for content and process adherence. This information could then be used in peer supervision as a tool to discuss modifications and obstacles to adherence, to better understand the motivation for adaptations and generate solutions that suit their agency's needs and resources.

Future Research

These preliminary results highlight the need for future examination of adherence within the large-scale implementation of evidence-based programs such as Triple P. However, the current study did not examine the predictive validity of the PSRT so the link between adherence and clinical outcomes is unknown. It will be valuable to understand the threshold of adherence

that is required to produce positive effects. It is possible, for example, that although coders perceived missed opportunities to promote self-regulation, that the practitioners seized enough opportunities to promote self-regulation and ensure generalization of gains. Further, it will be beneficial to explore the differences in content adherence as compared to process adherence, as this may provide valuable information in identifying targets in training and supervision.

Examination of variables that may moderate content and process adherence will be beneficial. These could include parental adjustment, practitioner and agency characteristics, and session composition (e.g., sessions with more content may be more heavily impacted by both content reduction and increased directiveness undermining process adherence). Increased understanding of these may support practitioners in making low-risk content modifications and encourage process adherence. In future analyses it will be helpful to examine correlation across ratings (as the same practitioners submitted multiple adherence forms) using multilevel modeling. In addition, it may also be helpful to examine the rates of content and process adherence (as rated by self-report compared to external coders) in this sample compared to measures used in other parenting programs to further contextualize these results of the PSRT.

Finally, the PSRT was originally developed as both a planning and a reflection tool, although only the reflection tool was used in this study. In the future, it may be helpful to include the planning tool and evaluate its feasibility in a community setting. Further, it may strengthen the PSRT if modification categories are streamlined into those that augment (such as adaptations or additions) or reduce/diminish (such as eliminating content or reordering content in a way that changes its function) the program's implementation. This could simplify measure completion and bolster its usefulness in community settings. In line with the approach of Lau and colleagues (2017), this may support a more nuanced adherence evaluation while minimizing

burden.

Conclusions

This study, as part of a larger implementation of Triple P in community agencies, provided support for the feasibility, reliability, and convergent validity of a new self-report tool that assesses content and process adherence. This tool was designed to encourage practitioners to track adherence and modifications to the program's content and invites them to reflect on the degree to which they promote self-regulation. The comparison of ratings across practitioner self-report and coders yielded low correlations. This may highlight the different contexts from which informants identify and report on adherence, rather than a measure of reporting accuracy. Examination of these differences, whether between external raters and practitioners as in this study, or between practitioners and their supervisors or colleagues, may provide a useful springboard to discuss adherence to a parenting program as implemented in community agencies.

References

- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of crossinformant correlations for situational specificity. *Psychological Bulletin*, 101, 213–232. <http://dx.doi.org/10.1037/0033-2909.101.2.213>
- Allen, J. A., Shelton, R. C., Emmons, K. M., & Linnan, L. A. (2017). Fidelity and its relationship to implementation effectiveness, adaptation, and dissemination. In R. C. Brownson, G. A. Colditz, & E. K. Proctor (Eds.), *Dissemination and implementation research in health: Translating science to practice* (2nd ed., pp. 267–284). New York, NY: Oxford Scholarship Online. <http://dx.doi.org/10.1093/oso/9780190683214.003.0016>
- Asgary-Eden, V., & Lee, C. M. (2011). So now we've picked an evidence-based program, what's next? Perspectives of service providers and administrators. *Professional Psychology: Research and Practice*, 42, 169–175. <http://dx.doi.org/10.1037/a0022745>
- Barlow, J., Parsons, J., & Stewart-Brown, S. (2005). Preventing emotional and behavioural problems: The effectiveness of parenting programmes with children less than 3 years of age. *Child: Care, Health and Development*, 31, 33–42. <http://dx.doi.org/10.1111/j.1365-2214.2005.00447.x>
- Biglan, A., Flay, B. R., Embry, D. D., & Sandler, I. N. (2012). The critical role of nurturing environments for promoting human well-being. *American Psychologist*, 67, 257–271. <http://dx.doi.org/10.1037/a0026796>
- Carter, J. D., Frampton, C. M., Mulder, R. T., Luty, S. E., & Joyce, P. R. (2010). The relationship of demographic, clinical, cognitive and personality variables to the discrepancy between self and clinician rated depression. *Journal of Affective Disorders*,

124, 202–206. <http://dx.doi.org/10.1016/j.jad.2009.11.011>

Clément, M.-E., Gagné, M.-H., & Lee, C. M. (2014). Understanding and preventing child maltreatment. *Canadian Journal of Community Mental Health, 33*, 1–12.

<http://dx.doi.org/10.7870/cjcmh-2014-012>

Cuijpers, P., Li, J., Hofmann, S. G., & Andersson, G. (2010). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: A meta-analysis. *Clinical Psychology Review, 30*, 768–778.

<http://dx.doi.org/10.1016/j.cpr.2010.06.001>

Delawarde-Saïas, C., Gagné, M.-H., Brunson, L., & Drapeau, S. (2018). Implementing a multilevel prevention strategy under an intersectoral partnership: The case of the Triple P Program. *Children and Youth Services Review, 88*, 170–179.

<http://dx.doi.org/10.1016/j.chilyouth.2018.03.009>

De Los Reyes, A., Augenstein, T. M., & Aldao, A. (2017). Assessment issues in child and adolescent psychotherapy. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (3rd ed., pp. 537–554). New York, NY: Guilford Press.

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multiinformant approach to assessing child and adolescent mental health. *Psychological Bulletin, 141*, 858–900.

<http://dx.doi.org/10.1037/a0038498>

De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology, 37*, 637–652. <http://dx.doi.org/10.1007/s10802-009-9307-3>

- De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, 131, 483–509. <http://dx.doi.org/10.1037/0033-2909.131.4.483>
- De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. A. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, 9, 123–149. <http://dx.doi.org/10.1146/annurev-clinpsy-050212-185617>
- Dozois, D. J. A., Mikail, S., Alden, L. E., Bieling, P. J., Bourgon, G., Clark, D. A., . . . Johnston, C. (2014). The CPA presidential task force on evidence-based practice of psychological treatments. *Canadian Psychology*, 55, 153–160. <http://dx.doi.org/10.1037/a0035767>
- Eyberg, S. M. (2005). Tailoring and adapting Parent-Child Interaction Therapy to new populations. *Education and Treatment of Children*, 28, 197–201.
- Forehand, R., Dorsey, S., Jones, D. J., Long, N., & McMahon, R. J. (2010). Adherence and flexibility: They can (and do) coexist. *Clinical Psychology: Science and Practice*, 17, 258–264. <http://dx.doi.org/10.1111/j.1468-2850.2010.01217.x>
- Forgatch, M. S., & DeGarmo, D. S. (2011). Sustaining fidelity following the nationwide PMTO implementation in Norway. *Prevention Science*, 12, 235–246. <http://dx.doi.org/10.1007/s11121-011-0225-6>
- Forgatch, M. S., Patterson, G. R., & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of competent adherence to the Oregon model of parent management training. *Behavior Therapy*, 36, 3–13. [http://dx.doi.org/10.1016/S0005-7894\(05\)80049-8](http://dx.doi.org/10.1016/S0005-7894(05)80049-8)
- Gagné, M.-H., Thomas, F., Delawarde, C., Daly, S., & Brunson, L. (2017). *L'implantation du programme Triple P sur le territoire du CLSC Mercier-est/Anjou: Une histoire de cas* [The

- implementation of the Triple P program in the CLSC Mercier-est/Anjou territory: A case story]. Québec City, Québec, Canada: Chaire de partenariat en prévention de la maltraitance, Université Laval.
- Gershoff, E. T., Goodman, G. S., Miller-Perrin, C. L., Holden, G. W., Jackson, Y., & Kazdin, A. E. (2018). The strength of the causal evidence against physical punishment of children and its implications for parents, psychologists, and policymakers. *American Psychologist*, 73, 626–638. <http://dx.doi.org/10.1037/amp0000327>
- Hendriks, A. M., Van der Giessen, D., Stams, G. J. J. M., & Overbeek, G. (2018). The association between parent-reported and observed parenting: A multi-level meta-analysis. *Psychological Assessment*, 30, 621–633. <http://dx.doi.org/10.1037/pas0000500>
- Herschell, A. D., Quetsch, L. B., & Kolko, D. J. (2019). Measuring adherence to key teaching techniques in an evidence-based treatment: A comparison of caregiver, therapist, and behavior observation ratings. *Journal of Emotional and Behavioral Disorders*. Advance online publication. <http://dx.doi.org/10.1177/1063426618821901>
- Hogue, A., Dauber, S., Lichvar, E., Bobek, M., & Henderson, C. E. (2015). Validity of therapist self-report ratings of fidelity to evidence-based practices for adolescent behavior problems: Correspondence between therapists and observers. *Administration and Policy in Mental Health*, 42, 229–243. <http://dx.doi.org/10.1007/s10488-014-0548-2>
- Hunsley, J., & Mash, E. J. (2018). 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* (2nd ed., pp. 3–16). New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/medpsych/9780190492243.003.0001>
- Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of*

Psychology, 44, 23–52. <http://dx.doi.org/10.1146/annurev.ps.44.020193.000323>

Kendall, P. C., Gosch, E., Furr, J. M., & Sood, E. (2008). Flexibility within fidelity. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47, 987–993.

<http://dx.doi.org/10.1097/CHI.0b013e31817eed2f>

Kirby, J. N., & Sanders, M. R. (2014a). *Adherence measure for the process quality in standard Triple P*. Brisbane, Queensland, Australia: Parenting and Family Support Centre, The University of Queensland.

Kirby, J. N., & Sanders, M. R. (2014b). *Instruction manual for the adherence measure for the process quality in standard Triple P*. St Lucia, Australia: Parenting and Family Support Centre, The University of Queensland.

Knutson, N. M., Forgatch, M. S., Rains, L. A., & Sigmarsson, M. (2009). *Fidelity of Implementation Rating System (FIMP): The manual for PMTO*. Eugene, OR: Implementation Sciences International, Inc.

Lau, A., Barnett, M., Stadnick, N., Saifan, D., Regan, J., Wiltsey Stirman, S., . . . Brookman-Frazee, L. (2017). Therapist report of adaptations to delivery of evidence-based practices within a system-driven reform of publicly funded children's mental health services. *Journal of Consulting and Clinical Psychology*, 85, 664–675.

<http://dx.doi.org/10.1037/ccp0000215>

Lyon, A. R., & Budd, K. S. (2010). A community mental health implementation of Parent–Child Interaction Therapy (PCIT). *Journal of Child and Family Studies*, 19, 654–668.

<http://dx.doi.org/10.1007/s10826-010-9353-z>

Mazzucchelli, T. G., & Sanders, M. R. (2010). Facilitating practitioner flexibility within an empirically supported intervention: Lessons from a system of parenting support. *Clinical*

Psychology: Science and Practice, 17, 238–252. <http://dx.doi.org/10.1111/j.1468-2850.2010.01215.x>

McLeod, B. D., Southam-Gerow, M. A., & Kendall, P. C. (2017). Observer, youth, and therapist perspectives on the alliance in cognitive behavioral treatment for youth anxiety.

Psychological Assessment, 29, 1550–1555. <http://dx.doi.org/10.1037/pas0000465>

Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology*, 50, 462–480. <http://dx.doi.org/10.1007/s10464-012-9522-x>

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, 75, 829–841. <http://dx.doi.org/10.1037/0022-006X.75.6.829>

Rothenberg, W. A., Weinstein, A., Dandes, E. A., & Jent, J. F. (2019). Improving child emotion regulation: Effects of Parent–Child Interaction Therapy and emotion socialization strategies. *Journal of Child and Family Studies*, 28, 720–731. <http://dx.doi.org/10.1007/s10826-018-1302-2>

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. <http://dx.doi.org/10.1023/A:1021843613840>

Sanders, M. R. (2012). Development, evaluation, and multinational dissemination of the Triple P-Positive Parenting Program. *Annual Review of Clinical Psychology*, 8, 345–379. <http://dx.doi.org/10.1146/annurev-clinpsy-032511-143104>

- Sanders, M. R., Markie-Dadds, C., & Turner, K. M. T. (2000). *Practitioner's manual for standard Triple P*. Brisbane, Australia: Families International Publishing.
- Sanders, M. R., & Mazzucchelli, T. G. (2013). The promotion of self-regulation through parenting interventions. *Clinical Child and Family Psychology Review*, 16, 1–17.
<http://dx.doi.org/10.1007/s10567-013-0129-z>
- Sanders, M. R., Turner, K. M. T., & Metzler, C. W. (2019). Applying self-regulation principles in the delivery of parenting interventions. *Clinical Child and Family Psychology Review*, 22, 24–42. <http://dx.doi.org/10.1007/s10567-019-00287-z>
- Schoenwald, S. K., & Garland, A. F. (2013). A review of treatment adherence measurement methods. *Psychological Assessment*, 25, 146–156. <http://dx.doi.org/10.1037/a0029715>
- Schoenwald, S. K., Garland, A. F., Southam-Gerow, M. A., Chorpita, B. F., & Chapman, J. E. (2011). Adherence measurement in treatments for disruptive behavior disorders: Pursuing clear vision through varied lenses. *Clinical Psychology: Science and Practice*, 18, 331–341. <http://dx.doi.org/10.1111/j.1468-2850.2011.01264.x>
- Scott, S., & Gardner, F. (2015). Parenting programs. In A. Thapar, D. S. Pine, J. F. Leckman, S. Scott, M. J. Snowling, & E. Taylor (Eds.), *Rutter's child and adolescent psychiatry* (6th ed., pp. 483–495). London, England: Wiley-Blackwell.
<http://dx.doi.org/10.1002/9781118381953.ch37>
- Sheshko, D. M., Lee, C. M., & Gagné, M.-H. (2015). *Practitioner Session Reflection Tool*. Ottawa, Ontario, Canada: Family Psychology Laboratory, The University of Ottawa.
- Tada, M., Uchida, H., Suzuki, T., Abe, T., Pollock, B. G., & Mimura, M. (2014). Baseline difference between patients' and clinicians' rated illness severity scores and subsequent outcomes in major depressive disorder: Analysis of the sequenced treatment alternatives to

relieve depression data. *Journal of Clinical Psychopharmacology*, 34, 297–302.

<http://dx.doi.org/10.1097/JCP.0000000000000112>

Taylor, W. D., Asgary-Eden, V., Lee, C. M., & LaRoche, K. J. (2015). Service providers'

adherence to an evidence-based parenting program: What are they missing and why?

Journal of Child and Family Studies, 24, 50–56. <http://dx.doi.org/10.1007/s10826-013-9812-4>

Uher, R., Perlis, R. H., Placentino, A., Dernovšek, M. Z., Henigsberg, N., Mors, O., . . . Farmer,

A. (2012). Self-report and clinician-rated measures of depression severity: Can one replace the other? *Depression and Anxiety*, 29, 1043–1049. <http://dx.doi.org/10.1002/da.21993>

Table 1*Practitioner characteristics (N = 52) who delivered Group Triple P*

	Total (N = 52)		Submitted Audio (n = 23)	
Sex				
Male	4	(8.0%)	2	(8.7%)
Female	47	(90.0%)	20	(87.0%)
Missing	1	(1.9%)	1	(4.3%)
Highest education				
Post-secondary diploma	17	(33.0%)	10	(43.5%)
Post-secondary degree	23	(44.0%)	9	(39.1%)
MA or PhD	8	(15.0%)	3	(13.0%)
Missing	4	(8.0%)	1	(4.3%)
Sector				
Health & social services agency	39	(75.0%)	19	(82.6%)
Community organization	9	(17.0%)	3	(13.0%)
Missing	4	(8.0%)	1	(4.3%)
Agency and submission rates				
Mean (SD) number of practitioners per agency	4.25	(5.02)	3.29	(2.66)
Range of practitioners across agencies	1 – 18		1 – 9	
Total number of groups registered	75			
Total number of sessions	375			
Total number PSRTs submitted	370	(98.7%)		
Total number valid PSRTs	361	(96.3%)	230	(62.2%)
Total number audio recordings submitted			66	(17.6%)
Total number valid audio recordings			63	(16.8%)

Table 2*Content adherence: Self-reported frequency and type of modifications, PSRT part one (N = 361)*

Session component	Modified (% selected)	Modification Type (% selected)				
		Adapted	Removed	Dropped activity	Added	Replaced
Program introduction	13.7	20.0	60.0	20.0	.0	.0
Establish group rules	28.8	42.9	38.1	14.3	4.8	.0
Setting agenda	5.0	33.3	16.7	16.7	33.3	.0
Review last session	6.4	14.3	28.6	7.1	50.0	.0
Homework review	12.2	54.3	22.9	17.1	5.7	.0
Strategies: Set 1	14.9	60.4	30.2	7.6	.0	1.9
Strategies: Set 2	27.2	47.4	35.1	16.5	1.0	.0
Strategies: Set 3	25.7	42.4	39.1	17.4	.0	1.1
Strategies: Set 4	19.0	51.2	39.0	7.3	2.4	.0
Phase out program	1.5	100.0	.0	.0	.0	.0
Review progress	5.8	75.0	.0	25.0	.0	.0
Final evaluation	27.5	15.8	42.1	42.1	.0	.0
Session Wrap-up	12.3	9.1	68.2	13.6	9.1	.0
Total (2862 entries)	15.4	43.5	32.3	15.8	8.2	.2
* Note: a total of 159 entries (5.5%) could not be used due to insufficient information						
Reason for Modification		% selected				
1. I did not have enough time in session		48.7				
2. This activity was not relevant for parents in this group		2.3				
3. Forgot to do the activity		2.4				
4. Have found that the activity did not help parents in the past		.0				
5. I did not have the necessary material		1.8				
6. I did not think parents would be comfortable doing the activity		1.1				
7. I did not think parents would understand the activity		.0				
8. I was not comfortable doing the activity		.5				
9. I did not feel trained well enough		.0				
10. Was instructed not to do activity by my supervisor		.3				
11. I did not think it fit my client's culture		.0				
12. Other (please describe)		40.4				
13. Multiple reasons selected		1.1				
14. No reason selected (despite reporting modification)		1.3				

Integrated Discussion

The primary objective that connects the two studies in this dissertation is to inform evidence-based supports for diverse parents. First, a meta-analysis was conducted to update the literature on adjustment among immigrant and refugee children post-migration by synthesizing research that used psychometrically sound measures of adjustment. Results are presented separately by rater. In addition, this meta-analysis extended past work with adult migrants by examining parental adjustment and parenting. This meta-analysis was intended to inform parenting supports by building on past reviews of adjustment among migrant youth and adults to provide a quantitative comparison of immigrant and refugee samples, to synthesize what is known about the adjustment of parents specifically, and of their parenting post-migration.

The second study focused on the development and evaluation of a self-report measure of adherence to an evidence-based parenting program that is used with diverse families. Expanding on existing adherence tools, I developed a self-report measure, the Practitioner Session Reflection Tool (PSRT; Sheshko, Lee, & Gagné, 2015) to support practitioners' adherence to both the session content and the primary process of self-regulation theorized to underpin the Triple P – Positive Parenting Program (Sanders, 1999, 2012). Practitioners participating in a randomized controlled trial (RCT) in Quebec were asked to complete this measure after each group session. A coding team rated adherence based on recordings of sessions. Thus, this study provides a multi-informant and multimethod evaluation of adherence. Preliminary psychometric properties were assessed for evidence of the tools' reliability (by evaluating interrater reliability and internal consistency) and convergent validity (between two adherence tools). In the following sections, I present a review of findings and their integration with existing research, a discussion of limitations, implications for research and practice, future directions, and final

conclusions.

Findings: Adjustment and Parenting in Immigrant and Refugee Families

In the first study I conducted an updated meta-analysis including studies published up until the end of January 2020. This meta-analysis includes 25 studies that have not been included in prior meta-analyses; approximately one-third of these studies were published following the database searches of previous reviews. Following previous research, I hypothesized that immigrant/refugee youth and parents would report more adjustment and parenting problems than the normative comparison groups, and that results would vary by migrant group, raters, and adjustment constructs.

Adjustment among Immigrant and Refugee Children and Adolescents

Overall, migrant youth described themselves as having more symptoms of anxiety than do normative comparison groups. Migrant youth also reported greater self-esteem, overall problems, and lower symptoms of depression and trauma than normative samples; however, the confidence intervals for these estimated summary effect sizes included zero. Parents reported their children experiencing greater overall and internalizing, but fewer externalizing, problems than normative samples. The single study of teacher-rated overall problems indicated fewer problems among migrant youth compared to normative data, however, the confidence interval included zero. These overall findings including the most recent studies replicate those of previous meta-analyses and systematic reviews of immigrant (Curtis, Thompson, & Fairbrother, 2018; Dimitrova, Chasiotis, & van de Vijver, 2016) and refugee (Blackmore et al., 2020) children and adolescents and reinforce that migrant youth may be at increased risk of experiencing higher levels of problems following migration compared to their nonmigrant peers. These results are explored in greater detail when I review differences between migrant groups

below.

Previous reviews tended to consider only immigrant (Dimitrova et al., 2016) or refugee (Kien et al., 2019) youth individually, or did not provide direct comparisons (Curtis et al., 2018). Given the significant variability across studies, I examined results separately for each migrant group. This revealed some consistencies and some differences between migrant groups in relation to normative comparison samples. Both immigrant only and refugee only samples reported higher anxiety than normative groups, reinforcing results of past reviews that focused exclusively on studies using psychometrically strong measures of adjustment (e.g., Kien et al., 2019). Among immigrant youth self-reports, the confidence intervals for all other constructs included zero. Consistent with past reviews using psychometrically strong measures (e.g., Bronstein & Montgomery, 2011) refugee youth also self-reported more depressive symptoms than normative samples. The confidence intervals for the estimates of self-esteem, overall problems, and trauma included zero. Once again, consistent with other reviews that included only studies that used psychometrically sound measures (e.g., Kien et al., 2019) a combined sample of immigrant/refugee youth self-reported greater levels of trauma than that reported in normative samples. Parent reports indicated greater overall problems among both immigrant and refugee samples than comparison groups; parents also reported greater internalizing, but lower externalizing, problems in a combined sample of immigrant/refugee youth compared to normative data. Finally, teacher-reports were limited to refugee youth, and indicated fewer overall problems compared to the normative sample; however, the confidence interval included zero.

Taken together, greater self-reported anxiety and parent-reported overall problems among both immigrant and refugee youth (compared to normative samples) emerged as robust findings

across migrant groups. The confidence intervals for these estimates did not include zero and the studies were methodologically high quality, and both factors increase confidence in my results. The interpretation of the other adjustment constructs is limited by the availability of a single sample per construct, and/or because the confidence interval of the summary estimate included zero. These factors necessitate caution in the subsequent interpretation of results, as these findings provide equivocal evidence of differences in adjustment between immigrant or refugee youth adjustment and normative comparisons. Potential explanations are explored below, first across all informants of child adjustment among immigrant samples, then refugee, and combined groups.

Results of the current meta-analysis showed that the only adjustment construct for which immigrant youth self-reported greater problems than comparison groups is anxiety, replicating past reviews that examined anxiety among immigrant youth (Curtis et al., 2018). Parent reports among immigrant youth were limited to one study examining overall problems, and results were consistent with past research (Dimitrova et al., 2016), indicating more problems among immigrant youth than normative comparisons. The summary effect size estimate of all other self-report constructs yielded confidence intervals that included zero, necessitating caution in their interpretation. Results of self-reported overall problems and trauma symptoms were in the expected direction, indicating greater problems among immigrant youth than normative comparisons, however, they also self-reported fewer symptoms of depression than do nonmigrants. This contradicts the finding of increased depressive symptoms among immigrant youth reported by Curtis and colleagues (2018) who did not limit their analysis to studies that used measures with evidence of reliability and validity, and the measures of the included studies were not described. It is possible that my results differed due to the use of more stringent

inclusion criteria for adjustment measures.

Refugee youth self-reported greater symptoms of anxiety and depression than both immigrant and normative comparison samples, replicating findings from past reviews limited to studies using psychometrically strong measures (e.g., Blackmore et al., 2020; Bronstein & Montgomery, 2011; Kien et al., 2019). In the current review there were single studies that addressed self-reported self-esteem, self-reported trauma, and teacher-reported overall problems, and two studies of self-reported overall problems in refugee youth; for all these studies the confidence interval includes zero. Numerous reviews have reported higher rates of trauma symptoms among refugee youth than comparison groups (e.g., Tam, Houlihan, & Melendez-Torres, 2017), or greater mental health problems than nonmigrants (e.g., Porter & Haslam, 2005). These reviews, however, employed less stringent inclusion criteria as studies were not limited to those that used measures with evidence of reliability and validity (Porter & Haslam, 2005; Tam et al., 2017), or measures were not described (Curtis et al., 2018). Further, my findings are discrepant from past reviews that used psychometrically strong measures to assess PTSD (e.g., Bronstein & Montgomery, 2011). However, studies in the review conducted by Bronstein and Montgomery (2011) that compared PTSD symptoms among refugee youth with a nonrefugee comparison group were limited to two studies, one of which focused on unaccompanied minors, the other included both accompanied and unaccompanied youth. Consistent with past reviews, it is possible that these unaccompanied youth reported higher levels of PTSD symptoms than accompanied refugee youth (Kien et al., 2019). In sum, it is possible that the exclusive use of psychometrically strong measures of adjustment and retaining only refugee youth accompanied by their parents may account for differences between the findings in my meta-analysis and previous reviews. Additional potential explanations are

explored below.

Reports of trauma in the included sample were inconsistent with the hypothesis that refugee youth exhibit more problems across adjustment domains, and contradict past syntheses that reported greater symptoms of PTSD or trauma among refugee youth (e.g., Fazel, Wheeler, & Danesh, 2005). Further, the confidence interval included zero. In attempting to understand this discrepancy it is important to consider the population against which the refugee group is compared. Previous meta-analyses that assessed trauma using standardized measures compared rates of disorder in refugee samples to worldwide prevalence rates (e.g., Blackmore et al., 2020) or only had a handful of studies that included a nonmigrant comparison group (e.g., Kien et al., 2019). In my meta-analysis, the measure used in the included study to assess symptoms of trauma among refugee youth recruited the normative comparison group from a high-conflict area (Post-traumatic Stress Symptoms in Children; Ahmad, Sundelin-Wahlsten, Sofi, Qahar, & von Knorring, 2000). It is possible that the comparison group for this measure was similar to the included refugee sample. This interpretation is consistent with past meta-analyses attempting to understand variability across refugee samples, in that comparison groups who may have been exposed to high levels of conflict or violence may report adjustment challenges at rates comparable to refugee samples (e.g., Porter & Haslam, 2005). This highlights the importance of considering the comparison group used when examining refugee adjustment.

Finally, in one of the two studies that combined immigrant and refugee youth, parents described their children as experiencing fewer externalizing (but more internalizing) problems than nonmigrant peers (Aroian et al., 2009). This is inconsistent with a past meta-analysis of externalizing problems among immigrant youth (Dimitrova et al., 2016). This past meta-analysis, however, did not describe the measures used and the use of psychometrically sound

measures was not an inclusion criterion. In addition, this meta-analysis did not present results separately by rater and it is unclear whether this may also have contributed to the discrepancy with the current study. Another possible explanation may relate to characteristics of the community from which the families were recruited: all families were recruited from a culturally similar community in one city. It is also possible that the presentation of externalizing behaviour is discouraged within the context of community expectations of behaviour, which may have contributed to lower reported rates of externalizing problems (Aroian et al., 2009; Dwairy, 2006).

Adjustment among Immigrant and Refugee Parents

Among parents, results were consistent with hypotheses for all constructs of adjustment. Across samples of either immigrant or refugee parents, results were consistently in the hypothesized direction, and the effect sizes were consistently large. These results based on an expanded set of studies replicated findings from previous reviews of adult migrants (Fazel et al., 2005; Porter & Haslam, 2005) and of perinatal adjustment among mothers who have migrated (Anderson, Hatch, Comacchio, & Howard, 2017; Falah-Hassani, Shiri, Vigod, & Dennis, 2015; Fellmeth, Fazel, & Plugge, 2017). Furthermore, they extend previous research of adult migrants to describe adjustment among both mothers and fathers. Mirroring the pattern found in studies of child adjustment, estimates of parental adjustment were significantly heterogeneous, and I examined summary estimates for each migrant group separately.

When examined individually, results continued to be mainly consistent with past reviews, however, the confidence interval of the estimate for higher overall problems in one sample of immigrant parents included zero and so does not replicate. Although most of the findings indicated greater problems across both immigrant and refugee samples, one exception consisted

of a single study that combined immigrant and refugee mothers. These results indicated that migrant mothers reported fewer symptoms of depression than do the normative comparison group (Aroian et al., 2009). As described above with regards to parental reports of externalizing problems in a sample of combined immigrant and refugee youth, this sample was recruited from a culturally similar community with strong community connections. Results of this study suggested that mothers higher in social support reported significantly fewer depressive symptoms than mothers who reported lower social support. Low levels of social support is a well-recognized risk factor in the development of mental health difficulty (Falah-Hassani et al., 2015; Fellmeth et al., 2017), and it is possible that the significant relation between levels of social support and maternal distress in the current study served to protect mothers from developing higher rates of depressive symptoms. These results point to the benefits of assessing potential moderators of adjustment – both risk and protective – to contextualize results and attempt to clarify differences between samples.

Parenting among Immigrant and Refugee Families

The second goal of the meta-analysis was to evaluate parenting as reported among immigrant and refugee families following migration. I found that few studies had assessed this important construct. As increased parental distress may negatively impact parenting (e.g., Barlow, Smailagic, Huband, Roloff, & Bennett, 2014), and immigrant and refugee adults may be at increased risk of adjustment problems following migration (e.g., Porter & Haslam, 2005), I hypothesized that immigrant and refugee families would report greater problems than normative comparison groups. Further, I hypothesized that refugee families would report higher problems than immigrant families. Results of the few included studies were inconsistent with this hypothesis; further, the small number of included studies precluded comparison between migrant

groups.

Immigrant only and combined samples reported more positive family functioning among migrant samples than among normative comparison groups. Interestingly, one of these studies that also assessed youth adjustment using a psychometrically strong measure (Slonim-Nevo, Mirsky, Rubinstein, & Nauck, 2009) found that youth self-reported fewer overall problems than the normative sample. It is possible in this sample that lower rates of child adjustment problems were related to positive family functioning (or that these may work transactionally), a finding that is consistent with past research examining family functioning and child behaviour (e.g., Biglan, Flay, Embry, & Sandler, 2012). Further, refugee youth reported fewer problematic parenting behaviours than do normative comparisons (Mellor, Renzaho, Swinburn, Green, & Richardson, 2012); the confidence interval for parent reports includes zero, providing equivocal evidence of differences between refugee and normative samples. It is important to note that none of the studies included in the parenting meta-analysis also included psychometrically sound measures of parental adjustment, signifying a gap in examining the interrelation between child and parental adjustment and parenting in these diverse families.

Taken together, some of my results based on methodologically strong studies were consistent with previous reviews in finding that immigrant and refugee youth reported higher anxiety than normative comparison groups; in addition, my data replicate the findings with respect to depressive symptoms among refugee youth. Further, parent reports indicated greater overall problems in both immigrant and refugee samples compared to normative samples, as well as greater internalizing problems in a combined sample of immigrant/refugee youth as compared to nonmigrant peers. We also found differences in terms of self-reported depression symptoms, as refugee youth reported greater symptoms than normative data, whereas immigrant youth

reported fewer symptoms than comparison groups. Further, immigrant only and combined immigrant/refugee samples reported greater trauma symptoms than normative samples, however, refugee youth reported less trauma symptoms than normative data. Parent self-reports of their adjustment were mainly consistent with past reviews, and my investigation of parenting behaviour and the parent-child relationship yielded very few studies that used psychometrically sound measures.

All included studies used psychometrically strong measures and were rated as high in methodological quality. Some of my results that were inconsistent with past reviews (lower self-reported depression among immigrant youth, higher self-reported self-esteem and lower trauma symptoms, and lower teacher-reported overall problems among refugee youth), and may have differed in the psychometric strength of the measures used to assess these constructs. For instance, past reviews including depression among immigrant youth (e.g., Curtis et al., 2018), and among refugee youth reviews examining self-esteem (Marley & Mauki, 2019), trauma (Curtis et al., 2018; Tam et al., 2017), and overall distress (e.g., Porter & Haslam, 2005) did not restrict their studies to those that used measures with evidence and reliability and validity (Fazel et al., 2012; Marley & Mauki, 2019; Porter & Haslam, 2005; Reed et al., 2012; Tam et al., 2017; van Os et al., 2016), or measures were not described (e.g., Curtis et al., 2018; Dimitrova et al., 2016). It is possible that the inclusion of psychometrically weak or untested measures could account for some of these differences in results.

Both meta-analyses of child/adolescent and parental adjustment which included recent studies updated and replicated past reviews indicating greater problems among immigrant and refugee families. Results, however, also highlight that it is imperative to clearly identify the participants' migrant status, as adjustment may differ between immigrants and refugees.

Presenting results that collapsed across both groups may inadvertently mask important differences in adjustment. Further, given the well-established differences between raters, it is essential to examine findings separately by rater (Achenbach, McConaughy, & Howell, 1987; De Los Reyes, Thomas, Goodman, & Kunder, 2013). Furthermore, results of my parenting meta-analysis suggest a dearth of studies examining parenting among immigrant and refugee families using psychometrically strong measures, and that it also may be helpful to assess for both child and parental adjustment and parenting variables to contextualize results.

Limitations

Results of the meta-analysis indicate that overall, migrant children, adolescents, and parents are at an increased risk of adjustment problems, although this varied across adjustment domains, migrant groups, and informants. The interpretation of results, however, is limited by several methodological issues. First, although examining adjustment separately for each migrant group and rater yielded important differences across groups and informants, this resulted in a small number of studies available for each analysis. These results could not be explored with moderator analyses due to insufficient available data. This limited interpretation of the variability in reported rates of adjustment. Second, some of the summary effect sizes yielded confidence intervals that include zero, limiting the interpretation of these results, and do not replicate past reviews. Third, very few studies were located that assessed parenting with psychometrically adequate measures among immigrant and refugee families, reinforcing a gap identified in previous systematic reviews (van Os, Kalverboer, Zijlstra, Post, & Knorth, 2016). Finally, the included studies rarely assessed both adjustment and parenting, making it impossible to determine whether adjustment accounted for differences in parenting.

Implications for Research and Practice

My research extends previous work examining adjustment among immigrants and refugees by updating the synthesis of literature among children and adolescents, includes both immigrants and refugees, and expands our understanding of adjustment among adult migrants by focusing specifically on those who are mothers or fathers. Further, my meta-analysis highlights the paucity of research examining parenting among migrant families using psychometrically sound measures. Results point to meaningful research implications centered on the importance of presenting adjustment data separately for migrant groups and implications of assessing moderating variables, evaluation of methodological quality, methods for generating quantitative comparisons between migrant and nonmigrant groups, and use of psychometrically adequate measures with diverse samples.

First, meta-analytic results indicated important differences between immigrant, refugee, and combined samples, bolstering the necessity of understanding differences within migrants. In my meta-analysis, there was an inadequate number of studies to conduct moderator analyses. In addressing this, one option is to systematically present results separately by migrant group and to compare groups where data permit to explore potential moderators of adjustment. However, the definition of immigrants and refugees is a somewhat arbitrary one. It is possible that the experiences of migrants occur on a continuum and it may be most fruitful to conduct research on the determinants of adjustment among migrants rather than to continue to focus on legally defined groups.

Second, past research (e.g., Fazel et al., 2005; Porter & Haslam, 2005) found that studies with greater methodological rigour yielded lower estimates of adjustment problems than did lower quality studies, suggesting that studies lower in quality may report inflated effects. My assessment of methodological quality indicated that most studies in my meta-analysis were

strong, increasing confidence in my results. In continuing to synthesize emerging research with diverse families, it is imperative to assess the methodological strength of included studies to increase the accuracy of estimates and bolster the strength of results. One component of methodological strength is the use of a measure with demonstrated evidence of reliability and validity (e.g., Downs & Black, 1998). Despite repeated calls for the use of psychometrically sound measures (e.g., Anderson et al., 2017; Blackmore et al., 2020), many studies continue to use weak, untested, or unnamed measures, and/or do not report on the reliability and validity of the measures' use in their sample. Further, review of the publication dates of studies excluded due to measurement issues did not suggest a trend of improvement; instead, an alarming proportion of these excluded studies were from more recent publications.

Third, using measures' norms as a comparison group enabled me to include studies that did not have comparison groups built into the study. Studies that did not follow the scoring as described in the manual were excluded. This permitted my retention of studies that used psychometrically adequate measures as they were designed and increased confidence in my results comparing migrant groups to the normative samples. This also indicated the frequency with which psychometrically strong measures are modified. The characteristics of comparison groups are also necessary considerations when assessing adjustment among migrants, particularly for measures of trauma: if the manual reported data from a clinic referred as compared to a community or general population sample, or a sample known to have been exposed to high levels of conflict or disruption, then no differences between the migrant sample and comparison group suggest that they are comparable to those requiring treatment for trauma. Earlier reviews (e.g., Anderson et al., 2017; Porter & Haslam, 2005) reported that comparison groups with levels of exposure to conflict or disruption similar to that of the refugee/immigrant

samples potentially yielded smaller differences in adjustment (or results in the opposite direction than expected). Integrating a description of the comparison group's characteristics into discussions about migrant adjustment is essential to interpretation of findings.

A noteworthy finding of the meta-analysis is the large number of studies that failed to meet inclusion criteria. Hundreds of studies were excluded because they did not use a measure we identified as being psychometrically adequate. Additional studies used a measure for which there is evidence of psychometric adequacy, but either scored the measure or presented data in a manner inconsistent with the norms. Further, many potentially eligible studies were excluded because they did not report on the reliability of measures used with immigrant and refugee samples or reported inadequate levels of reliability with their samples. This is an essential step when using psychometrically adequate measures, as psychometric adequacy achieved with the normative sample does not guarantee adequacy in subsequent samples (Hunsley & Mash, 2018). Overall, nearly 1,100 articles assessed for eligibility had problematic measurement issues. This suggests that a large proportion of studies in this area suffer from inadequate measurement of key constructs. These results indicated a crucial gap and reinforced the importance of reporting these data that could help to clarify variability across samples. In addition to research implications, study results highlight implications for service delivery to diverse families. The results of the meta-analysis suggest that although immigrant and refugee youth may be at heightened risk of developing adjustment problems, there is great variability across informants, the type of adjustment construct, and migrant group. Results also highlight the infrequency with which parenting constructs are assessed among these families using psychometrically sound measures.

Taken together, these results mirror the pattern in the general population, indicating that

there is a wide range in need for support. First, in assessing immigrant/refugee youth and their parents, it is important to consider the source of the rating. Results indicate that youth, parents, and external informants perceive things differently. Results of the meta-analysis also indicated variability across adjustment constructs between migrant groups. This speaks to the benefits of taking into consideration multiple adjustment constructs across family members to identify specific areas of resilience or difficulty that may interact with parenting, and to use this information to inform decision-making about supports.

Second, there may also be different periods of heightened vulnerability which may vary with migrant status: for example, it is possible that, in line with the “immigrant paradox,” that some immigrants who initially present with healthy adjustment may go on to develop difficulty over time (e.g., Vang et al., 2015, 2017), whereas refugee youth may arrive in higher levels of distress which remits over time as they gain greater distance from the pre-migratory events that precipitated their flight (e.g., Blackmore et al., 2020; Curtis et al., 2018). It is also possible that for the rater, the comparison group may be different at different stages. For example, a person fleeing an area of genocide may initially consider themselves fortunate to have escaped; post settlement they may become aware that the symptoms they regularly experience (such as vigilance, startle, or concentration difficulties) are different from the experiences of others in the host country. Further, preliminary results indicate that people do better when they settle in a community where they experience increased social support (Aroian et al., 2009) and peer acceptance (Kovacev & Shute, 2004), reiterating important avenues for supports. These results suggest that it is inaccurate to make assumptions based simply on immigrant or refugee status. It is also important to assess for factors that may exacerbate their risk or protect them, and to offer supports to bolster protective variables (such as promoting social support connections).

Each of these considerations has important implications for informing policies or decision-making for delivering evidence-based parenting supports to diverse families. In line with a public health framework of service delivery, it is essential to offer supports of varying intensities to best meet the needs of diverse families: a public health approach for parenting supports promotes the use of effective and supportive parenting practices at a population level, can accommodate multiple levels of functioning and risk, and requires flexibility to accommodate diverse needs (Sanders & Prinz, 2018). Operating within this framework may involve beginning with the least intensive level that emphasizes healthy living and information. If that is insufficient to meet a family's needs, then to offer services of increasing intensity (Sanders & Prinz, 2018). This is in line with the concept of minimal sufficiency, in offering the lightest touch needed for an individual family to pursue goals they identify as important. It will also be crucial to identify individuals who can serve as guides to support families: in navigating a new system to access information and supports and simultaneously promote the family's autonomy in their decision-making. It will be important for these guides to have a firm understanding of a family's individual needs, values, and characteristics to work collaboratively in identifying goals and providing information about available options. This is congruent with both the use and promotion of self-regulation (Mazzucchelli & Ralph, 2019; Sanders et al., 2019), as well as evidence-based program implementation that tailors services to an individual family's needs and characteristics (American Psychological Association, 2006, 2008; Canadian Psychological Association Task Force on Evidence-Based Practice of Psychological Treatments, 2012).

Future Research

Future research will be strengthened by integrating methodological recommendations

focused on the comparison groups used in quantifying adjustment and parenting in diverse families and presenting the results of psychometrically adequate measures separately by informant. Previous reviews have tended to examine only immigrant (Dimitrova et al., 2016; Falah-Hassani et al., 2015) or refugee (Kien et al., 2019; Fazel et al., 2005) samples, or if both were included direct comparisons were not offered (Anderson et al., 2017; Curtis et al., 2018). Although the migration experience may differ between migrant groups, some variables (such as exposure to violence) may render these groups more comparable (Porter & Haslam, 2005). Future research may benefit from including both immigrants and refugees in the same study, and to compare their experiences where data permit.

Further, it is important to integrate a nonmigrant comparison group more consistently to quantify the degree of difficulty experienced by immigrant and refugee children, adolescents, and parents following migration. It is also essential to describe the comparison group's characteristics to tease apart the potential effects of stressors that may be uniquely associated with the migration experience. If a nonmigrant comparison group is not feasible, then using the norms for a given measure is an option. However, this requires that the measure be used and scored precisely as it was with normative samples, and that results are presented in a way that aligns with the available norms (e.g., if norms are only available for each gender separately, and researchers collapse across their full sample, this study may be excluded from future meta-analyses) to facilitate inclusion in future systematic reviews and meta-analyses.

Finally, some past reviews have struggled to identify trends in adjustment among migrant youth (Stevens & Vollebergh, 2008), were limited to self-report (Bronstein & Montgomery, 2011; Curtis et al., 2018), or did not systematically differentiate between or define informants (Dimitrova et al., 2016; Marley & Mauki, 2019; Tam et al., 2017). As the differences between

raters are well-established (Achenbach et al., 1987; De Los Reyes et al., 2013), it is essential to clearly identify who is reporting on adjustment or parenting, in both primary articles and in systematic reviews or meta-analyses. Incorporating these recommendations could increase the likelihood of inclusion in future meta-analyses, resulting in a greater number of included studies which could permit more reliable moderator analyses.

The Public Health Approach: Working with Diverse Families

The results of my meta-analysis demonstrating variability across informants reiterate that children and parents perceive things differently. This highlights the importance of incorporating multiple perspectives – and to examine them separately – when working with families. As such, flexibility is required when working with families to accommodate multiple informants, and materials are needed that can help parents with varying levels of problems, from those experiencing comparatively minor disruptions to those who have experienced trauma and/or are at high levels of distress. A public health approach to parenting supports is based on the idea that parenting can be challenging for many families and that all parents could benefit from some level of support (Prinz & Sanders, 2007; Sanders & Prinz, 2018). In considering supports, many people are resilient in the face of adversity (e.g., systematic review of resilience among refugee youth conducted by Marley & Mauki, 2019), but families do not need to experience high levels of problems to benefit from parenting programs. This is in line with the potential needs identified in the meta-analysis, indicating varying degrees of adjustment problems across immigrant and refugee youth. The use of self-regulation is thus a core component in working with diverse families in parenting supports, as this encourages parents to reflect on their individual family's needs and goals and supports parents in their decision-making (Sanders et al., 2019). Given this key feature, it is essential that practitioners delivering parenting interventions

are supported in their promotion of self-regulation. To date, however, evaluation of the flexible implementation of a parenting program, both in terms of the program's content and in its promotion of self-regulation, is rarely evaluated as a self-report to adherence in parenting programs, and has yet to be done in the Triple P program.

Findings: Adherence

Adherence Measurement in Parenting Programs

The third goal of this dissertation was to develop and evaluate a new self-report measure of content and process adherence to the Triple P – Positive Parenting Program (Sanders, 1999, 2012) to support practitioners working with diverse families. The aims of the study were primarily exploratory, in first assessing the feasibility of the measure's use in a community implementation, and then to present preliminary psychometric evidence of the reliability and validity of two tools of adherence, using both self-report and observational ratings. Following this, I examined reported rates of adherence by both self-reports and external coders, and based on past research (Herschell, Quetsch, & Kolko, 2019; Hogue & Dauber, 2013), I hypothesized that practitioners would report higher rates of adherence than would coders.

Prior to exploring the results, I will first discuss the sample size of this study. Across the implementation of Group Triple P, each practitioner completed multiple self-reports of adherence. For adherence data to be independent, analyses would need to account for multiple levels: each practitioner (1) submitted up to five measures of adherence, one for each session with a particular group of parents; (2) may have conducted the program multiple times; (3) is grouped within a particular agency; and (4) there are multiple agencies involved. As such, depending on the analysis pursued, each measurement could require nesting within its implementation, practitioner, and agency.

The G*Power program (Faul, Erdfelder, Lang, & Buchner, 2007) is a well-recognized tool to calculate sample size in social and behavioral research. However, analyses that require nesting of this magnitude are too complex for a power analysis using G*Power. As an example, I completed a power analysis for a simpler analysis using logistic regression examining practitioners' level of self-regulation (as the independent variable) and the categorical dependent variable of likelihood of making low-risk versus high-risk adaptations to the program's content. To complete the power analysis, I used the z test Logistics regression in G*Power (version 3.1.9.7). A one-tailed test was used to accommodate a directional hypothesis (e.g., higher levels of self-regulation will be associated with low-risk modifications). An odds ratio of 1.8 was selected in order to improve the likelihood of identifying a significant effect, and the Power ($1-\beta$ err prob) was set at 0.80 (in line with Cohen's, 1988, recommendation for power). Results of the G*Power analysis are available in [Appendix I](#). A total of 119 participating practitioners was recommended, which is in line with other adherence evaluations of Triple P (e.g., Taylor, Asgary-Eden, Lee, & LaRoche, 2015).

The final sample size of the current study consisted of 52 practitioners who delivered Group Triple P and submitted self-reports of adherence. A total of 68 practitioners were trained in the program, yielding a 76.5% participation rate among trained practitioners. Of the 16 practitioners who declined participation, it is unclear whether they delivered the program (and declined to participate in the research project), if they were not actively delivering the program despite completing the training (and did not have materials to submit), or if they had left their position with the participating organization. These data were not available to the researcher. Of the participating practitioners, 23 also submitted audio-recordings to permit external ratings of adherence, yielding a 48% participation rate. Of those who declined to submit recordings,

information was available to the researcher to determine the number of group sessions that were impacted: one agency did not consent to record sessions for any group sessions; parents did not consent in 12 groups (impacting the recording of 60 sessions), individual practitioners or their agency were not comfortable with recording in two groups (10 sessions), and issues with recording impacted an additional 12 sessions. The above power analysis suggests that the number of practitioners in the current study may not provide adequate power for analyses. Although this sample size is smaller than studies that collected a single evaluation of adherence among parenting programs (such as a practitioner's most recent session: e.g., Lau et al., 2017; Taylor et al. 2015), it is comparable to studies that obtained repeated adherence measures from the same practitioners (e.g., Eames et al., 2008; Hogue & Dauber, 2013). Further, this power analysis does not account for repeated submissions of adherence reports by practitioners across each implementation of the program, totaling over 360 useable self-reports. The current study focused on descriptive results in the preliminary examination of this new self-report adherence tool; future analyses with this data set will require analyses that permit nesting data to account for multiple submissions by the same practitioner, as described in a subsequent section.

Consistent with expectations, participating practitioners had a high completion rate indicating that the tool could feasibly be used in a community implementation. Further, scores for both the self-report and observational measures of adherence exhibited good rates of interrater reliability and internal consistency. Ratings from the external coder version of the self-report tool and the observational process measure also demonstrated adequate convergent validity, indicating that both tools assessed a similar construct. These results provided preliminary support for the reliability and validity of this new self-report measure.

It is well-established that programs implemented with high levels of adherence produce

stronger effects than implementation that deviates from the program's evidence base (Forgatch & DeGarmo, 2011). Furthermore, evaluation of adherence is essential to accurately assess the efficacy of an intervention (Fixsen, Blase, Duda, Naoom, & Van Dyke, 2010; McArthur, Riosa, & Preyde, 2012). In Triple P, program adherence requires adherence to both session content and to promoting the underlying process of self-regulation. Competence, also captured under implementation fidelity, is the skill with which practitioners deliver the program (Schoenwald, Garland, Southam-Gerow, Chorpita, & Chapman, 2011). When practitioners implement Triple P, an important consideration in evaluating adherence is the concept of minimal sufficiency, which refers to offering the minimal support needed to encourage change. In Triple P, this reflects practitioners' use of the minimum intensity of support necessary to encourage parents' self-reflection on their goals, use of the strategies, and progress (Sanders, 1999, 2012; Sanders et al., 2019). In differentiating between lower levels adherence and the use of minimal sufficiency, it is crucial to consider whether the essential message or skill was communicated as it was intended and is it conducted in a self-regulatory manner. This is then distinguished from lower levels of competence if the core information is still presented but is delivered in a directive fashion; higher competence would occur when self-regulation is adeptly and frequently promoted with parents.

In terms of adherence rates, results were consistent with the hypothesis, as practitioners reported higher content adherence than that rated by coders (84.6% compared to 49.9%). In addition, practitioners reported more frequent promotion of self-regulation with parents than coders' rating: practitioners rated that they promoted self-regulation in 80.0 to 95.0% of the available opportunities, whereas coders rated between 30 – 55%. This resulted in poor levels of agreement between practitioners and coders and is explored in greater depth below. Across self-

reports, practitioners reported a high level of adherence. A comprehensive evaluation of adherence ought to consider variability in adherence rates across self-reports, and this requires data to be nested by implementation, practitioner, and agency. This was beyond the scope of the current project and its preliminary examination of a new adherence tool. It is an important area of future research, as described below. This study supported the development of a self-report tool of content and process adherence that can feasibly be used by practitioners in a community setting and reinforced the benefits of using multiple reports of adherence.

Limitations

In many areas of research, individuals tend to rate themselves more positively than do external raters (e.g., Hendriks, Van der Giessen, Stams, & Overbeek, 2018). In adherence research, practitioners often rate themselves as more adherent than do external coders, and these discrepancies are often explained as over-estimations among self-reports (Herschell et al., 2019; Hogue & Dauber, 2013). As described above, the level of agreement between self-report and coder ratings of adherence was low. This finding was consistent with past research comparing multi-informant ratings of adherence (e.g., Hogue & Dauber, 2013). There is a robust literature that suggests that self-evaluations tend to be higher than external ratings (e.g., Herschell et al., 2019). It may also point to a limitation in the measure's completion: it is possible that practitioners and external coders approached ratings from different perspectives using the construct of minimal sufficiency (Sanders, 1999, 2012).

From an adherence perspective, in considering the potential relation of minimal sufficiency to informants' ratings of adherence, it is first important to consider the different perspectives from which adherence ratings occurred in the RCT: coders referenced the manual throughout coding as they assessed the extent to which practitioners used opportunities to

promote self-regulation. Triple P practitioners, on the other hand, were managing the multiple demands of conducting a group and self-evaluated adherence after the session. Practitioner ratings may reflect their impressions of whether they promoted enough self-regulation.

Furthermore, adherence measurement in the RCT was one of many components that practitioners were asked to assess, and all tools were introduced by the research team's implementation support person, rather than by the researcher. As such, I did not have direct contact with practitioners to explain the PSRT and outline distinctions between reporting a session component completed as described in the manual or as modified. Thus, it is possible that practitioners' reports reflected covering the gist of a session component, whereas coders rated content adherence much more strictly. Further, among the options describing modifications, I was not able to describe the nuances between adding, removing, adapting, or replacing elements of session content, or eliminating a session component entirely. Practitioners reported adapting content more frequently than did coders, whereas coders rated removing content more frequently than did self-reports. Once again, the use of minimal sufficiency may be reflected in practitioners' reports in covering the key elements of session content, whereas coders may have coded these as removing content. This pattern may also apply to ratings of process adherence in using the self-regulation checklist (PSRT, part two): practitioners' reports may relate to promoting self-regulation often enough to keep the construct firmly in mind, whereas coders may have rated whether practitioners seized every available opportunity.

These differences, rather than interpreting them as measures of accuracy or inaccuracy between raters, may reflect variability in how raters conceptualize adherence using minimal sufficiency. To strengthen future use of the PSRT, it may be helpful to first expand instructions to describe the differences between completing content as described in the manual and

modifying its presentation. Further, consistent with adherence measurement across multiple parenting interventions (Lau et al., 2017), streamlining the PSRT (part one) into fewer modification categories may be beneficial. For example, it may be helpful to offer categories that reflect changes that either boost (such as adaptations or additions) or reduce/diminish (such as eliminating content or reordering content in a way that changes its function) the program's implementation and effectiveness to simplify its completion by practitioners.

The adherence study presented preliminary results of psychometric adequacy of a self-report measure of content and process adherence, and that it can successfully be used with diverse families in a community implementation. The low rate of agreement between self-reports and external coders may reflect over-estimation by practitioners, potentially related to practitioners' management of multiple demands during a session's implementation. The low rate of agreement may also reflect differences in perspectives between raters' conceptualization of adherence and the use minimal sufficiency, and lead to important implications for research and practice in evaluating adherence in parenting interventions.

Implications for Research and Practice

Across parenting programs, assessment of adherence to program content is far more common than evaluation of adherence to a program's underlying processes. Further, evaluation of process adherence is rare among self-report measures. My study is the first to use both self-report and external coder ratings to assess content and process adherence to the Triple P program. The results point to meaningful research implications, including the importance of assessing both content and process adherence in parenting supports and differences across raters. In the evaluation of content adherence, the PSRT expands existing adherence checklists to invite practitioners to track changes they make to flexibly implement the program. To promote self-

regulation, the PSRT extends self-report measures of adherence to incorporate an estimation of process adherence. In evaluating this tool, differences between self-reports and external coders indicate that different informants may rate adherence using the construct of minimal sufficiency, which has important implications for adherence evaluation.

In service delivery for parenting supports, the PSRT aligns with the self-reflective practice that is integral to self-regulation and to the Triple P program (Sanders, 1999, 2012). To date, however, we have not provided an option for practitioners to track and reflect upon content modifications made to better align with client's needs or self-evaluate their promotion of self-regulation with parents throughout their clinical practice. This is a gap the PSRT aims to diminish in both research and practice. It may also be helpful to integrate the PRST into training (Mazzucchelli & Ralph, 2019), as discussion about tracking and monitoring one's implementation – and opportunities for tailoring the program to better suit parents' needs – are in line with the program's self-reflective practice. A further practice implication is that use of the PSRT may be particularly well-suited in supervision activities: best practices in clinical supervision indicate that direct observation permits increased accuracy and helpfulness of supervisory feedback (American Psychological Association, 2014). When direct observation is not feasible, a tool such as the PSRT, which enables tracking of each session component and encourages examples for discussion, could support a fruitful and collaborative discussion between practitioners in peer-supervision, or between a practitioner and a direct supervisor. A key component of the Triple P program is for practitioners to self-reflect on their implementation of the program. Integrating the PSRT into discussions among practitioners about what modifications worked well with their clients and challenges they encountered in implementing session content and in the promotion of self-regulation with clients could prove helpful in

identifying trends in their implementation, as well as highlight areas of strength and those in need of support.

Future Research

The results of this study highlight the need for future examination of both content and process adherence within the large-scale implementation of evidence-based programs such as Triple P. Future research examining adherence in Triple P may benefit from streamlining the PSRT as described above to further simplify its completion in community settings. In addition, it is important to evaluate the tool's predictive validity by assessing the potential impact of content and process adherence rates on clinical outcomes. It will be helpful to evaluate adherence rates separately for self-reports and external coders to understand the threshold of adherence that is required to produce positive effects, and to explore the potential role of minimal sufficiency in adherence measurement. Further, it will be essential in future analyses to examine the correlation across ratings (as the same practitioners submitted multiple adherence forms) using multilevel modeling to account for the nonindependence of data, and permit examination of variability across practitioners' self-reported rates of content and process adherence. In addition, parents participating in the adherence study were recruited from culturally diverse communities, however, information about the families' cultural identity and immigration/generational status was not collected. This information could be helpful moving forward in tracking modifications to program content and process by practitioners. Finally, the findings of the current study require replication: both adherence measures were evaluated in this study in French and the English versions have yet to be examined. The content checklists used in PSRT were already available in French produced by program developers and did not require translation. Terms used in the existing French Triple P materials were used in development of

part two of the PSRT, or the self-regulation checklist, which was developed in French for this study. Similarly, terms from the French version of Triple P were used in the translation of the observational coding system (AMPQ; Adherence Measure for Process Quality; Kirby & Sanders, 2014a, 2014b) from English to French by the research team. Both the PSRT and AMPQ were reviewed by a bilingual expert in the Triple P program. To strengthen confidence in the use of these measures, the French versions require replication, and the English versions require initial examination followed by replication.

Conclusion

Results across studies were consistent with past research indicating that multiple informants tend to perceive the same circumstances differently, and this held true across raters of adjustment and adherence. Differences emerged when reports of adjustment and adherence were compared and reiterated the importance of incorporating multi-informant methodology. This also reinforced the necessity of not only clearly identifying raters, but also to present raters' results separately to clarify variability across reports. Across this dissertation research, the overarching theme is that evidence-based practice can make a positive impact on diverse families navigating challenging times. As the relations between parenting and both child and parental adjustment are well-established (e.g., Barlow & Coren, 2018; Barlow et al., 2014; Collins, Maccoby, Steinberg, Hetherington, & Bornstein, 2000; Grusec, 2011), offering supports that can be delivered flexibly with diverse families are important, as are tracking the program's implementation to ensure its adherence the program's evidence base. Self-regulation is a key component in working with diverse families. Expanding self-reported adherence measurement in Triple P to include both content and process adherence is an essential step to support practitioners' promotion of self-regulation in working with diverse families with a range of

needs. In sum, as communities continue to become increasingly diverse, it is important to continue integrating research findings to identify emerging trends that may impact adjustment and parenting in diverse families. Further, it is essential to continue to refine our tools for evaluating program adherence to inform these services so that we may most effectively support families across a spectrum of needs within a public health approach to support families.

References

- Abdullah, T., & Brown, T. L. (2011). Mental illness stigma and ethnocultural beliefs, values, and norms: An integrative review. *Clinical Psychology Review, 31*, 934-948.
<https://doi.org/10.1016/j.cpr.2011.05.003>
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232. <https://doi.org/10.1037/0033-2909>
- Ahmad, A., Sundelin-Wahlsten, V., Sofi, M. A., Qahar, J. A., & von Knorring, A.-L. (2000). Reliability and validity of a child-specific crosscultural instrument for assessing posttraumatic stress disorder. *European Child & Adolescent Psychiatry, 9*, 285-294.
<https://doi.org/10.1007/s007870070032>
- Allen, J., Shelton, R., Emmons, K., & Linnan, L. (2017). *Fidelity and its relationship to implementation effectiveness, adaptation, and dissemination*. Oxford University Press.
<https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190683214.001.0001/oso-9780190683214-chapter-16>.
- American Psychological Association. (2006). Presidential task force on evidence-based practice. *American Psychologist, 61*, 271-285. <https://doi.org/10.1037/0003-066X.61.4.271>
- American Psychological Association task force on evidence-based practice for children and adolescents. (2008). *Disseminating evidence-based practice for children and adolescents: A systems approach to enhancing care*. American Psychological Association.
- American Psychological Association. (2014). *Guidelines for clinical supervision in health service psychology*. Retrieved from <http://apa.org/about/policy/guidelines-supervision.pdf>
- Anderson, F. M., Hatch, S. L., Comacchio, C., & Howard, L. M. (2017). Prevalence and risk of

mental disorders in the perinatal period among migrant women: a systematic review and meta-analysis. *Archives of Women's Mental Health*, 20, 449-462.

<https://doi.org/10.1007/s00737-017-0723-z>

Aroian, K., Hough, E. S., Templin, T. N., Kulwicki, A., Ramaswamy, V., & Katz, A. (2009). A model of mother-child adjustment in Arab Muslim immigrants to the US. *Social Science & Medicine*, 69, 1377-1386. <https://doi.org/10.1016/j.socscimed.2009.08.027>

Azar, S. T., & Wolfe, D. A. (2006). Child physical abuse and neglect. In E. J. Mash & R. A. Barkley (Eds.), *Treatment of childhood disorders* (3rd ed., pp. 595-646). Guilford.

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

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.

Bandura, A. (1995). Social learning. In A. S. R. Manstead & M. Hewstone (Eds.), *Blackwell encyclopedia of social psychology* (pp. 600-606). Blackwell.

Barlow, J., & Coren, E. (2018). The effectiveness of parenting programs. *Research on Social Work Practice*, 28, 99-102. <https://doi.org/10.1177/1049731517725184>

Barlow, J., Smailagic, N., Huband, N., Roloff, V., & Bennett, C. (2014). Group-based parent training programmes for improving parental psychosocial health. *Cochrane Database of Systematic Reviews*, CD002020. <https://doi.org/10.1002/14651858.CD002020.pub4>

Beiser, M. (2010). The mental health of immigrant and refugee children in Canada: A description and selected findings from the New Canadian Children and Youth Study (NCCYS). *Canadian Issues, Summer*, 103–107.

Beiser, M., Goodwill, A. M., Albanese, P., McShane, K., & Nowakowski, M. (2014). Predictors of immigrant children's mental health in Canada: selection, settlement contingencies,

culture, or all of the above? *Social Psychiatry and Psychiatric Epidemiology*, 49, 743-756.

<https://doi.org/10.1007/s00127-013-0794-8>

Beiser, M., Hamilton, H., Rummens, J. A., Oxman-Martinez, J., Ogilvie, L., Humphrey, C., & Armstrong, R. (2010). Predictors of emotional problems and physical aggression among children of Hong Kong Chinese, Mainland Chinese and Filipino immigrants to Canada. *Social Psychiatry and Psychiatric Epidemiology*, 45, 1011-1021.

<https://doi.org/10.1007/s00127-009-0140-3>

Beiser, M., & Hou, F. (2016). Mental health effects of premigration trauma and postmigration discrimination on refugee youth in Canada. *The Journal of Nervous and Mental Disease*, 204, 464-470. <https://doi.org/10.1097/NMD.0000000000000516>

Beiser, M., Ogilvie, L., Rummens, J. A., Armstrong, R., & Oxman-Martinez, J. (2005). The New Canadian Children and Youth Study: Research to fill a gap in Canada's children's agenda. *Canadian Issues*, Spring, 21-24. <https://doi.org/10.7939/R3D795C1W>

Beiser, M., Puente-Duran, S., & Hou, F. (2015). Cultural distance and emotional problems among immigrant and refugee youth in Canada: Findings from the New Canadian Child and Youth Study (NCCYS). *International Journal of Intercultural Relations*, 49, 33-45. <https://doi.org/10.1016/j.ijintrel.2015.06.005>

Beiser, M., Taa, B., Fenta-Wube, H., Baheretibeb, Y., Pain, C., & Araya, M. (2012). A comparison of levels and predictors of emotional problems among preadolescent Ethiopians in Addis Ababa, Ethiopia, and Toronto, Canada. *Transcultural Psychiatry*, 49, 651-677. <https://doi.org/10.1177/1363461512457155>

Beiser, M., Zilber, N., Simich, L., Youngmann, R., Zohar, A. H., Taa, B., & Hou, F. (2011). Regional effects on the mental health of immigrant children: Results from the New

Canadian Children and Youth Study (NCCYS). *Health Place*, 17, 822-829.

<https://doi.org/10.1016/j.healthplace.2011.03.005>

Betancourt, J. R., Green, A. R., Carrillo, J. E., & Ananeh-Firempong, O. (2003). Defining cultural competence: A practical framework for addressing racial/ethnic disparities in health and health care. *Public Health Reports*, 118, 293–302.

<https://doi.org/10.1093/phr/118.4.293>

Biglan, A., Flay, B. R., Embry, D. D., & Sandler, I. N. (2012). The critical role of nurturing environments for promoting human well-being. *American Psychologist*, 67, 257-271.

<https://doi.org/10.1037/a0026796>

Blackmore, R., Gray, K. M., Boyle, J. A., Fazel, M., Ranasinha, S., Fitzgerald, G., Misso, M., & Gibson-Helm, M. (2020). Systematic review and meta-analysis: The prevalence of mental illness in child and adolescent refugees and asylum seekers. *Journal of the American Academy of Child and Adolescent Psychiatry*, 59, 705-714.

<https://doi.org/10.1016/j.jaac.2019.11.011>

Bodenmann, G., Cina, A., Ledermann, T., & Sanders, M. R. (2008). The efficacy of the Triple P-Positive Parenting Program in improving parenting and child behavior: A comparison with two other treatment conditions. *Behaviour Research and Therapy*, 46, 411-427.

<https://doi.org/10.1016/j.brat.2008.01.001>

Breitenstein, S. M., Fogg, L., Garvey, C., Hill, C., Resnick, B., & Gross, D. (2010). Measuring implementation fidelity in a community-based parenting intervention. *Nursing Research*, 59, 158-165. <https://doi.org/10.1097/NNR.0b013e3181dbb2e2>

Bronstein, I., & Montgomery, P. (2011). Psychological distress in refugee children: A systematic review. *Clinical Child and Family Psychology Review*, 14, 44-56.

<https://doi.org/10.1007/s10567-010-0081-0>

Canadian Psychological Association Task Force on Evidence-Based Practice of Psychological Treatments. (2012). *Evidence-based practice of psychological treatments: A Canadian perspective*.

http://www.cpa.ca/docs/File/Practice/Report_of_the_EBP_Task_Force_FINAL_Board_Aproved_2012.pdf

Carr, A. (2014). The evidence base for family therapy and systemic interventions for child-focused problems. *Journal of Family Therapy*, 36, 107-157. <https://doi.org/10.1111/1467-6427.12032>

Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K. S., Crits-Christoph, P., Baker, M., Johnson, B., Woody, S. R., Sue, S., Beutler, L., Williams, D. A., & McCurry, S. (1996). An update on empirically validated therapies. *The Clinical Psychologist*, 49, 5-18.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Erlbaum.

Collins, W. A., Maccoby, E. E., Steinberg, L., Hetherington, E. M., & Bornstein, M. H. (2000). Contemporary research on parenting: The case for nature and nurture. *American Psychologist*, 55, 218–232. <https://doi.org/10.1037/0003-066X.55.2.218>

Cunningham, S. A., Ruben, J. D., & Narayan, K. M. (2008). Health of foreign-born people in the United States: A review. *Health Place*, 14, 623-635. <https://doi.org/10.1016/j.healthplace.2007.12.002>

Curtis, P., Thompson, J., & Fairbrother, H. (2018). Migrant children within Europe: A systematic review of children's perspectives on their health experiences. *Public Health*, 158, 71-85.

<https://doi.org/10.1016/j.puhe.2018.01.038>

de Graaf, I., Speetjens, P., Smit, F., de Wolff, M., & Tavecchio, L. (2008a). Effectiveness of the Triple P Positive Parenting Program on behavioral problems in children: A meta-analysis.

Behavior Modification, 32, 714-735. <https://doi.org/10.1177/0145445508317134>

de Graaf, I., Speetjens, P., Smit, F., de Wolff, M., & Tavecchio, L. (2008b). Effectiveness of the Triple P Positive Parenting Program on parenting: A meta-analysis. *Family Relations*, 57,

553-566. <https://doi.org/10.1111/j.1741-3729.2008.00522.x>

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, 131, 483-509.

<https://doi.org/10.1037/0033-2909.131.4.483>

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*,

9, 123-149. <https://doi.org/10.1146/annurev-clinpsy-050212-185617>

De Maio, F. G., & Kemp, E. (2010). The deterioration of health status among immigrants to

Canada. *Global Public Health*, 5, 462-478. <https://doi.org/10.1080/17441690902942480>

Dimitrova, R., Chasiotis, A., & van de Vijver, F. (2016). Adjustment outcomes of immigrant children and youth in Europe: A meta-analysis. *European Psychologist*, 21, 150-162.

<https://doi.org/10.1027/1016-9040/a000246>

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

<https://doi.org/10.1136/jech.52.6.377>

- Dozois, D. J. A., Mikail, S. F., Alden, L. E., Bieling, P. J., Bourgon, G., Clark, D. A., Drapeau, M., Gallson, D., Greenberg, L., Hunsley, J., & Johnston, C. (2014). The CPA Presidential Task Force on Evidence-Based Practice of Psychological Treatments. *Canadian Psychology/Psychologie canadienne*, 55, 153-160. <https://doi.org/10.1037/a0035767>
- Dretzke, J., Davenport, C., Frew, E., Barlow, J., Stewart-Brown, S., Bayliss, S., Taylor, R. S., Sandercock, J., & Hyde, C. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: A systematic review of randomised controlled trials. *Child and Adolescent Psychiatry and Mental Health*, 3, 7. <https://doi.org/10.1186/1753-2000-3-7>
- Dwairy, M. (2006). *Counseling and psychotherapy with Arabs And Muslims: A culturally sensitive approach*. Columbia University.
- Eames, C., Daley, D., Hutchings, J., Hughes, J. C., Jones, K., Martin, P., & Bywater, T. (2008). The Leader Observation Tool: A process skills treatment fidelity measure for the Incredible Years parenting programme. *Child: Care, Health and Development*, 34, 391-400. <https://doi.org/10.1111/j.1365-2214.2008.00828.x>
- Erford, B. T., Paul, L. E., Oncken, C., Kress, V. E., & Erford, M. R. (2014). Counseling outcomes for youth with oppositional behavior: A meta-analysis. *Journal of Counseling & Development*, 92, 13-24. <https://doi.org/10.1002/j.1556-6676.2014.00125.x>
- Errázuriz, P., Cerfogli, C., Moreno, G., & Soto, G. (2016). Perception of Chilean parents on the Triple P Program for improving parenting practices. *Journal of Child and Family Studies*, 25, 3440-3449. <https://doi.org/10.1007/s10826-016-0492-8>
- European Migration Network. (2018). *Asylum and migration glossary: A tool for better comparability produced by the European Migration Network (May 2018)*.

https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/networks/european_migration_network/docs/interactive_glossary_6.0_final_version.pdf

Eyberg, S. M. (2005). Tailoring and adapting parent-child interaction therapy to new populations. *Education and Treatment of Children*, 28, 197-201.

<https://doi.org/10.2307/42899841>

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 & Adolescent Psychology*, 37, 215-237. <https://doi.org/10.1080/15374410701820117>

Falah-Hassani, K., Shiri, R., Vigod, S., & Dennis, C. L. (2015). Prevalence of postpartum depression among immigrant women: A systematic review and meta-analysis. *Journal of Psychiatric Research*, 70, 67-82. <https://doi.org/10.1016/j.jpsychires.2015.08.010>

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191. <https://doi.org/10.3758/BF03193146>

Fazel, M., Reed, R. V., Panter-Brick, C., & Stein, A. (2012). Mental health of displaced and refugee children resettled in high-income countries: risk and protective factors. *The Lancet*, 379, 266-282. [https://doi.org/10.1016/s0140-6736\(11\)60051-2](https://doi.org/10.1016/s0140-6736(11)60051-2)

Fazel, M., Wheeler, J., & Danesh, J. (2005). Prevalence of serious mental disorder in 7000 refugees resettled in western countries: A systematic review. *The Lancet*, 365, 1309-1314. [https://doi.org/10.1016/s0140-6736\(05\)61027-6](https://doi.org/10.1016/s0140-6736(05)61027-6)

Feldmann, M., Heinrichs, N., Hahlweg, K., & Bertram, H. (2007). Bibliothérapie zur verbesserung elterlicher erziehungskompetenz: Eine randomisierte evaluationsstudie [Bibliotherapy for improvement of parental competencies: A randomized evaluation

- study]. *Verhaltenstherapie*, 17, 26-35. <https://doi.org/10.1159/000099629>
- Fellmeth, G., Fazel, M., & Plugge, E. (2017). Migration and perinatal mental health in women from low- and middle-income countries: A systematic review and meta-analysis. *British Journal of Obstetrics and Gynaecology*, 124, 742-752. <https://doi.org/10.1111/1471-0528.14184>
- Fernandez, M. A., & Eyberg, S. M. (2009). Predicting treatment and follow-up attrition in parent-child interaction therapy. *Journal of Abnormal Child Psychology*, 37, 431-441. <https://doi.org/10.1007/s10802-008-9281-1>
- Fixsen, D. L., Blase, K. A., Duda, M. A., Naoom, S. F., & Van Dyke, M. (2010). Implementation of evidence-based treatments for children and adolescents: Research findings and their implications for the future. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (2nd ed., pp. 435–450). Guilford Press.
- Forgatch, M. S., & DeGarmo, D. S. (2011). Sustaining fidelity following the nationwide PMTO implementation in Norway. *Prevention Science*, 12, 235-246. <https://doi.org/10.1007/s11121-011-0225-6>
- Forgatch, M. S., & Patterson, G. R. (2010). Parent management training—Oregon model: An intervention for antisocial behavior in children and adolescents. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-Based Psychotherapies for Children and Adolescents* (2nd ed., pp. 159–178). Guilford.
- Forgatch, M. S., Patterson, G. R., & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of competent adherence to the Oregon model of parent management training. *Behavior Therapy*, 36, 3–13. [https://doi.org/10.1016/S0005-7894\(05\)80049-8](https://doi.org/10.1016/S0005-7894(05)80049-8)
- Foster, E. M., Prinz, R. J., Sanders, M. R., & Shapiro, C. J. (2008). The costs of a public health

- infrastructure for delivering parenting and family support. *Children and Youth Services Review*, 30, 493-501. <https://doi.org/10.1016/j.childyouth.2007.11.002>
- 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*, 2, 1-344. <https://doi.org/10.1002/14651858.CD008225.pub2>
- Garbacz, L. L., Brown, D. M., Spee, G. A., Polo, A. J., & Budd, K. S. (2014). Establishing treatment fidelity in evidence-based parent training programs for externalizing disorders in children and adolescents. *Clinical Child and Family Psychology Review*, 17, 230-247. <https://doi.org/10.1007/s10567-014-0166-2>
- Gardner, F., Montgomery, P., & Knerr, W. (2016). Transporting evidence-based parenting programs for child problem behavior (Age 3-10) between countries: Systematic review and meta-analysis. *Journal of Clinical Child & Adolescent Psychology*, 45, 749-762. <https://doi.org/10.1080/15374416.2015.1015134>
- George, M. A., & Bassani, C. (2013). Does living in a neighbourhood with others of the same ethnic background contribute to health of Canada's immigrant children? *Canadian Journal of Public Health*, 104, e482-e486. <https://doi.org/10.17269/cjph.104.4101>
- George, M. A., & Bassani, C. (2016). The health of immigrant children who live in areas with high immigrant concentration. *Ethnicity & Health*, 21, 426-438. <https://doi.org/10.1080/13557858.2015.1066762>
- George, M. A., & Bassani, C. (2018). Influence of Perceived Racial Discrimination on the Health of Immigrant Children in Canada. *Journal of International Migration and Integration*, 19, 527-540. <https://doi.org/10.1007/s12134-018-0539-3>

- George, M. A., Bassani, C., & Armstrong, R. W. (2012). Influence of perceived racial discrimination on health and behaviour of immigrant children in British Columbia. *International Journal of Population Research*, 2012, 1-9.
<https://doi.org/10.1155/2012/274963>
- Guo, M., Morawska, A., & Sanders, M. R. (2016). A randomized controlled trial of Group Triple P with Chinese parents in Mainland China. *Behavior Modification*, 40, 825-851.
<https://doi.org/10.1177/0145445516644221>
- Hamilton, H. A., Marshall, L., Rummens, J. A., Fenta, H., & Simich, L. (2011). Immigrant parents' perceptions of school environment and children's mental health and behavior. *Journal of School Health*, 81, 313-319. <https://doi.org/10.1111/j.1746-1561.2011.00596.x>
- Haslam, D. M., & Mejia, A. (2018). Accommodating race and ethnicity in parenting interventions. In M. R. Sanders & T. G. Mazzucchelli (Eds.), *The power of positive parenting: Transforming the lives of children, parents, and communities using the Triple P system* (pp. 332–343). Oxford University Press. [https://doi.org/10.1093/med-
psych/9780190629069.003.0030](https://doi.org/10.1093/med-psych/9780190629069.003.0030)
- Herschell, A. D., Quetsch, L. B., & Kolko, D. J. (2019). Measuring adherence to key teaching techniques in an evidence-based treatment: A comparison of caregiver, therapist, and behavior observation ratings. *Journal of Emotional and Behavioral Disorders*, 28, 92-103.
<https://doi.org/10.1177/1063426618821901>
- Hogue, A., & Dauber, S. (2013). Assessing fidelity to evidence-based practices in usual care: The example of family therapy for adolescent behavior problems. *Evaluation and Program Planning*, 37, 21-30. <https://doi.org/10.1016/j.evalprogplan.2012.12.001>
- Hou, Y., Benner, A. D., Kim, S. Y., Chen, S., Spitz, S., Shi, Y., & Beretvas, T. (2020).

Discordance in parents' and adolescents' reports of parenting: A meta-analysis and qualitative review. *American Psychologist*, 75, 329-348.

<https://doi.org/10.1037/amp0000463>

Houlding, C., Schmidt, F., Stern, S. B., Jamieson, J., & Borg, D. (2012). The perceived impact and acceptability of Group Triple P Positive Parenting Program for Aboriginal parents in Canada. *Children and Youth Services Review*, 34, 2287-2294.

<https://doi.org/10.1016/j.childyouth.2012.08.001>

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.

<https://doi.org/10.1080/15374410701820174>

Huey, S. J., Jr., Tilley, J. L., Jones, E. O., & Smith, C. A. (2014). The contribution of cultural competence to evidence-based care for ethnically diverse populations. *Annual Review of Clinical Psychology*, 10, 305-338. <https://doi.org/10.1146/annurev-clinpsy-032813-153729>

Hunsley, J., & Mash, E. J. (2018). 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* (2nd ed., pp. 3-16). Oxford University Press.

<https://doi.org/10.1093/med-psych/9780190492243.003.0001>

Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology*, 44, 23-52. <https://doi.org/10.1146/annurev.ps.44.020193.000323>

Kazdin, A. E. (2005). *Parent management training: Treatment for oppositional, aggressive, and antisocial behavior in children and adolescents*. Oxford University Press.

Kazdin, A. E., & Blase, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science*, 6, 21-37.

<https://doi.org/10.1177/1745691610393527>

Kendall, P. C., Gosch, E., Furr, J. M., & Sood, E. (2008). Flexibility within fidelity. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47, 987-993.

<https://doi.org/10.1097/CHI.0b013e31817eed2f>

Kien, C., Sommer, I., Faustmann, A., Gibson, L., Schneider, M., Krczal, E., Jank, R., Klerings, I., Szelag, M., Kerschner, B., Brattstrom, P., & Gartlehner, G. (2019). Prevalence of mental disorders in young refugees and asylum seekers in European Countries: A systematic review. *European Child & Adolescent Psychiatry*, 28, 1295-1310.

<https://doi.org/10.1007/s00787-018-1215-z>

Kirby, J. N., & Sanders, M. R. (2014a). *Adherence measure for the process quality in standard Triple P*. Parenting and Family Support Centre.

Kirby, J. N., & Sanders, M. R. (2014b). *Instruction manual for the adherence measure for the process quality in standard Triple P*. Parenting and Family Support Centre.

Knutson, N. M., Forgatch, M. S., Rains, L. A., & Sigmarisdóttir, M. (2009). *Fidelity of Implementation Rating System (FIMP): The manual for PMTO*. Implementation Sciences International.

Kovacev, L., & Shute, R. (2004). Acculturation and social support in relation to psychosocial adjustment of adolescent refugees resettled in Australia. *International Journal of Behavioral Development*, 28, 259-267. <https://doi.org/10.1080/01650250344000497>

Kumpfer, K. L., Alvarado, R., Smith, P., & Bellamy, N. (2002). Cultural sensitivity and adaptation in family-based prevention interventions. *Prevention Science*, 3, 241-246.

<https://doi.org/1389-4986/02/0900-0241/1>

Lau, A., Barnett, M., Stadnick, N., Saifan, D., Regan, J., Wiltsey Stirman, S., Roesch, S., &

- Brookman-Frazee, L. (2017). Therapist report of adaptations to delivery of evidence-based practices within a system-driven reform of publicly funded children's mental health services. *Journal of Consulting and Clinical Psychology, 85*, 664-675.
<https://doi.org/10.1037/ccp0000215>
- Lee, C. M., Horvath, C., & Hunsley, J. (2013). Does it work in the real world? The effectiveness of treatments for psychological problems in children and adolescents. *Professional Psychology: Research and Practice, 44*, 81-88. <https://doi.org/10.1037/a0031133>
- Lee, C. M., Smith, P. B., Stern, S. B., Piché, G., Feldgaier, S., Ateah, C., Clément, M.-È., Gagné, M.-H., Lamonde, A., Barnes, S., Dennis, D., & Chan, K. (2014). The International Parenting Survey–Canada: Exploring access to parenting services. *Canadian Psychology/Psychologie canadienne, 55*, 110-116. <https://doi.org/10.1037/a0036297>
- Leung, C., Sanders, M. R., Ip, F., & Lau, J. (2006). Implementation of Triple P-Positive Parenting Program in Hong Kong: Predictors of program completion and clinical outcomes. *Journal of Children's Services, 1*, 4-17.
<https://doi.org/10.1108/17466660200600010>
- Lyon, A. R., & Budd, K. S. (2010). A community mental health implementation of Parent-Child Interaction Therapy (PCIT). *Journal of Child and Family Studies, 19*, 654-668.
<https://doi.org/10.1007/s10826-010-9353-z>
- Marley, C., & Mauki, B. (2019). Resilience and protective factors among refugee children post-migration to high-income countries: A systematic review. *The European Journal of Public Health, 29*, 706-713. <https://doi.org/10.1093/eurpub/cky232>
- Matsumoto, Y., Sofronoff, K., & Sanders, M. R. (2010). Investigation of the effectiveness and social validity of the Triple P Positive Parenting Program in Japanese society. *Journal of*

Family Psychology, 24, 87-91. <https://doi.org/10.1037/a0018181>

Maughan, D. R., Christianseri, E., Jensori, 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.

<https://doi.org/10.1080/02796015.2005.12086287>

Mazzucchelli, T. G., & Ralph, A. (2019). Self-regulation approach to training child and family practitioners. *Clinical Child and Family Psychology Review*, 22, 129-145.

<https://doi.org/10.1007/s10567-019-00284-2>

Mazzucchelli, T. G., & Sanders, M. R. (2010). Facilitating practitioner flexibility within an empirically supported intervention: Lessons from a system of parenting support. *Clinical Psychology: Science and Practice*, 17, 238-252. [https://doi.org/10.1111/j.1468-](https://doi.org/10.1111/j.1468-2850.2010.01215.x)

[2850.2010.01215.x](https://doi.org/10.1111/j.1468-2850.2010.01215.x)

McArthur, B. A., Riosa, P. B., & Preyde, M. (2012). Treatment fidelity in psychosocial intervention for children and adolescents with comorbid problems. *Child and Adolescent Mental Health*, 17, 139-145. <https://doi.org/10.1111/j.1475-3588.2011.00635.x>

Mejia, A., Calam, R., & Sanders, M. R. (2015). A pilot randomized controlled trial of a brief parenting intervention in low-resource settings in Panama. *Prevention Science*, 16, 707-717. <https://doi.org/10.1007/s11121-015-0551-1>.

Mellor, D., Renzaho, A., Swinburn, B., Green, J., & Richardson, B. (2012). Aspects of parenting and family functioning associated with obesity in adolescent refugees and migrants from African backgrounds living in Australia. *Australian and New Zealand Journal of Public Health*, 36, 317-324. <https://doi.org/10.1111/j.1753-6405.2012.00894.x>

Merikangas, K. R., He, J.-P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C.,

- Georgiades, K., & Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication–Adolescent supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, 49, 980-989. <https://doi.org/10.1016/j.jaac.2010.05.017>
- Mihalopoulos, C., Vos, T., Pirkis, J., & Carter, R. (2011). The economic analysis of prevention in mental health programs. *Annual Review of Clinical Psychology*, 7, 169-201. <https://doi.org/10.1146/annurev-clinpsy-032210-104601>
- Miller, D. A. A., Ronis, S. T., Slaunwhite, A. K., Audas, R., Richard, J., Tilleczeck, K., & Zhang, M. (2020). Longitudinal examination of youth readmission to mental health inpatient units. *Child and Adolescent Mental Health*, 25, 238-248. <https://doi.org/10.1111/camh.12371>
- Miranda, J., Bernal, G., Lau, A., Kohn, L., Hwang, W. C., & LaFromboise, T. (2005). State of the science on psychosocial interventions for ethnic minorities. *Annual Review of Clinical Psychology*, 1, 113-142. <https://doi.org/10.1146/annurev.clinpsy.1.102803.143822>
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 2693-2698. <https://doi.org/10.1073/pnas.1010076108>
- Newbold, K. B. (2005). Self-rated health within the Canadian immigrant population: Risk and the healthy immigrant effect. *Social Science & Medicine*, 60, 1359-1370. <https://doi.org/10.1016/j.socscimed.2004.06.048>
- Ng, E., Wilkins, R., Gendron, F., & Berthelot, J.-M. (2005). Healthy today, healthy tomorrow? Dynamics of immigrants' health in Canada: Evidence from the National Population Health

Survey. *Statistics Canada, Catalogue 82-618-MWE2005002*.

Norcross, J. C. (2002). Empirically supported therapy relationships. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 3-16). Oxford University Press.

Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy*, 55, 303-315. <https://doi.org/10.1037/pst0000193>

Norcross, J. C., & Wampold, B. E. (2011). Evidence-based therapy relationships: Research conclusions and clinical practices. *Psychotherapy*, 48, 98-102.
<https://doi.org/10.1037/a0022161>

Norcross, J. C., & Wampold, B. E. (2018). A new therapy for each patient: Evidence-based relationships and responsiveness. *Journal of Clinical Psychology*, 74, 1889-1906.
<https://doi.org/10.1002/jclp.22678>

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.
<https://doi.org/10.1007/s10567-008-0033-0>

Oxman-Martinez, J., & Choi, Y. R. (2014). Newcomer children: Experiences of inclusion and exclusion, and their outcomes. *Social Inclusion*, 2, 23-37.
<https://doi.org/10.17645/si.v2i4.133>

Oxman-Martinez, J., Rummens, A. J., Moreau, J., Choi, Y. R., Beiser, M., Ogilvie, L., & Armstrong, R. (2012). Perceived ethnic discrimination and social exclusion: Newcomer immigrant children in Canada. *American Journal of Orthopsychiatry*, 82, 376-388.
<https://doi.org/10.1111/j.1939-0025.2012.01161.x>

- Patterson, G. R. (1982). *Coercive family process*. Castalia.
- Patterson, G. R. (2005). The next generation of PMTO models. *The Behavior Therapist*, 28, 27-33.
- Pelham, W. E., Jr., & Fabiano, G. A. (2008). Evidence-based psychosocial treatments for attention-deficit/hyperactivity disorder. *Journal of Clinical Child & Adolescent Psychology*, 37, 184-214. <https://doi.org/10.1080/15374410701818681>
- Pereplechikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology*, 75, 829-841. <https://doi.org/10.1037/0022-006X.75.6.829>
- Perruchoud, R., & Redpath-Cross, J. (2011). Glossary on migration, 2nd edition. *International Migration Law*, 25. International Organization for Migration.
- Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology & Psychiatry*, 56, 345-365. <https://doi.org/10.1111/jcpp.12381>
- Porter, M., & Haslam, N. (2005). Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: A meta-analysis. *Journal of the American Medical Association*, 294, 602 - 612. <https://doi.org/10.1001/jama.294.5.602>
- Prinz, R. J., & Sanders, M. R. (2007). Adopting a population-level approach to parenting and family support interventions. *Clinical Psychology Review*, 27, 739-749. <https://doi.org/10.1016/j.cpr.2007.01.005>
- Prinz, R. J., Sanders, M. R., Shapiro, C. J., Whitaker, D. J., & Lutzker, J. R. (2009). Population-

based prevention of child maltreatment: The U.S. Triple P system population trial.

Prevention Science, 10, 1-12. <https://doi.org/10.1007/s11121-009-0123-3>

Raby, K. L., Lawler, J. M., Shlafer, R. J., Hesemeyer, P. S., Collins, W. A., & Sroufe, L. A.

(2015). The interpersonal antecedents of supportive parenting: A prospective, longitudinal study from infancy to adulthood. *Developmental Psychology*, 51, 115-123.

<https://doi.org/10.1037/a0038336>

Reed, R. V., Fazel, M., Jones, L., Panter-Brick, C., & Stein, A. (2012). Mental health of

displaced and refugee children resettled in low-income and middle-income countries: Risk and protective factors. *The Lancet*, 379, 250-265. [https://doi.org/10.1016/s0140-](https://doi.org/10.1016/s0140-6736(11)60050-0)

[6736\(11\)60050-0](https://doi.org/10.1016/s0140-6736(11)60050-0)

Richard, J. F., Thériault, M., Audas, R., Ronis, S., Tilleczek, K., Zhang, M., Bell, B., Slaunwhite,

A., & Poirier, N. (2019). Les obstacles et les facilitateurs dans l'accès aux services pour les enfants et les adolescents autistes francophones du Nouveau-Brunswick [Obstacles and facilitators in accessing services for French-speaking children and adolescents with autism in New Brunswick]. *Reflets*, 25, 90-117. <https://doi.org/10.7202/1067045ar>

Ruane, A., Carr, A., Moffat, V., Finn, T., Murphy, A., O'Brien, O., Groarke, H., & O'Dwyer, R.

(2019). A randomised controlled trial of the Group Stepping Stones Triple P training programme for parents of children with developmental disabilities. *Clinical Child Psychology and Psychiatry*, 24, 728-753. <https://doi.org/10.1177/1359104519827622>

Ryan, S. M., & Ollendick, T. H. (2018). The interaction between child behavioral inhibition and

parenting behaviors: Effects on internalizing and externalizing symptomology. *Clinical Child and Family Psychology Review*, 21, 320-339. <https://doi.org/10.1007/s10567-018-0254-9>

- 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.
<https://doi.org/10.1023/A:1021843613840>
- Sanders, M. R. (2008). Triple P-Positive Parenting Program as a public health approach to strengthening parenting. *Journal of Family Psychology*, 22, 506-517.
<https://doi.org/10.1037/0893-3200.22.3.506>
- Sanders, M. R. (2010). Adopting a public health approach to the delivery of evidence-based parenting interventions. *Canadian Psychology/Psychologie canadienne*, 51, 17-23.
<https://doi.org/10.1037/a0018295>
- Sanders, M. R. (2012). Development, evaluation, and multinational dissemination of the Triple P-Positive Parenting Program. *Annual Review of Clinical Psychology*, 8, 345-379.
<https://doi.org/10.1146/annurev-clinpsy-032511-143104>
- Sanders, M. R., Burke, K., Prinz, R. J., & Morawska, A. (2017). Achieving population-level change through a system-contextual approach to supporting competent parenting. *Clinical Child and Family Psychology Review*, 20, 36-44. <https://doi.org/10.1007/s10567-017-0227-4>
- Sanders, M. R., Markie-Dadds, C., & Turner, K. M. T. (2000). *Practitioner's manual for standard Triple P*. Families International Publishing.
- Sanders, M. R., Markie-Dadds, C., Rinaldis, M., Firman, D., & Baig, N. (2007). Using household survey data to inform policy decisions regarding the delivery of evidence-based parenting interventions. *Child: Care, Health and Development*, 33, 768-783.
<https://doi.org/10.1111/j.1365-2214.2006.00725.x>

- Sanders, M. R., & Mazzucchelli, T. G. (2013). The promotion of self-regulation through parenting interventions. *Clinical Child and Family Psychology Review*, 16, 1-17.
<https://doi.org/10.1007/s10567-013-0129-z>
- Sanders, M. R., & Mazzucchelli, T. G. (2018). Core principles and techniques of positive parenting. In M. R. Sanders & T. G. Mazzucchelli (Eds.), *The power of positive parenting: Transforming the lives of children, parents, and communities using the Triple P system* (pp. 63-80). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190629069.003.0004>
- Sanders, M. R., & Prinz, R. J. (2018). Emergence of a population approach to evidence-based parenting support. In M. R. Sanders & T. G. Mazzucchelli (Eds.), *The power of positive parenting: Transforming the lives of children, parents, and communities using the Triple P system* (pp. 32-62). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190629069.003.0003>
- Sanders, M. R., Turner, K. M. T., & Metzler, C. W. (2019). Applying self-regulation principles in the delivery of parenting interventions. *Clinical Child and Family Psychology Review*, 22, 24-42. <https://doi.org/10.1007/s10567-019-00287-z>
- Schoenwald, S. K. (2011). It's a bird, it's a plane, it's ... fidelity measurement in the real world. *Clinical Psychology: Science and Practice*, 18, 142-147. <https://doi.org/10.1111/j.1468-2850.2011.01245.x>
- Schoenwald, S. K., Garland, A. F., Southam-Gerow, M. A., Chorpita, B. F., & Chapman, J. E. (2011). Adherence measurement in treatments for disruptive behavior disorders: Pursuing clear vision through varied lenses. *Clinical Psychology: Science and Practice*, 18, 331-341.
<https://doi.org/10.1111/j.1468-2850.2011.01264.x>

- Schoenwald, S. K., Sheidow, A. J., & Letourneau, E. J. (2004). Toward effective quality assurance in evidence-based practice: Links between expert consultation, therapist fidelity, and child outcomes. *Journal of Clinical Child and Adolescent Psychology, 33*, 94-104. https://doi.org/10.1207/S15374424JCCP3301_10
- Sheshko, D. M., Lee, C. M., & Gagné, M.-H. (2015). *Practitioner Session Reflection Tool*. Family Psychology Laboratory, The University of Ottawa.
- Sheshko, D. M., Lee, C. M., & Gagné, M.-H. (2020). Multimethod adherence measurement in an evidence-based parenting program. *Practice Innovations, Advance online publication*. <https://doi.org/10.1037/pri0000110>
- Slagt, M., Dubas, J. S., Dekovic, M., & van Aken, M. A. (2016). Differences in sensitivity to parenting depending on child temperament: A meta-analysis. *Psychological Bulletin, 142*, 1068-1110. <https://doi.org/10.1037/bul0000061>
- Slonim-Nevo, V., Mirsky, J., Rubinstein, L., & Nauck, B. (2009). The impact of familial and environmental factors on the adjustment of immigrants: A longitudinal study. *Journal of Family Issues, 30*, 92-123. <https://doi.org/10.1177/0192513x08324575>
- Smith, D. K., & Chamberlain, P. (2010). Multidimensional treatment foster care for adolescents: Processes and outcomes. In J. R. Weisz & A. E. Kazdin (Ed.), *Evidence-based psychotherapies for children and adolescents* (2nd ed., pp. 243-258). Guilford.
- Smith, P. B., Brown, J., Feldgaier, S., & Lee, C. M. (2015). Supporting parenting to promote children's social and emotional well-being. *Canadian Journal of Community Mental Health, 34*, 129-142. <https://doi.org/10.7870/cjcmh-2015-013>
- Smith, T. B., Rodriguez, M. D., & Bernal, G. (2011). Culture. *Journal of Clinical Psychology, 67*, 166-175. <https://doi.org/10.1002/jclp.20757>

- Stafford, M., Newbold, B. K., & Ross, N. A. (2011). Psychological distress among immigrants and visible minorities in Canada: A contextual analysis. *International Journal of Social Psychiatry*, 57, 428-441. <https://doi.org/10.1177/0020764010365407>
- Statistics Canada. (1998). *National Longitudinal Survey of Children and Youth (NLSCY) Cycle 2: 1996 - 1997*. <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=4630>
- Statistics Canada. (2013). *Immigration and ethnocultural diversity in Canada: National Household Survey, 2011*. <http://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-010-x/99-010-x2011001-eng.pdf>
- Statistics Canada. (2016). *Quarterly demographic estimates, April to June 2016*. <http://www.statcan.gc.ca/pub/91-002-x/91-002-x2016002-eng.pdf>
- Steinberg, A. M., Brymer, M. J., Decker, K. B., & Pynoos, R. S. (2004). The University of California at Los Angeles Post-traumatic Stress Disorder Reaction Index. *Current Psychiatry Reports*, 9, 96-100. <https://doi.org/10.1007/s11920-004-0048-2>
- Stern, S. B., Alaggia, R., Watson, K., & Morton, T. R. (2008). Implementing an evidence-based parenting program with adherence in the real world of community practice. *Research on Social Work Practice*, 18, 543-554. <https://doi.org/10.1177/1049731507308999>
- Stevens, G. W., & Vollebergh, W. A. (2008). Mental health in migrant children. *Journal of Child Psychology and Psychiatry*, 49, 276-294. <https://doi.org/10.1111/j.1469-7610.2007.01848.x>
- Stormont, M. (2002). Externalizing behavior problems in young children: Contributing factors and early intervention. *Psychology in the Schools*, 39, 127-138. <https://doi.org/10.1002/pits.10025>
- Sue, S., Zane, N., Nagayama Hall, G. C., & Berger, L. K. (2009). The case for cultural

- competency in psychotherapeutic interventions. *Annual Review of Psychology*, 60, 525-548. <https://doi.org/10.1146/annurev.psych.60.110707.163651>
- Sumargi, A., Sofronoff, K., & Morawska, A. (2015). A randomized-controlled trial of the Triple P-Positive Parenting Program seminar series with Indonesian parents. *Child Psychiatry & Human Development*, 46, 749-761. <https://doi.org/10.1007/s10578-014-0517-8>
- Tam, S. Y., Houlihan, S., & Melendez-Torres, G. J. (2017). A systematic review of longitudinal risk and protective factors and correlates for posttraumatic stress and its natural history in forcibly displaced children. *Trauma Violence Abuse*, 18, 377-395. <https://doi.org/10.1177/1524838015622437>
- Taylor, W. D., Asgary-Eden, V., Lee, C. M., & LaRoche, K. J. (2015). Service providers' adherence to an evidence-based parenting program: What are they missing and why? *Journal of Child and Family Studies*, 24, 50-56. <http://dx.doi.org/10.1007/s10826-013-9812-4>
- 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. <https://doi.org/10.1023/A:1021848315541>
- Tehrani-Doost, M., Shahrivar, Z., Mahmoudi Gharaie, J. M., & Alaghband-Rad, J. (2009). Efficacy of positive parenting on improving children's behavior problems and parenting styles. *Iranian Journal of Psychiatry and Clinical Psychology*, 14, 371-379.
- Turner, K. M. T., Hodge, L. M., Forster, M., & McIlduff, C. D. (2018). Working effectively with indigenous families. In M. R. Sanders & T. G. Mazzucchelli (Eds.), *The power of positive parenting: Transforming the lives of children, parents, and communities using the Triple P system* (pp. 321-331). Oxford University Press. <https://doi.org/10.1093/med->

[psych/9780190629069.003.0029](https://doi.org/10.1111/j.1440-1754.2007.01053.x)

Turner, K. M. T., Richards, M., & Sanders, M. R. (2007). A randomised clinical trial of a group parent education program for Australian indigenous families. *Journal of Paediatrics and Child Health*, 43, 429-437. <https://doi.org/10.1111/j.1440-1754.2007.01053.x>

Turner, K. M. T., Sanders, M. R., Keown, L. J., & Shepherd, M. (2018). A collaborative partnership adaptation model. In M. R. Sanders & T. G. Mazzucchelli (Eds.), *The power of positive parenting: Transforming the lives of children, parents, and communities using the Triple P system* (pp. 310-320). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190629069.003.0028>

United Nations High Commissioner for Refugees. (2006). Master glossary of terms, Rev. 1. *Division of International Protection Services Status Determination and Protection Information Section*. <http://www.unhcr.org/refworld>

United Nations High Commissioner for Refugees. (2016). *Global trends: Forced displacement in 2015*. <http://www.unhcr.org/576408cd7.pdf>

United Nations High Commissioner for Refugees. (2020). *Global trends: Forced displacement in 2019*. <https://www.unhcr.org/globaltrends2019/>

United Nations Office on Drugs and Crime. (2010). *Compilation of evidence-based family skills training programs*. <https://www.unodc.org/documents/prevention/family-compilation.pdf>

van Os, E. C., Kalverboer, M. E., Zijlstra, A. E., Post, W. J., & Knorth, E. J. (2016). Knowledge of the unknown child: A systematic review of the elements of the best interests of the child assessment for recently arrived refugee children. *Clinical Child and Family Psychology Review*, 19, 185-203. <https://doi.org/10.1007/s10567-016-0209-y>

Vang, Z. M., Sigouin, J., Flenon, A., & Gagnon, A. (2015). The Healthy immigrant effect in

Canada: A systematic review. *Population change and lifecourse strategic knowledge cluster discussion paper series/Un réseau stratégique de connaissances changements de population et parcours de vie document de travail*, 3. <http://ir.lib.uwo.ca/pclc/vol3/iss1/4>

Vang, Z. M., Sigouin, J., Flenon, A., & Gagnon, A. (2017). Are immigrants healthier than native-born Canadians? A systematic review of the healthy immigrant effect in Canada. *Ethnicity & Health*, 22, 209-241. <https://doi.org/10.1080/13557858.2016.1246518>

Webster-Stratton, C. (1998). Preventing conduct problems in Head Start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology*, 66, 715-730. <https://doi.org/10.1037/0022-006X.66.5.715>

Wessels, I., & Ward, C. L. (2016). Battered women and parenting: Acceptability of an evidence-based parenting programme to women in shelters. *Journal of Child and Adolescent Mental Health*, 28, 21-31. <https://doi.org/10.2989/17280583.2015.1132425>

World Health Organization. (2009). *Preventing violence through the development of safe, stable and nurturing relationships between children and their parents and caregivers*. https://apps.who.int/iris/bitstream/handle/10665/44088/9789241597821_eng.pdf;jsessionid=9B28C2C7B306A296888005761174BF3F?sequence=1

Appendix A: Sample Triple P Session Checklist

GROUP TRIPLE P SESSION 1 CHECKLIST

Use this as a guide and as record of what you covered in the session. Indicate with a checkmark (✓) in the “YES” box if the item was covered, or the “NO” box if the item was omitted.

Date: _____ Name of facilitator: _____ Name of co-facilitator: _____

Start time: _____ Finish time: _____ # of caregivers attended: _____

Content Checklist

	YES	NO
1. Introduction		
• Congratulate parents on making the commitment to participate	☐	☐
• Provide an overview of Group Triple P	☐	☐
• If any parent did not complete <i>Assessment Booklet One</i> or is just joining the group have him/her complete the assessment packet before starting this session	☐	☐
2. Agenda (outline proposed session goals and gain consent from the group)	☐	☐
• Behavior Rating Scales		
• Working as a group		
• What is positive parenting		
• Causes of child behavior problems		
• Goals for change		
• Keeping track of children's behavior		
• Homework tasks		
3. Behavior Rating Scales		
• Distribute and have caregivers complete the Behavior Rating Scales (~2 mins)	☐	☐
• Immediately collect Behavior Rating Scales	☐	☐
4. Working as a Group		
• Exercise 1: Setting basic ground rules for the group	☐	☐
• Exercise 2: Getting to know you	☐	☐
• Exercise 3: What you would like to get out of the group session	☐	☐
5. What is Positive Parenting?		
• Discuss the challenges of being a parent	☐	☐
• Show video: Opening segment	☐	☐
• Introduce positive parenting		
○ Ensuring a safe and interesting environment	☐	☐
○ Creating a positive learning environment	☐	☐
○ Using assertive discipline	☐	☐
○ Having realistic expectations	☐	☐
○ Taking care of oneself	☐	☐
• Show video Part 1: What is Positive Parenting?	☐	☐

- | | | |
|---|--------------------------|--------------------------|
| • Exercise 4: What is positive parenting? | ☐ | ☐ |
| 6. Causes of Child Behavior Problems | | |
| • Introduce factors know to influence children's behavior | | |
| ○ Genetic make-up | ☐ | ☐ |
| ○ The family environment | ☐ | ☐ |
| ○ Influences outside the home | ☐ | ☐ |
| 6. Causes of Child Behavior Problems (continued) | | |
| • Show video Part 2: Causes of Child Behavior Problems | ☐ | ☐ |
| • Exercise 5: Identifying causes of child behavior problems | ☐ | ☐ |
| 7. Goals for Change | | |
| • Exercise 6: What skills should we encourage in children? | ☐ | ☐ |
| • Exercise 7: Setting goals for change | ☐ | ☐ |
| 8. Keeping Track of Children's Behavior | | |
| • Introduce monitoring of children's behavior | ☐ | ☐ |
| • Show video Part 2: Goals for Change, Keeping Track | ☐ | ☐ |
| • Exercise 8: Keeping track | ☐ | ☐ |
| 9. Session Close | | |
| • Review the main points covered in the session | ☐ | ☐ |
| • Explain why the homework tasks are important | ☐ | ☐ |
| • Explain homework tasks (monitoring one or two child behavior) | ☐ | ☐ |
| • Ask if any parent would like to take the DVD home to review | ☐ | ☐ |
| ○ List name of parent checking out video_____. | | |
| • Outline the content of the next session and give a reminder of the day and time | ☐ | ☐ |

Session Notes

Additional Agenda Items (note any additional content or major deviation from the set program)

Return any extra *Assessment Booklet One* to the Parent Training Institute

Signed: _____
(facilitator)

Date completed: _____

Signed: _____
(co-facilitator)

Signed: _____
(co-facilitator)

Appendix B: Search Terms

Database Search Terms						
PsycINFO	Medline (PubMed)	Embase	Cochrane	Cinahl	Scopus	Web of Science
1	Immigra*	“Emigrants and Immigrants” OR “Emigration and Immigration”	“Immigrant” and “Immigratio n”	Immigra*	“Emigration and Immigration”	Immigra*
	Refuge*	“Refugees”	Refuge* "displaced people" or "displaced person" or "displaced persons"		Refuge*	Refuge*
	Displace*	“Displaced person OR displaced people”	Displace*	“Refugees” [incl. Displaced Persons]	“Displace d person” OR “Displace d people”	“Displaced person” OR “Displaced people”
	Asylum- seek*	“Asylum- seeker OR Asylum- seeking”	Asylum- seek*	Asylum-seek*	“asylum- seek”	Asylum- seek*
AND						
2	Parent*	“Parents” OR "Parenting"	“parent”	Parent*	“Parents”	Parent*
	Youth		“juvenile”	Youth		Youth
	Adolescen*	"Adolescent"	“adolescenc e” OR “adolescent”	Adolescen*	“Adolescence ” [incl. Youth]	Adolescen*
	Child* Famil*	"Child" "Family"	“child/” “Family”	Child* Famil*	“Child” “Family”	Child* Famil*
AND						
3	Parent child relation*	“Parent-Child Relations”	“child parent relation”	Parent child relation*	“Parent-Child Relations”	“Parent child relation*”
	Parental Psychopath ology	“Parental Psychopathol ogy”	Parental Psychopatho logy	Parental Psychopatholog y	“Parental Psychopathol ogy”	“Parental Psychopath ology”
	Well being OR Mental health	“Well being”	“wellbeing”	Well being	“Psychologica l Well-Being”	“Well being” or “Well- being”

Continued

Appendix B continued

	PsycINFO	Medline (PubMed)	EMbase	Cochrane	Cinahl	Scopus	Web of Science
		“Mental Health”	“mental health”	Mental health	“Mental Health”	“Mental health”	“Mental health”
		“Adjustment Disorders” OR “Social Adjustment”	“adjustment ”	Adjust*		Adjustme nt	Adjustment
3	Adjust* OR Adapt*	“Resilience, Psychologica l” OR “adapt*”	“adaptive behavior”	Adapt*	“Adaptation, Psychological ” [incl Adjust*]	“adapt” OR “adaptive behavior” OR “adaptatio n”	“adapt” OR “adaptive behavior” OR “adaptation ”

Note. Search terms queried with the Boolean operator * yielded results with any variant of the end of the search terms. Search terms queried in “*search term*” conducted a keyword search; other search terms consisted of searches using mapped headings; please note mapped heading categories/thesaurus varies across databases.

Appendix C: Meta-analysis Inclusion Protocol

Meta-analysis: Inclusion Protocol

Construct	Coding rules applied to article's abstract	Data entry
Determine article type	1 = Case Report 2 = Theory/model 3 = Study of Adjustment (Quantitative) 4 = Treatment study (Quantitative) 5 = Narrative review 6 = Systematic review; Meta-analysis 7 = Other	Enter: the number associated with the article type
Determine <i>Newcomer</i> status	Is the sample's average number of years in the host country less than 10 years? Sample is classified as newcomers if the abstract specified length of time in host country or described as "recently arrived"	Enter: Yes (newcomer) No ? (insufficient information to code; review full-text)
Determine if child sample included	Does the abstract specify that children/adolescents/youth participated in the study?	Enter: Yes (child/adolescent) No ? (insufficient information to code; review full-text)
Determine if parent sample included	Does the abstract specify that adults described as parents or as members of a family participated in the study?	Enter: Yes (child/adolescent) No ? (insufficient information to code; review full-text)
Determine status in host country	Does the abstract specify the status of the participants in the host country?	Enter: Immigrant Refugee Asylum seeker Displaced persons Migrant Combination of groups ? (insufficient information to code; review full-text)
Determine measures	Does the abstract specify if the following constructs were measured? • Child adjustment • Parent adjustment • Parenting	Enter: for each construct Yes No ? (insufficient information to code; review full-text)

Appendix D: Meta-analysis Data Extraction Protocol

Meta-analysis: Data Extraction Protocol

Construct	Description
Full text Review: Determining Eligibility	
To locate article: search by first author's last name in file of PDF folders:	
• Search Google Drive for PDF; if not in Drive, check EndNote Library file for is PDF attachment • If no PDF but there is a <u>website link</u>, follow it to the PDF • If it is <u>easy to access the PDF through the website link</u>, save a copy of the PDF to our shared folder on Dropbox ** Please do not spend time searching for articles – I can ask a volunteer in Ottawa to help me with that ☺ **	
Determining Eligibility: Go to “Sample” tab on Excel and complete the first section “Full text: Meets Inclusion criteria?” = Go to article’s METHOD section:	
1) Includes measures of Adjustment (child or parent) or Parenting from list? a. If the measure is on our list, but the authors used a modified version (e.g., fewer items), proceed with coding, but list in this column that it is a modified version (include the reference used by authors) b. Specify which types of measures are included from the Measures List (select from drop-down list): (i) child adjustment only, (ii) parental adjustment only, (iii) parenting behaviour only, (iv) child & parental adjustment, or (v) either adjustment & parenting 2) Are participants children (Age Criterion: mean of sample ≤ 18; age range ≤ 19)? Or parents of children aged 19 years or less? 3) Newcomer Criterion (mean number of years in host country ≤ 10 years) 4) Children migrated with their parents (any separation between children and parents ≤ 2 years) 5) Are Data available for target group? Can we extract data for the newcomer immigrant/refugee group of interest? E.g., If data are collapsed across Newcomer and non-Newcomer groups, or across adolescents and young adults, then we <u>cannot</u> extract data for the target group, and this is coded “No” * See column “Notes from Abstract Review:” I noted the main questions that retained this article for further review after reading the abstract. It may be helpful to check on these questions first when determining the full-text’s eligibility	
If YES to all 4 of these criteria, enter “YES” in “Code-able?” column	
If NO to ANY of these criteria, enter “NO” in “Code-able?” column; list reason in “Reason” column (e.g., “data collapsed across eligible and ineligible groups)	
• Additional Screening Questions: If NO to any of these, exclude article, and list these reasons in “Reason” column ○ Is the article in English? – If NO: Enter “Not available in English”	
Articles with eligible and ineligible groups:	
If includes both eligible and ineligible groups such as, unaccompanied minors, extract data only for the <u>relevant immigrant/refugee group</u>	
Emailing authors:	
If information is missing (as described below) for an article published within the past 10 years , enter “No” in column “Code-able?” and list “Please Email author” in “Reason/Flag for group” column – briefly describe why: e.g., “mean years in host country not reported” or “not reported if child/youth migrated with parents” or “data collapsed across newcomers and non-newcomers” (Dana will email author with email template below).	
For missing values: enter 999	

Entered for data that you searched for, but were not reported/not asked in the study	
Construct	Description
Full-text?	Enter Yes/No if we have the full text available through EndNote/Google Drive
Quantitative Study?	Are quantitative data of adjustment/parenting available? - If YES: Enter “Yes” in “Code-able?” column; Proceed with Screening Questions - If NO: “Code-able?” column: Enter “No”; “Reason” column: list “No quantitative data” and describe type of article: e.g., Case Report; Theoretical/model/discussion paper; Narrative/Literature review; Systematic review; Meta-analysis; Qualitative study; Introduction to Special Edition; Editorial
Construct	Description
Code-able?	Enter yes or no if article is code-able
Reason	If article is not code-able, list reason (see <i>Comment</i> in “Reason” Excel cell for examples)
Group Status/Description	If the article is code-able, use this column to enter a description of the sample, especially if there are multiple groups presented separately: for example - Study includes both an experimental and control group of immigrant youth: each group will have a line of data in the Excel. For Group Status, the first group’s description is “Experimental” and the second is “Control” - Data presented separately two newcomer immigrant group: (1) 1–2 years and (2) 3–5 years - Data presented separately for two refugee newcomer groups: (1) Syrian and (2) Libyan
Sample Characteristics	
Status	Best descriptor: (i) Immigrant; (ii) Refugee; (iii) Immigrant/refugee; (iv) Asylum seeker; (v) Detainees (refugee); (vi) Displaced person; (vii) Migrant Note: If sample includes participants with a different status (e.g., arrived on a temporary Work/Tourist/Student visa, then stayed) and data collapsed across all groups: if this subcategory is 10% or less of the total sample, include full sample. If subcategory is more than 10%, exclude sample (“No” in “Code-able,” “Email author” in “Reason” columns)
Origin Country	Country emigrated <u>from</u>: * If 2+ origin countries are listed, select “Multiple” and then type the countries listed in next column. * If origin country not listed, select “Other” and list countries in next column. * If origin country is a <u>Region</u> (e.g., South America, Middle-Eastern, South Asian), or a former country not listed (e.g., former Yugoslavia), select “Other” and list region in next column.
Origin Country: List	List names of all origin countries/regions
Host Country	Code final destination. Enter “Multiple” if sample includes participants who immigrated to various countries. Note: For Multiple Host Countries: If article presents data from multiple host countries, and data are presented <u>separately for each country group</u>, insert an additional line for each country group, and treat each line as a unique article - E.g., Article comparing immigrants to Germany and Israel; Means, SDs, and N are presented for each country group separately. When entering data, all demographic and child adjustment data for German sample entered on row 1 (listed as Author Names: A); all data for Israeli sample entered on row 2 (Author Names: B). <u>Insert this additional line for all excel sheets</u> (extract data separately for each country group) - If multiple host countries are listed, and data are collapsed across all countries, select “Multiple” (keep article’s entry to one line in database); list countries in next column - If host country not listed in excel options, select “Other”& list countries in next column.

Host Country: List	List names of all host countries (if selected “Multiple” or “Other” in previous column)																																		
Host Country Characteristics																																			
For this section: do the authors raise any of the following topics in describing the sample? Use information in the description of the Sample/Participants. For all variables: • If yes: enter description as stated in article (e.g., 40% families were two-parent households) • If no: enter “not specified”																																			
Two-parents	Is the proportion of two-parent families reported for the sample? (may be described portion of as “Marital status/cohabitation,” indicating proportion of households with two adults)																																		
Siblings	Are the presence or number of siblings for child/adolescent participants reported for the sample? (may also be described as the number of children in the home)																																		
Housing	Are participants’ housing reported? (e.g., detention centres, community housing)																																		
Settlement	Are access to settlement/reception services reported? (e.g., language classes, employment information/supports, child care)																																		
Community	Are the proportion of participants who settled in communities with a similar language or culture to their country of origin reported?																																		
Youth Demographics																																			
Enter youth demographic data as presented in the article: • If the data are presented for the entire youth sample, enter data in one line. • If the data are presented separately (complete demographic and adjustment measure data) for multiple groups (e.g., separated by age group; gender; ethnicity; country of origin; host country; experimental/control group), <u>enter data separately by inserting a new line into the excel spreadsheet, and treat group as a separate entry:</u> ○ Be sure to insert the additional line into each tab/sheet in the excel, and label the lines accordingly (e.g., Aroian et al., A; Aroian et al., B, etc) • If data are presented separately by gender for the demographic data, but are collapsed for the child adjustment measures, <u>enter youth demographic data separately by gender (into the appropriate gender sections in the excel sheet, in the same line), and enter the adjustment data for the full group.</u>																																			
If entering demographic data in this section, enter the number of girls/women and boys/men as it is reported in the article (all age and host country info presented for full group): • E.g. 1: Total youth sample is $N = 200$, girls = 110, boys = 90; all 1st generation • E.g. 2: Total youth sample is $N = 200$, girls = 55%; all 1st generation <table border="1"> <thead> <tr> <th rowspan="2">N: Combined Youth</th><th rowspan="2">Gender: Youth (% female)</th><th rowspan="2">N: Female</th><th rowspan="2">N: Male</th><th rowspan="2">% 2nd gen.</th><th colspan="4">Age: Combined Youth</th><th rowspan="2">etc</th></tr> <tr> <th>Mean</th><th>SD</th><th>Min (Age Range)</th><th>Max (Age Range)</th></tr> </thead> <tbody> <tr> <td>200</td><td></td><td>110</td><td>90</td><td>0</td><td>12.50</td><td>1.10</td><td>10.00</td><td>15.00</td><td></td></tr> <tr> <td>200</td><td>55</td><td></td><td></td><td>0</td><td>12.50</td><td>1.10</td><td>10.00</td><td>15.00</td><td></td></tr> </tbody> </table>		N: Combined Youth	Gender: Youth (% female)	N: Female	N: Male	% 2 nd gen.	Age: Combined Youth				etc	Mean	SD	Min (Age Range)	Max (Age Range)	200		110	90	0	12.50	1.10	10.00	15.00		200	55			0	12.50	1.10	10.00	15.00	
N: Combined Youth	Gender: Youth (% female)						N: Female	N: Male	% 2 nd gen.	Age: Combined Youth				etc																					
		Mean	SD	Min (Age Range)	Max (Age Range)																														
200		110	90	0	12.50	1.10	10.00	15.00																											
200	55			0	12.50	1.10	10.00	15.00																											
If study is longitudinal, enter demographic information at Time 1																																			
Construct	Description																																		
N: Youth	Sample Size: Youth - Total number of youth																																		
% 2 nd	If sample consists of both 1 st and 2 nd generation youth (both of which meet																																		

generation	Newcomer status), enter percentage of sample that is 2 nd generation. Note: If children are 2 nd gen, but were born within 10 years of parents immigrating, count as Newcomer (enter all relevant data). However, if <u>parent duration in country not specified</u> , and children are 2 nd generation, cannot assume youth meet Newcomer = Exclude	
Age: Youth	Mean in years of youth (children/adolescents); convert months into part of year: e.g., Mean of 12 years and 3 months is entered as Mean of 12.25 If age is presented differently (e.g., Median), and age range is within our limits (0 – 19), enter mean age as missing (999), extract data from child adjustment measures	
	SD: Standard deviation of youth age If no SD is reported enter SD as missing (999)	
	Range: lowest of age range	If sample age range not listed, but eligibility criteria describes “children aged x-x,” enter range described as recruitment criterion.
	Range: highest of age range	
Years in Host Country: Youth	Mean of sample - If no mean duration is reported... ... but sample described as in host country for “less than one year,” enter as “up to one year”; article means inclusion criterion, but mean unavailable. ... but study’s recruitment criteria = living in host country for less than 10 years, enter as “up to 10 years”; article means inclusion criterion, but mean unavailable. - If years in host country is not reported, but think they may be Newcomers (e.g., recruited from agencies that tend to service new immigrants): Email author - If data are <u>Current Mean Age</u> & <u>Mean Age at Immigration</u>: Subtract = years in host country. - If sample is split into groups with different durations: - E.g., group 1 in host country for 2 years or less, group 2 in host country for 3 years or more (mean of overall/both groups is less than 10 years): Email author - E.g., group 1 in host country for 5 years or less, group 2 in host country for 6 years or more (mean of group 1 counts as Newcomer, mean of group 2 is over 10 years and is ineligible) = only extract data for eligible group (group 1) - If authors present time in host country in months, convert to years prior to entering data (to facilitate data importation into meta-analysis program)	
	SD: standard deviation of mean years in host country	
	Range: lowest of range	
	Range: highest of range	
	Average Age at Immigration (M): If, instead of reporting number of years in host country, study reports participants’ average age (and SD) when they immigrated to host country, enter here. If not listed: Email author	
	Average Age at Immigration (SD): Standard deviation of the average age of participants when immigrating to new country If not listed: Email author	
Parent Demographics		
When the parent demographic data are presented separately for Mothers and Fathers, enter the demographic data separately.		
If study is longitudinal, enter demographic information at Time 1		
Construct	Description	
N: Parent	Sample Size: Parent - Total number of parents	
Age: Parent	Mean in years of parents; convert months into part of year: e.g., Mean of 40 years	

	and 3 months is entered as Mean of 40.25
	SD: Standard deviation of parent age
	Range: lowest of age range
	Range: highest of age range
Gender: Parent	Proportion of each gender: Enter % of women
Years in Host Country: Parent	Mean of sample (if sample consists of both eligible and ineligible participants, enter only for eligible)
	SD: standard deviation
	Range: lowest of range
	Range: highest of range
	Average Age at Immigration (M): If, instead of reporting number of years in host country, study reports participants' average age (and SD) when they immigrated to host country, enter here.
	Average Age at Immigration (SD): Standard deviation of the average age of participants when immigrating to host country
Methodology	
Construct	Description
Data Set	If applicable/known: are these data set part of a larger study? Yes, No, Possibly
Data Set Name	If applicable/known: identify if study is part of a larger data set (checking for overlap across studies) (independence of data): (i) select from list of large data sets, (ii) identify if study name is provided (but not included on our list), or (iii) select that larger study not named.
Data Set Name: List	If the sample is part of a larger study, list the name of the study (if not listed in "Data Set Name"). If unclear, but suspect there are (potentially) related studies, list the potentially related article(s). E.g., Aroian et al, 2011: states that method previously described in Aroian, 2009
Study Design	Indicate if study is (i) Cross-sectional, (ii) Longitudinal (non-experimental), or (iii) RCT/Experimental
Comparison Group?	Indicate if the study includes a comparison group (compared to the immigrant/refugee group of interest). Select: (i) Other; (ii) Yes (non-immigrant/ refugee comparison); (iii) Yes (experimental-control group comparison); or (iv) No (no comparison group)
Follow-up	If longitudinal, duration of follow-up interval (in months), regardless of number of follow-up assessments. As only Time 1 data are being entered, follow-up interval is entered descriptive purposes. E.g., Two year study with Time 1, follow-up at 12 months, and follow-up at 24 months: enter 24 as the duration of follow-up interval.
Method: Measurement	Select the best description of the type of measurement used: Self-report, Parent-report, Observations, Interview, Other, Multi-method
NOTES (describe)	Use this section to describe (potentially) related studies, particularly if article may be part of a larger study, but it is unclear. E.g., Aroian et al, 2011: method previously described in Aroian, 2009
Data Extraction: ALL measures	
Extract data for EACH MEASURE that is on the list of measures in its own line in the Excel :	
1) <u>Measure</u> : List the name of the measure (confirm it is on the Measures List)	
2) <u>Rater</u> : List who provided the data (select from drop-down list)	

- For each rater, insert a new row in the Excel to extract the data (e.g., Youth Self-report and Parent report)
- 3) **Score:** Specify whether this is the total score of the measure, or a subscale score
 - For each total and/or subscale reported, insert a new row in the Excel to extract the data (e.g., ASEBA-Child Behavior Checklist: Overall/Total score reported, as are the Internalizing and Externalizing subscales = each one gets its own row for data extraction)
- 4) **Age:** Specify the age range that this specific measure was used to assess in this article
- 5) **Psychometrics:** Are any reliability or validity data presented for the use of this measure with this specific immigrant/refugee sample? (e.g., using the ASEBA-CBCL with a sample Syrian refugees in Canada: are data on the measure's internal consistency presented for this sample?)

Additional information for specific circumstances:

- **If there are multiple eligible samples in one paper:** For studies that present data for distinct samples (e.g., Slonim-Nevo et al, 2009: Israeli and German samples), extract data for the measure separately for each eligible group **if the authors presented demographic data separately for each eligible group**
 - As described in "Host Country" country, insert additional lines for each country group, and treat as a unique article (e.g. row 1 = "Study ID: A"; row 2 = "Study ID: B")
- If study is **longitudinal**, enter **Time 1** data only
- If data are presented as a **Composite Score** that combines parent and youth reports, **Email author for separate scores**

General information for **Extracting Data:**

When extracting data, the priority is to extract **Means, SD, and N for the measure of interest**. If that information is not available, move to other forms of data (x^2 , Odds, etc)

- **Sample Size:** when extracting the Means and SDs for a particular measure, include the number of participants who contributed to that Mean and SD.
- If there is a discrepancy between the *N* as reported in the demographic info and the actual number of participants who completed the measure:
 - **INCLUDE:** Enter data if at least **80%** of full sample completed measure
 - **EXCLUDE:** if less than 80% of full sample completed measure
- If *N* used to calculate Mean and SD is not specifically reported, assume the full sample completed the measure (unless examination of degrees of freedom suggests otherwise)
- If mean score on a measure of child adjustment is presented, but **SD is not included** (and no other data are available to extract for meta-analysis): **Email author**
- If the only child adjustment data that are available as presented as **Correlation or Regression** (i.e., data are not presented in a way where we can calculate Means and SDs): **Email author**

Data Extraction: All Measures of Adjustment or Parenting

Enter data as it is presented in the article: if presented separately for boys and girls (or women and men), enter separately in the same line of database. Only use "Combined Gender" section if the data are presented for all youth.

- If the article has divided the data into groups (by age, country of origin, etc), and complete demographic and adjustment/parenting measurement data are available for each group, insert a new line into the database for each group and treat as individual studies.
- If data are presented for 2+ groups that are all eligible immigrant/refugee groups, but they are in distinct experimental and control groups, insert a new line into the database for each group and treat as individual

studies - but only enter the Pretest assessment data for each group.	
Construct	Description
<i>Ideally, extract Means, SDs, and N for each measure</i>	
Mean	Mean
SD	SD
N	Sample size included in this specific analysis
<i>If Ms and SDs are not available, extract the N and any of the following: d, F, t, x2, Odds, Corr.</i>	
P	P value significance
D	Effect size
F	F value
T	t value
χ^2	Chi Square
Odds	Odds ratio
Corr.	Correlation (if data are only presented as a correlation: Email author)
Df	Degrees of Freedom

Appendix E: Meta-analysis Measures List

Measures List: Child Adjustment, Parental Adjustment, Parenting

Acronym	Child Adjustment Measure	Validated for years	Report	Overall	Internalizing	Externalizing	Depression; Suicide	Anxiety	Trauma	Self-esteem	Peer Social Support
ADIS	Anxiety Disorders Interview Schedule	7 - 17	Youth; Parent;					X			
	ADIS - Clinician Rating Scale		Clinician								
ASEBA: CBCL	Child behavior checklist	1.5 – 18	Parent	X	X	X					
ASEBA: TRF	Teacher report form CBCL	1.5 – 18	Teacher	X	X	X					
ASEBA: YSR	Youth Self-report CBCL	11 – 18	Youth	X	X	X					
ASPD	Antisocial Process Screening Device	6 - 13	Parent; Teacher			X					
BASC-2/3	Behavior Assessment System for Children (2 nd ed; 3 rd ed) Parent or Teacher Rating Scales (PRS/TRS); Self-report of Personality (SRP; youth)	2-21 (P, T) 8-21 (Y)	Parent; Teacher; Youth		X	X					
BCS	Behavioral Coding System	Children	Clinician-administered			X					
BDI	Beck Depression Inventory	13+	Youth SR				X				
BHS	Beck Hopelessness Scale	13+	Youth				X				
BPC	Brief Problem Checklist	7 – 13	Youth; Parent	X							
BSI	Brief Symptom Inventory	13+	Youth	X							
BSS	Beck Scale for Suicidal Ideation	adolescents	Youth				X				
CADBI	Child & Adolescent Disruptive Behavior Inventory	3-18	Parent; Teacher			X					
CAPA	Child and Adolescent Psychiatric Assessment	9 – 17	Interview				X				
CAPES	Child Adjustment and Parent Efficacy Scale	2 – 12	Parent	X							
CDI	Children's Depression Inventory	7 – 17	Youth; Parent; Teacher				X				
CDRS-R	Children's Depression Rating Scale-Revised	6 – 12	Clinician-administered				X				
CES-D	Center for Epidemiological Studies Depression Scale	12 – 18	Youth; Parent				X				
CES-DC	Center for Epidemiological Studies Depression Scale for Children	6 – 23	Youth; Parent				X				
C-GAS	Child Global Assessment Scale	4 – 16	Clinician-administered	X							
CPSS	Child PTSD Symptom Scale	8 – 18	Youth; Clinician						X		
CPTSDI	Children's PTSD Inventory	6 – 18	Clinician						X		
CPTS-RI;	Children's PTSD Reaction Index	8 – 18	Youth;						X		

Acronym	Child Adjustment Measure	Validated for years	Report	Overall	Internalizing	Externalizing	Depression; Suicide	Anxiety	Trauma	Self-esteem	Peer Social Support
CPTSD-RI			Clinician								
CRS / CRS-R / CBRS (CPRS/C TRS)	<u>Conners Rating Scales</u> : Conners Rating Scales – Revised; Conners Behavior Rating Scales; Conners-3 (Conners Parent Rating Scales; Conners Teachers Rating Scales)	P/T: 6 – 18 Y: 8 - 18	Youth; Parent; Teacher	X							
CSPS	Child Suicide Potential Scales	8 – 12	Clinician-administered				X				
CY-BOCS	Children's Yale-Brown Obsessive-Compulsive Scale	6 – 17	Parent; Clinician-admin					X			
DISC-IV	Diagnostic Interview Schedule for Children-4 th edition	P: 6 – 17 Y: 9 – 17	Clinician-administered		X	X	X	X	X		
DPICS	Dyadic Parent-Child Interaction Coding System	3 – 12	Observation			X					
DSRS / DSRSC	Depression Self-Rating Scale / DSRSC for children	8 – 14	Youth				X				
ECBI/SES BI-R	Eyberg Child Behavior Inventory / Sutter-Eyberg Child Behavior Inventory-Revised	2 – 16	Parent			X					
FSSC-R	Fear Survey Schedule for Children-Revised	7 – 16	Youth					X			
GHQ / GHQ-12	General Health Questionnaire	11 – 15	Youth	X							
HRSD	Hamilton Rating Scale for Depression	adolescents	Youth				X				
	How-I-Feel Questionnaire	8 – 15	Youth								X
HSC-37A/ HSCL-37A	Hopkins Symptom Checklist-37 for Adolescents	12 – 18	Youth	X							
IES / IES-R	Impact of Event Scale / Impact of Event Scale - Revised	8 – 18	Youth						X		
KHS / HSC	Kazdin Hopelessness Scale / Hopelessness Scale for Children	5 – 13	Youth				X				
KSADS	Schedule for Affective Disorders and Schizophrenia for School-Age Children (Present & Lifetime Version)	6 – 18	Youth; Parent		X	X	X	X	X		
LSS	Langner Symptom Survey	15 – 18	Youth	X							
MASC / MASC-2	Multidimensional Anxiety Scale for Children	Children & adolescents	Youth; Parent					X			
	Multidimensional Anxiety Scale for Children (2 nd edition)	8 – 19	Youth; Parent					X			
MAST	Multi-Attitude Suicide Tendency Scale	Adolescents & adults	Youth				X				
MFQ	Mood and Feelings Questionnaire	8 – 18	Youth; Parent				X				
MSPSE	Multidimensional Scales of	11 – 18	Youth							X	

Acronym	Child Adjustment Measure	Validated for years	Report	Overall	Internalizing	Externalizing	Depression; Suicide	Anxiety	Trauma	Self-esteem	Peer Social Support
	Perceived Self-efficacy										
MSLSS	Multi-dimensional Student Life Satisfaction Scale	8 – 18	Youth							X	
PANSI	Positive and Negative Suicide Ideation Inventory	14 – 17	Youth				X				
PDR	Parent Daily Report	4 – 10	Parent Observation			X					
PEDS	Pediatric Emotional Distress Scale	2 – 10	Parent						X		
PHCSCS	Piers-Harris Children's Self-Concept Scale, 2 nd edition	7 – 18	Youth							X	
PHQ-9	Patient Health Questionnaire – 9 items (depression scale)	13+	Youth				X				
Y-PSC / PSC	Pediatric Symptom Checklist / PSC-Youth Report	Y: 9 – 14 P: 3 – 18	Parent; Youth	X							
PSSM	Psychological Sense of School Membership	9 – 16	Youth								X
PSWQ-C	Penn State Worry Questionnaire for Children	7 – 17	Youth					X			
PTSS-C	Post-traumatic Stress Symptoms in Children	6 – 18	Semi-structured interview						X		
RADS	Reynolds Adolescent Depression Scale	13 – 18	Youth; Parent				X				
RATS	Reactions of Adolescents to Traumatic Stress	10+	Youth						X		
RCADS	Revised Children's Anxiety & Depression Scale	6 – 18	Youth; Parent		X		X	X			
RCDS	Reynolds Child Depression Scale	8 – 12	Youth; Parent				X				
RCMAS	Revised Children's Manifest Anxiety Scale	6 – 19	Youth; Parent					X			
RFL-A	Reasons for Living Inventory for Adolescents	14 – 18	Youth				X				
RSES	Rosenberg Self-Esteem Scale/Inventory	13 – 19	Youth							X	
SASC-R	Social Anxiety Scale for Children-Revised	7 – 18	Youth					X			
SCARED	Screen for Child Anxiety Related Emotion Disorders	6 – 18	Youth; Parent					X			
SCAS	Spence Children's Anxiety Scale	7 – 19	Youth; Parent					X			
SCL-90-R	Symptom Checklist-90-Revised	13+	Youth	X							
SDQ	Strengths & Difficulties Questionnaire	P & T: 3 – 16 Y: 11 – 16	Youth; Parent; Teacher	X							
SDQ-I	Self-Description Questionnaire (preadolescents)	Gr. 2 – 6	Youth							X	
SDQ-II	Self-Description Questionnaire (early, middle adolescents)	Gr. 7 – 12	Youth							X	

Acronym	Child Adjustment Measure	Validated for years	Report	Overall	Internalizing	Externalizing	Depression; Suicide	Anxiety	Trauma	Self-esteem	Peer Social Support
SEI	Self-esteem Inventory (Coopersmith SEI)	8 – 15	Youth							X	
SET/PYIT	Social Evaluative Task/Parent-Youth Interaction Task	Youth	Observation					X			
SIQ	Suicidal Ideation Questionnaire (high school students)	Gr. 10 – 12	Youth				X				
SIQ-Jr	Suicide Ideation Questionnaire (middle school students)	Gr. 7 - 9	Youth				X				
SITBI	Self-injurious Thoughts & Behaviors Interview	12 – 19	Interview				X				
SOCSS	Survey of Children's Social Support	Gr. 3 – 5	Youth								X
SPAIC	Social Phobia and Anxiety Inventory for Children	9 – 14	Youth					X			
SPPCA	Self-Perception Profile for Children & Adolescents	Children & adolescents	Youth; Teacher							X	
SPS	Suicide Probability Scale	Adolescents	Youth				X				
SSI	Scale for Suicide Ideation	adolescents	Youth (interview)				X				
SSMS	Social Support Microsystems Scales (peer, family, school)	10 – 18	Youth								X
STAI-C	State-Trait Anxiety Inventory for Children	9 – 12	Youth					X			
TSCC	Trauma Symptom Checklist for Children	8+	Youth						X		
TSCYC	Trauma Symptom Checklist for Young Children	3 – 12	Parent						X		
TP	Youth Top Problems	7 – 13	Youth; Parent; Clinician	X							
UCLA PTSD-RI	UCLA PTSD Reaction Index for DSM IV	7 – 18	Youth						X		

Acronym	Parent Adjustment Measure	Report (SR = Self-report)	Overall	Internalizing	Externalizing	Depression; Suicide	Mania	Anxiety	Trauma/PTSD	Aggression	Substance Use	Social Support
ADIS-IV	Anxiety Disorders Interview Schedule, Lifetime version for DSM-IV	Interview				X		X				
ADIS-VL	Anxiety and Related Disorders Interview Schedule for DSM-5 – Lifetime Version	Interview				X						
ADS	Alcohol Dependence Scale	Adult SR									X	
ASRM	Altman Self-Rating Mania Scale	Adult SR					X					
AUDIT	Alcohol Use Disorders Identification Test	Adult SR									X	
BAI	Beck Anxiety Inventory	Adult SR						X				
BDI-II	Beck Depression Inventory, 2 nd edition	Adult SR				X						
BDI	Beck Depression Inventory	Adult SR				X						
BSI	Brief Symptom Inventory	Adult SR	X									
CAGE	CAGE (screening scale for alcohol)	Adult SR									X	
CAPS/CAPS-IV/CAPS-5	Clinician-Administered PTSD Scale/ for DSM-IV/ for DSM-5	Clinician administered							X			
CAQ	Cognitive Avoidance Questionnaire	Adult SR						X				
CES-D	Center for Epidemiological Studies Depression Scale	Adult SR				X						
CIDI-SAM	Composite International Diagnostic Interview-Substance Abuse Module	Interview									X	
CORE-OM	Clinical Outcomes in Routine Evaluation - Outcome Measure	Adult SR				X						
C-SSRS	Columbia-Suicide Severity Rating Scale	Clinician rating										
CUDOS	Clinically Useful Depression Outcome Scale	Adult SR				X						
CUXOS	Clinically Useful Anxiety Outcome Scale	Adult SR						X				
DASS/DASS-21	Depression Anxiety Stress Scales	Adult SR				X						
DUSI-R	Drug Use Screening Inventory-Revised	Adult SR									X	
EPDS	Edinburgh Postnatal Depression Scale	Adult SR				X						
EQ	Experiences Questionnaire	Adult SR				X						
FAST	FAST (screening scale for alcohol)	Adult SR									X	
FQ	Fear Questionnaire	Adult SR						X				
GAD-7	Generalized Anxiety Disorder Screener	Adult SR						X				

Acronym	Parent Adjustment Measure	Report (SR = Self-report)	Overall	Internalizing	Externalizing	Depression; Suicide	Mania	Anxiety	Trauma/PTSD	Aggression	Substance Use	Social Support
PSS	PTSD Symptom Scale	Interview; Adult SR							X			
PSWQ	Penn State Worry Questionnaire	Adult SR						X				
QIDS(-SR)	Quick Inventory of Depressive Symptomatology (Self-rated) (shortened version of IDS)	Adult SR; Clinician				X						
QLQ	Quality of Life Questionnaire	Adult SR						X				
QOLI	Quality of Life Inventory	Adult SR				X		X				
RAPS	Rapid Alcohol Problems Screen	Adult SR									X	
RAS	Recovery Assessment Scale	Self-report & Clinician Rating	X									
RFL	Reasons for Living Inventory	Adult SR				X						
SAAST	Self-Administered Alcoholism Screening Test	Adult SR									X	
SASII	Suicidal Attempt Self-Injury Interview	Structured interview				X						
SBQ	Suicidal Behaviors Questionnaire	Adult SR				X						
SBQ-R	Suicide Behaviors Questionnaire – Revised	Adult SR										
SCL-90-R	Symptom Checklist-90-Revised	Adult SR	X									
SCID-5/SCID-IV	Structured Clinical Interview for DSM-5, DSM-IV / DSM-IV Personality Disorders	Clinician administered	X	X	X	X	X	X	X		X	
SITBI	Self-injurious Thoughts & Behaviors Interview	Interview				X						
SPIN	Social Phobia Inventory	Adult SR						X				
SSI	Scale for Suicide Ideation	Interview				X						
THQ	Trauma History Questionnaire	Adult SR; Clinician							X			
THS	Trauma History Screen	Adult SR							X			
TLEQ	Traumatic Life Events Questionnaire	Adult SR							X			
TOP	Treatment Outcome Package	Adult SR				X						
TWEAK	TWEAK (screening measure for alcohol)	Adult SR									X	
WAQ	Worry and Anxiety Questionnaire	Adult SR						X				
WHODAS 2.0	World Health Organization Disability Assessment Schedule 2.0	Adult SR				X						
WW-II	Why Worry-II	Adult SR						X				
YMRS	Young Mania Rating Scale	Clinician rating					X					

Acronym	Parenting Measure	Report
APQ	Alabama Parenting Questionnaire	Parent; Youth
BICS, BISS	Beavers Interactional Competence Scale; BI Style Scale	Observation
CAPI	Child Abuse Potential Inventory	Parent
CARE-Index	Child and Adult Relational Experimental Index	Observation
Circumplex CRS	Circumplex Clinical Rating Scale	Observation
DPCICS	Dyadic Parent–Child Interaction Coding Syst.	Observation
FACES IV	Family Adaptability and Cohesion Evaluation Scales - Fourth Edition (based on Circumplex Model)	Parent
FAD/MFAD	Family Assessment Device/McMaster Family Assessment Device	Parent
GARF	Global Assessment of Relational Functioning	Clinician rating
MCBS	Management of Children’s Behavior Scale	Parent
McMaster CRS	McMaster Clinical Rating Scale	Observation
McSIFF	McMaster Structured Interview for Family Functioning	Structured interview
PAC	Parent–Adolescent Communication Scale	Youth; Parent
PBI; CRPBI	Parent Behaviour Inventory (aka: Children’s Report on Parent Behavior Inventory; CRPBI)	Youth
PBIQ	Parent Behavior Importance Questionnaire	Parent
PIRGAS	Parent Infant Rel. Global Assessment Scale	Clinician rating
PS	Parenting Scale	Parent
PSD (PSDQ)	Parenting Styles and Dimensions (Questionnaire)	Parent
PSI	Parenting Stress Index	Parents of children aged 1 - 12
SFI	Self-report Family Inventory (Beavers)	Parent

Appendix F: Methodological Quality Coding Criteria

Methodological Quality: Coding Criteria

Components drawn from the “Checklist for Measuring Quality” (Downs & Black, 1998)

Article Authors:

Year:

Coder:

Reporting	Yes	No	n/a
1. <i>Is the hypothesis/aim/objective of the study clearly described?</i> If the authors describe the formative research, theoretical basis(es) or constructs upon which the intervention was developed the question should be answered yes*.	1	0	n/a
2. <i>Are the main outcomes to be measured clearly described in the Introduction or Methods section?</i> If main outcomes are first mentioned in the Results section, question should be answered no.	1	0	n/a
3. <i>Are the characteristics of the participants included in the study clearly described?</i> In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and the source for controls should be given.	1	0	n/a
4. <i>Are the main findings of the study clearly described?</i> Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical tests which are considered below).	1	0	n/a
5. <i>Does the study provide estimates of the random variability in the data for the main outcomes?</i> In non-normally distributed data, the <u>inter-quartile range</u> of results should be reported. In normally distributed data the <u>standard error, standard deviation or confidence intervals</u> should be reported. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered yes.	1	0	n/a
6. <i>Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?</i> Focus on the adjustment outcomes of interest to the meta-analysis (depression, trauma, anxiety, externalizing/internalizing behaviour, etc)	1	0	n/a
External validity: All the following criteria attempt to address the representativeness of the findings of the study and whether they may be generalised to the population from which the study subjects were derived.	Yes	No	U/D
7. <i>Were the participants asked to participate in the study representative of the entire population from which they were recruited?</i> The study must identify the source population for patients and describe how the patients were selected. Patients would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample. Random sampling is only feasible where a list of all members of the relevant population exists. <u>Where a study does not report the proportion of the source population from which the patients are derived, the question should be answered as unable to determine.</u>	1	0	0
8. <i>Were those participants who were prepared to participate representative of the entire population from which they were recruited?</i> The <u>proportion of those asked who agreed should be stated</u> . Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.	1	0	0
Internal validity – bias	Yes	No	U/D

9. <i>If any of the results of the study were based on “data dredging”, was this made clear?</i> Any analyses NOT planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.	1	0	0
10. <i>Were the statistical tests used to assess the main outcomes appropriate?</i> The statistical techniques used must be appropriate to the data. For example, nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes. ** As with any consideration of statistical analysis, consideration should be given to whether there was a more appropriate alternate statistical method that could have been used. The methods section should be detailed enough for reviewers to <u>identify which analytical techniques were used</u> (in particular, regression or stratification) and how specific confounders were measured. For studies utilizing regression analysis, it is useful to identify if the study identified which variables were included and how they related to the outcome. If stratification was the analytical approach used, were the strata of analysis defined by the specified variables? It is also important to assess the appropriateness of the analytical strategy in terms of the assumptions associated with the approach as differing methods of analysis are based on differing assumptions about the data and how it will respond.**	1	0	0
11. <i>Were the main outcome measures used accurate (valid and reliable)?</i> For studies where the outcome measures are clearly described, the question should be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes. – This was a selection criteria	1 – p.	0	0
Power	Yes	No	
12. <i>* Did the study mention having conducted a power analysis to determine the sample size needed to detect a significant difference in effect size for one or more outcome measures?*</i> Yes, 2 or more measures = score 2; Yes, 1 measure = score 1; No = score 0	2	1	0
TOTAL (out of 13)			

Note: n/a = Not Applicable; U/D = Unable to Determine

Items drawn from the “Checklist for Measuring Quality” (Downs & Black, 1998)

* Item/Explanation drawn from: Kennelly, J. (2011) Methodological Approach to Assessing the Evidence. In A. Handler et al. (eds.), Reducing Racial/Ethnic Disparities in Reproductive and Perinatal Outcomes: The Evidence from Population-Based Interventions (pp. 7–19). DOI 10.1007/978-1-4419-1499-6_2

** Explanation drawn from: The Joanna Briggs Institute Critical Appraisal tools for use in JBI Systematic Reviews: Checklist for Analytical Cross Sectional Studies (https://joannabriggs.org/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Analytical_Cross_Sectional_Studies2017_0.pdf) – recommended by NICE guidelines

Appendix G: Sample Practitioner Session Reflection Tool

Session Reflection Tool: Triple P Level 4 Group, Session 1

<i>Session Components: Please indicate whether you completed each component of the session as described in the manual, or if you modified it</i>	<i>Activity completed as described</i>	<i>Activity modified (How?)</i> Check all that apply
1. Introduction: • Congratulate parents on making the commitment to participate • Provide an overview of Group Triple P • If any parent did not complete Assessment Booklet One or is just joining the group have him/her complete the assessment packet before starting this session	☐	☐ I added elements to this activity ☐ I removed elements from this activity ☐ I adapted this activity ☐ I replaced an element of this activity with another element ☐ I dropped this activity entirely
2. Agenda (outline proposed session goals and gain consent from group) • Behavior Rating Scales • Working as a group • What is positive parenting • Causes of child behavior problems • Goals for change • Keeping track of children's behavior • Homework tasks	☐	☐ I added elements to this activity ☐ I removed elements from this activity ☐ I adapted this activity ☐ I replaced an element of this activity with another element ☐ I dropped this activity entirely
3. Behavior Rating Scales • Distribute and have caregivers complete the Behavior Rating Scales (~2 mins) • Immediately collect Behavior Rating Scales	☐	☐ I added elements to this activity ☐ I removed elements from this activity ☐ I adapted this activity ☐ I replaced an element of this activity with another element ☐ I dropped this activity entirely
4. Working as a Group • Exercise 1: Setting basic ground rules for the group • Exercise 2: Getting to know you • Exercise 3: What you would like to get out of the group session	☐	☐ I added elements to this activity ☐ I removed elements from this activity ☐ I adapted this activity ☐ I replaced an element of this activity with another element ☐ I dropped this activity entirely
5. What is Positive Parenting? • Discuss the challenges of being a parent • Show video: Opening segment • Introduce positive parenting: ○ Ensuring a safe and interesting environment ○ Creating a positive learning environment ○ Using assertive discipline ○ Having realistic expectations ○ Taking care of oneself • Show video Part 1: What is Positive Parenting? • Exercise 4: What is positive parenting?	☐	☐ I added elements to this activity ☐ I removed elements from this activity ☐ I adapted this activity ☐ I replaced an element of this activity with another element ☐ I dropped this activity entirely
6. Causes of Child Behavior Problems • Introduce factors know to influence children's behavior ○ Genetic make-up ○ The family environment ○ Influences outside the home • Show video Part 2: Causes of Child Behavior Problems • Exercise 5: Identifying causes of child behavior problems	☐	☐ I added elements to this activity ☐ I removed elements from this activity ☐ I adapted this activity ☐ I replaced an element of this activity with another element ☐ I dropped this activity entirely

Session Components: Please indicate whether you completed each component of the session as described in the manual, or if you modified it	Activity completed as described	Activity modified (How?) Check all that apply
7. Goals for Change - Exercise 6: What skills should we encourage in children? - Exercise 7: Setting goals for change	☐	☐ I added elements to this activity ☐ I removed elements from this activity ☐ I adapted this activity ☐ I replaced an element of this activity with another element ☐ I dropped this activity entirely
8. Keeping Track of Children's Behavior - Introduce monitoring of children's behavior - Show video Part 2: Goals for Change, Keeping Track - Exercise 8: Keeping track	☐	☐ I added elements to this activity ☐ I removed elements from this activity ☐ I adapted this activity ☐ I replaced an element of this activity with another element ☐ I dropped this activity entirely
9. Session Close - Review the main points covered in the session - Explain why the homework tasks are important - Explain homework tasks (monitoring one or two child behavior) - Ask if any parent would like to take the DVD home to review - List name of parent checking out video _____ - Outline the content of the next session and give a reminder of the day and time	☐	☐ I added elements to this activity ☐ I removed elements from this activity ☐ I adapted this activity ☐ I replaced an element of this activity with another element ☐ I dropped this activity entirely

The next set of questions are about the modifications you made¹:

¹ Categories drawn from Taylor et al. (2015) and Asgary-Eden & Lee (2011); *Service Providers' Reasons for Omitting Session Objectives*: Asgary-Eden & Lee, p. 171

Why: What is your rationale for making this modification? Check all that apply

1) I did not have enough time in session 2) This activity was not relevant for parents in this group 3) Forgot to do the activity 4) Have found that the activity did not help parents in the past 5) I did not have the necessary material 6) I did not think parents would be comfortable doing the activity	7) I did not think parents would understand the activity 8) I was not comfortable doing the activity 9) I did not feel trained well enough 10) Was instructed not to do activity by my supervisor 11) I did not think it fit my client's culture 12) Other (please describe):	Details: Text response: "Please describe how [you added/changed]" (provides qualitative data)
---	--	---

Item	Please pick the response that best describes the way you worked with the parents during this session¹ :	Not at all	A little bit	A lot	Not applicable in this session
1.	I invited the parents to develop their own parenting goals	☐	☐	☐	☐
2.	I invited the parents to monitor their own behaviors	☐	☐	☐	☐
3.	I invited the parents to monitor the behavior of their children	☐	☐	☐	☐
4.	I invited the parents to select the strategies they want to employ	☐	☐	☐	☐
5.	When discussing parenting strategies employed by the parents, I invited them to identify what went well	☐	☐	☐	☐
6.	When difficulties were noted in using the parenting strategies, I invited the parents to identify what they could do differently	☐	☐	☐	☐
7.	I invited the parents to recognize the gains they have made	☐	☐	☐	☐
8.	I invited the parents to use the parenting practices I introduced across different contexts	☐	☐	☐	☐

¹ Items drawn from Sanders & Mazzucchelli (2013) discussion of guiding parents in learning Self-management Tools (p. 2) to promote Self-regulation: Item 1 (self-determination of parenting goals); Items 2, 3 (self-monitoring of specific parent and child behaviors over course of intervention); Item 4 (self-selection of change strategies); Items 5, 6 (self-evaluation of whether a parent achieved some performance criterion); Item 7 (self-reward); Item 8 (skills generalization)

Appendix H: Practitioner and Coder Ratings of Adherence

Practitioner and Coder Ratings of Adherence

	RATER	
Adherence	Practitioner	Coder
Content	PSRT-part 1 Includes items from session checklist plus questions about changes Completed after session	CSRT-part 1 Includes items from session checklist plus questions about changes Completed while reviewing recording
	PSRT-part 2 Includes items about promotion of self- regulation Completed after session	CRST-part 2 Includes items about promotion of self- regulation Completed after reviewing recording
Process		AMPQ Includes items about promotion of self- regulation Completed after reviewing recording

Appendix I: Power Analysis

Power Analysis: Logistic Regression

