

Evaluation of the Positive Discipline in Everyday Parenting Program

Elena Gallitto

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

School of Psychology
Faculty of Social Sciences
University of Ottawa

© Elena Gallitto, Ottawa, Canada, 2021

Abstract

The Positive Discipline in Everyday Parenting (PDEP) program is a group-based parenting program aimed at eliminating the use of physical and emotional punishment of children and promote positive discipline. My dissertation consists of three separate studies. Study 1 evaluated the psychometric qualities of the scales included in the current version of the PDEP questionnaire by way of item-level analyses conducted using the Item Response Theory (IRT) method. Pending the results, scale items were modified, and the revised version of the questionnaire was administered to two expert samples (e.g., researchers and professionals as well as parents) to assess content validity. Following the content validity analyses, the revised version of the PDEP questionnaire was created. Study 2 was conducted to assess the factor structure of the scales of the new revised version of the PDEP questionnaire via IRT confirmatory analyses by using data from a sample of 400 parents recruited through an online platform. Moreover, the reliability of the PDEP questionnaire was assessed in terms of both internal consistency and test re-test reliability. For the latter, one-fourth of the total sample of parents ($n = 100$) filled out the same questionnaire after four weeks from the initial administration. In addition, the sample data were also used to gather evidence to support the convergent validity of the final revised version of the questionnaire. For this purpose, the PDEP questionnaire scales were correlated to similar psychological constructs to determine the degree of association among the selected measures. Finally, Study 3 examined the efficacy of the PDEP program in decreasing parents' approval of punishment and reducing negative attributions of typical child behaviours and in improving parenting self-efficacy, as well as other constructs of interest including parental emotion regulation, parenting stress, problem-solving, and parenting practices. This evaluation was carried out by way of a quasi-experimental design that included a wait-list control group.

Acknowledgements

I would like to express my gratitude and appreciation for all the families that took the time to participate in this research, and without whom I would have no content for my thesis.

I am especially indebted to my supervisor Dr. Elisa Romano, whose guidance, support and encouragement have been invaluable throughout this study. Thank you for helping me expand my abilities as a researcher and for pushing me to my best work. Thank you to each of the members of my Dissertation Committee for their generosity and suggestions. I am also grateful to all of the students with whom I have had the pleasure to work during this and other related projects.

I also owe much gratitude to my mother and father, Iolanda Placenti and Eugenio Gallitto, for offering their support and unconditional love throughout this process. You have always believed in me and comforted me during my many struggles in life. Thank you for teaching me the importance of practicing self-care and being open and receptive to love. This is one of many qualities that have helped me immensely during these years as a student, wife and mother. Through you, I have also learned to embrace the joy of learning and seeking new opportunities. Without the courage and persistence that you have instilled in me, I would not be here today. You have allowed me to go beyond what I know to understand myself and others; only then could be a better parent to my children. Some special words of gratitude to my husband Craig who has always been a major source of support when things would get a bit discouraging. Lastly, thank you to my two children, Rocco and Raffaello who have been the light of my life for the last seven years and continue to provide unending inspiration. To you, I dedicate this work.

Content of Thesis and Contributions of Authors

This dissertation follows a multiple-article format and is divided into three studies: (1) The preliminary analysis and content validity of the current version of Positive Discipline and Everyday Parenting (PDEP) questionnaire; (2) Item-level analysis, reliability and construct validity assessment of the revised version of the PDEP questionnaire; (3) Quasi-experimental evaluation of the PDEP program. Study 1 and 2 are currently in the process of being written up as a manuscript to be submitted to a peer-review journal. The writer of the thesis will appear as the first author and the thesis supervisor will appear as a co-author. Ms. Gallitto took the lead in every aspect of this dissertation, including the literature review and conceptualization of the project, development and implementation of study procedures and methods, ethics review board applications, data collection, data scoring and analysis, and writing of the thesis. Dr. Romano adopted an advisory role throughout the process and oversaw all abovementioned activities.

Table of Contents

<i>Abstract</i>	<i>ii</i>
<i>Acknowledgements.....</i>	<i>iii</i>
<i>Content of Thesis and Contributions of Authors</i>	<i>iv</i>
<i>.....</i>	<i>iv</i>
<i>General Introduction.....</i>	<i>1</i>
Physical Punishment and Children's Rights	1
Physical Punishment Rates in Canada and Links to Children's Outcomes	3
Parenting Programs to Reduce the Use of Physical Punishment	5
Taking Action to Eliminate Physical Punishment: The PDEP Program.....	11
Dissertation Objectives	16
<i>Study 1: Evaluating and Modifying the Current Positive Discipline in Everyday Parenting (PDEP)</i>	
<i>Questionnaire</i>	<i>18</i>
Abstract.....	18
Introduction.....	19
The PDEP Program: Objectives and Theoretical Framework	19
The PDEP Questionnaire.....	22
Statistical Methods for Evaluating Measurement Scales: Classical Test Theory and Item Response Theory	24
Study Objectives	27
<i>Methods.....</i>	<i>28</i>
Sample and Procedures.....	28
Measures.....	31
Sociodemographic questionnaire.	31
The PDEP questionnaire (Durrant et al., 2014).	31
Content validity form.....	32
Data Analyses	33
Data screening procedures.	33
Testing IRT assumptions: Unidimensionality and local independence.	35
IRT analyses: The graded response model.	36
Content validity analyses.	38
<i>Results</i>	<i>38</i>
Unidimensional Analysis.....	38
IRT Results for the Approval of Physical Punishment Scale	39
IRT results for Subjective Norms Scale	40
IRT Results for the Parenting Self-Efficacy Scale	41
Summary of Preliminary IRT Findings and Initial Modifications to the PDEP Questionnaire	42
PDEP Item Ratings for Experts and Parents	45

Parent Focus Group Findings	49
Summary of Content Validity Analysis Results	50
Discussion	52
<i>Study 2: Examining the Psychometric Properties of the Revised Positive Discipline in Everyday Parenting Questionnaire</i>	<i>55</i>
Abstract	55
Study Hypotheses	62
<i>Method</i>	<i>63</i>
Procedures and Sample	63
Measures	65
Socio-demographic questionnaire. Participants answered several socio-demographic questions (i.e., age, gender, ethnic/racial background, income), as well as questions concerning their recruitment (e.g., community centres, friends, social media)	65
Revised PDEP questionnaire	65
Attitudes Toward Spanking Scale (ATS; Holden et al., 1995)	66
Parent Cognition Scale (PCS; Snarr, Slep, & Grande, 2009)	66
Self-Efficacy for Parenting Tasks Scale – Discipline subscale (SEPTI; Coleman & Karraker, 2000)	67
Data Analyses	67
Data screening for the IRT analysis	67
Testing IRT assumptions	67
IRT model assessment	68
Internal consistency, test-retest reliability, and convergent validity.	68
<i>Results</i>	<i>69</i>
Data Screening	70
Testing the Assumptions of Unidimensionality and Local Dependence	70
IRT Results for the Approval of Physical Punishment Scale	71
IRT Results for the Approval of Non-Physical Punishment Scale	72
IRT Results for Subjective Norms Scale	73
IRT Results for the Parenting Self-Efficacy Scale	74
Summary of the IRT Results	75
<i>Discussion</i>	<i>77</i>
<i>Study 3: A Quasi-Experimental Evaluation of the Positive Discipline in Everyday Parenting Program</i>	<i>82</i>
Abstract	82
Physical Punishment: A Threat to Children’s Mental Health and Well-Being	83
Physical Punishment and Family Violence	85
Mechanisms Underlying the Impact of Physical Punishment on Children’s Development	86
Use of Non-Physical Punishment: A Current Controversial Issue	87
Use of time-out	89
Parental psychological control	91
Parental verbal aggression	92

Determinants of Physical Punishment Use	93
Attitudes toward physical punishment	93
Attributions about children's behaviours	94
Parenting self-efficacy	95
Parenting stress	96
Problem-solving and disciplinary strategies	97
Parental emotion regulation	98
The Positive Discipline in Everyday Parenting Program	99
Research Questions and Hypotheses	102
<i>Method</i>	103
Participants and Procedures	103
Measures	107
Socio-demographics	107
Cognitive/affective processes	107
Action goals.	109
Parental Behaviour.	111
Treatment attendance	113
Data Analysis	113
Standard data screening procedures.	113
Screening procedures for hierarchical linear modeling	114
Interclass correlation coefficient and power analysis: Determining the number of clusters in hierarchical linear modelling.	115
Growth curve modelling: Measuring the rate of change in the outcome variables.	116
Results	120
Differences between the Experimental and Wait-list Control Group on Pre-PDEP Outcome Variables	120
Correlations Among Variables at Pre-PDEP	121
Changes in Outcome Variables Over Time Between the PDEP Experimental and Wait-list Control Group ..	122
Bonferroni Adjustment for Multiple Comparisons: Follow-up to the Analyses on the Changes in the Outcome Between the Experimental and Control Groups.	129
Treatment Attendance as Moderator of the Change in Outcomes for the PDEP-Experimental Group.	129
Discussion	129
Limitations and Future Research Directions	135
General Discussion	136
Implications	138
Theoretical/research implications	138
Clinical/applied implications	139
References	141

List of Tables

Table 1. Mean, Standard Deviation, Skewedness and Kurtosis of the PDEP items (current version)	167
Table 2. Model Fit Indices for the PDEP Scale before the Revisions to the Questionnaire.....	168
Table 3. Discrimination and Threshold Parameters for Approval of Physical Punishment Items.....	169
Table 4. SD- χ^2 Values of Approval of Physical Punishment Items.....	170
Table 5. Discrimination and Threshold Parameters for Approval of Non-physical Punishment Items.....	171
Table 6. SD- χ^2 Values of Approval of Non-Physical Punishment Items.....	172
Table 7. Discrimination and Threshold Parameters for Subjective Norms Items.....	173
Table 8. SD- χ^2 Values of Subjective Norms Items.....	174
Table 9. Discrimination and Threshold Parameters for Parenting Self-Efficacy Items.....	175
Table 10. SD- χ^2 Values of Parenting Self-Efficacy Items.....	176
Table 11. PDEP Questionnaire – First revisions following Preliminary Analyses.....	177
Table 12. Content Validity Assessment Content Experts ($n = 8$).....	181
Table 13. Content Validity Assessment Parents ($n = 12$).....	189
Table 14. Changes to PDEP Items Based on the Content Experts' and Parents' Feedback.....	199
Table 15. Demographic Characteristics of Parents Participating in the Online Survey ($n = 400$).....	202
Table 16. Demographic Characteristics of Parents Participating in the Test-Retest Study ($n = 100$).....	205
Table 17. Mean, Standard Deviation, Skewedness and Kurtosis of the Revised PDEP Questionnaire Items.....	207
Table 18. Model Fit Indices for the PDEP Scales after Revisions.....	209
Table 19. Discrimination and Threshold Parameters	

for Approval of Physical Punishment Items.....	210
Table 20. SD χ^2 Values of Approval of Physical Punishment Items.....	212
Table 21. Discrimination and Threshold Parameters for Approval of Non-Physical Punishment Items.....	213
Table 22. SD- χ^2 Values of Approval of Non-Physical Punishment Items.....	215
Table 23. Discrimination and Threshold Parameters for Subjective Norms Items.....	216
Table 24. SD- χ^2 Values of Subjective Norms Items.....	218
Table 25. Discrimination and Threshold Parameters for Parenting Self-Efficacy Items.....	219
Table 26. SD- χ^2 Values of Parenting Self-Efficacy Items.....	221
Table 27. Means and Standard Deviations of All Study Variables (Convergent Validity Assessment).....	222
Table 28. Demographic Characteristics of Parents Participating in the Quasi-Experimental Evaluation Study (PDEP Experimental vs. Wait-List Control Group).....	223
Table 29. Unconditional/Null Growth Curve Models for Parenting Outcomes.....	225
Table 30. Means (Standard Deviations) of All Variables at Pre-Program for the Two PDEP Groups.....	230
Table 31. Correlation Among All Outcome Variables at the Pre-PDEP Data Collection.....	232
Table 32. Fixed Effects of the Conditional Growth Curve Models for All the Parenting Outcomes.....	234
Table 33. Fixed Effects of the Conditional Growth Curve Models for the PDEP Experimental group across the Parenting Outcomes with Treatment Attendance as Moderator.....	238

List of Figures

Figure 1. PDEP conceptual model.....	243
Figure 2. Information Functions for Approval of Physical Punishment Items (PDEP original version).....	244
Figure 3. Information Functions for Approval of Non-physical Punishment Items (PDEP original version).....	245
Figure 4. Information Functions for Subjective Norms Items (PDEP original version).....	246
Figure 5. Information Functions for Parenting Self-Efficacy Items (PDEP original version).....	247
Figure 6. Information Functions for Approval of Physical Punishment Items (PDEP revised version).....	248
Figure 7. Information Functions for Approval of Non-Physical Punishment Items (PDEP revised version).....	249
Figure 8. Information Functions for Approval of Subjective Norms (PDEP revised version).....	250
Figure 9. Information Functions for Parenting Self-Efficacy Items (PDEP revised version).....	251
Figure 10. Flow chart sample size Study.....	252

List of Appendices

Appendix A: Ethics Approval Study 1 (Objective 1)	253
Appendix B: Ethics Approval Study 1 (Objective 2)	254
Appendix C: Recruitment Notice Study 1, Objective 2 (Content Validity Experts)	255
Appendix D: Study 1, Objective 2 (Consent Form Experts)	256
Appendix E: Recruitment notice Study 1, Objective 2 (Content Validity Parents)	258
Appendix F: Study 1 – Objective 2 (Consent Form Parents)	259
Appendix G: Interview Protocol for the Focus Group.....	261
Appendix H: Sociodemographic Questionnaire (Experts).....	262
Appendix I: Sociodemographic Questionnaire (Parents).....	264
Appendix J: Content Validity Forms (Experts and Parents)	267
Appendix K: Recruitment Notice Study 1 (Convergent Validity)	273
Appendix L: Consent Form- Validity and Reliability Assessment.....	274
Appendix M: Parent Cognition Scale.....	276
Appendix N: Discipline subscale of the Self-Efficacy for Parenting Tasks scale.....	277
Appendix O: Ethics Approval Study 3.....	278
Appendix P: Recruitment Notice Study 3.....	279
Appendix Q: Consent Form- Study 3	280
Appendix R: Facilitator Fidelity Forms	284
Appendix S: Socio-Demographic Questionnaire – Study 3.....	305
Appendix T: DERS -SF- Reduced version.....	309
Appendix U: Parenting Daily Hassles Scale.....	310
Appendix V: Problem Solving Inventory.....	311
Appendix W: Parenting Young Children Scale -Modified.....	312

General Introduction

Physical Punishment and Children's Rights

Physical punishment is a concerning societal problem that is inherent to children's rights to protection from violence. The United Nations Committee on the Rights of the Child (2006) defines physical punishment as

“any punishment in which corporal force is used and intended to cause some degree of pain or discomfort, however, light. Most involve hitting (“smacking”, “slapping”, “spanking”) children, with the hand or with an implement – whip, stick, belt, shoe, wooden spoon, etc. But it can also involve, for example, kicking, shaking or throwing children, scratching, pinching, biting, pulling hair or boxing ears, forcing children to stay in uncomfortable positions, burning, scalding or forced ingestion (for example, washing children's mouths out with soap or forcing them to swallow hot spices)” (p. 4).

Physical punishment contravenes the United Nations Convention on the Rights of the Child, which was ratified by Canada in 1991. There are several articles in the Convention that pertain to physical punishment and other forms of violence against children. More specifically, Article 19 states that children have the right to be protected against any form of violence (whether physical or psychological) and Article 37 which states that children have the right to be protected from any form of cruel and harmful punishment (UN General Assembly, 1989).

In addition to the Convention, several organizations have condemned physical punishment as a public health issue leading to severe negative consequences for children's development, including the World Health Organization (Krug, Mercy, Dahlberg, & Zwi, 2002). In Canada, more than 500 child- and family-serving organizations have adopted formal positions recommending the elimination of physical punishment. These include, but are not limited to, the following: Canadian Academy of Child and Adolescent Psychiatry; Canadian Institute of Child Health; Canadian Mental Health Association; Canadian Public Health Association; Canadian Pediatric Society; Canadian Psychological Association; Centre for Addiction and Mental Health; Child and Youth Health Network for Eastern Ontario;

Children's Mental Health Ontario; Infant Mental Health Promotion; Mental Health Commission of Canada; National Institute of Families for Child and Youth Mental Health; and the Ontario Centre of Excellence for Child and Youth Mental Health (Durrant & Ensom, 2004). The Joint Statement on Physical Punishment of Children and Youth (Durrant & Ensom, 2004) was a Canadian initiative developed by a national coalition of six organizations facilitated by the Children's Hospital of Eastern Ontario (Ottawa, Ontario) to inform the public of the deleterious consequences of child physical punishment and to move toward its elimination. Additionally, contemporary global initiatives have supported the idea that physical punishment is a serious violation of children's rights and that it is our social responsibility to ensure children's protection from all forms of violence (Global Initiative to End All Corporal Punishment of Children, 2017).

Despite efforts to end physical punishment, Section 43 of the Canadian Criminal Code (Criminal Code, 1985) gives certain individuals the right to use physical punishment under various circumstances. It reads as follows: "Every school teacher, parent, or person standing in the place of the parent is justified in using force by way of correction toward the pupil or child, as the case may be, who is under his care, if the force does not exceed what is reasonable under the circumstances" (s.43, Criminal Code, 1985). In 2004, the Supreme Court imposed several limitations to Section 43, such as restricting the use of force to children aged 2 to 12 years and prohibiting the use of objects/blows to the head. Recently, Canada has taken an important step toward the elimination of physical punishment. In December 2015, the Liberal Government committed to addressing all 94 calls to action outlined in the Truth and Reconciliation Report, and the 6th call to action specified the repeal of Section 43 (Mas, 2015). In the same month, Senator Céline Hervieux-Payette tabled a private member's bill (Bill S-206) to repeal Section 43. The bill passed the second reading in the House of Parliament in May 2018 and is currently under review by the Legal and Constitutional Affairs Committee (Parliament of Canada, 2019).

Physical Punishment Rates in Canada and Links to Children's Outcomes

Despite efforts aimed at the elimination of physical punishment, this form of discipline is still widely used in Canada and throughout the world. Approximately 80% of children worldwide experience some form of physical or psychological discipline in their homes (UNICEF, 2014). In Canada, representative data from the National Longitudinal Survey of Children and Youth indicate that 1 in 4 parents (approximately half a million) report having used physical punishment with children aged 2-11 years (Fréchette & Romano, 2015). Other Canadian studies have found even higher rates, ranging between 59% to 71% (Durrant & Ensom, 2012; Durrant, Rose-Krasnor, & Broberg, 2003). This variability may be a function of methodological differences among studies. For example, Fréchette and Romano (2015) examined physical punishment in a sample of 2-11-year-olds using a single 5-point Likert item, whereas Durrant et al. (2003) used the 10-item Parental Response to Misbehavior (Holden, Coleman, & Schmidt, 1995) scale involving a slightly younger sample of children aged 2-6 years. Given that the available information regarding physical punishment prevalence is based on parental self-reports, it is likely to be underreported due to recall biases and unwillingness to disclose sensitive family issues (Straus & Stewart, 1999).

In addition to its high prevalence, physical punishment can have deleterious effects on children's mental health and well-being. There is solid research showing that physical punishment exerts a negative toll on children's adjustment across multiple domains of functioning (e.g., emotional, interpersonal, cognitive) on both a short- and long-term basis. In a meta-analysis that included 88 studies, Gershoff (2002) found 10 detrimental effects of physical punishment out of the 11 that were tested. In childhood, physical punishment was related to decreased moralization ($d = -.33$), increased aggression ($d = .36$), delinquent and antisocial behaviour ($d = .42$), decreased mental health ($d = -.42$), lower relationship quality with the parent ($d = -.58$), and increased risk of physical abuse victimization ($d = .69$). In adulthood, physical punishment was linked to increased aggression ($d = .57$), criminal and

antisocial behaviour ($d = .42$), decreased mental health ($d = -.09$), and increased risk of being a perpetrator of physical abuse ($d = .13$). The only positive outcome of physical punishment was the children's immediate compliance ($d = 1.13$). One limitation of Gershoff's (2002) meta-analysis was the adoption of a broad definition of physical punishment that included potentially injurious and abusive parental behaviour (e.g., hitting children with objects) with more common forms of discipline (e.g., spanking).

In a more recent meta-analysis, Gershoff and Grogan-Kaylor (2016) selected 75 studies that exclusively reported on spanking as the primary form of corporal discipline. Results indicated that spanking was related to 13 of the 17 outcomes investigated (e.g., low moral internalization, high levels of child aggression, negative parent-child relationship). Moreover, 99% of the effect sizes (78 of the 79 statistically significant effect sizes) indicated the presence of an association between physical punishment use and child detrimental outcomes. In contrast to the results of the previous meta-analysis (Gershoff, 2002), no statistically significant relationship between physical punishment and children's immediate and long-term compliance was found which may have been a function of the technique used to analyze the data (Gershoff & Grogan-Kaylor, 2016). In her initial meta-analysis, Gershoff (2002) analyzed pre- and post- intervention compliance rates from a series of experiments conducted by Roberts and Powers (1990) that compared children exposed to spanking vs. time-outs following an act of disobedience. Following the incidents, those children were then asked to respond to 30 simple commands from their mothers. However, when analyzing the data, Gershoff (2002) failed to account for baseline differences in the rates of compliance for the comparison groups, which were substantial for at least 2 of the 5 experiments reported, when considering post-intervention compliance rates. In their most recent meta-analysis, Gershoff and Grogan Kaylor (2016) achieved clearer results because they were able to account for those baseline differences by subtracting pre-to-post change scores for the spanking group from pre-to-post change scores for the non-spanking group (i.e., time-out only) and as

such, they were able to account for the differential effect of initial levels of compliance on children's behaviour.

Parenting Programs to Reduce the Use of Physical Punishment

A number of parenting programs exist to raise awareness of the negative effects of physical punishment on children and help parents who are struggling with discipline. These programs are based on different theoretical models and modalities and can be divided into four main categories: individual- and family-level; hospital or clinic-led; group-based interventions; and multilevel programs (for a recent review of these programs, see Gershoff & Lee, 2020). Individual-level parenting programs usually involve a program facilitator working one-on-one with a parent (with or without a partner). The individualized sessions have the advantage of focusing on specific issues identified jointly by the facilitator and the parent, and therefore have a customized approach tailored to the needs of the family (Holland & Holden, 2020). Moreover, with an individualized approach, it is possible for the facilitator (depending on their professional qualifications) to work closely with the parent on issues that might directly or indirectly impact parenting behaviours (e.g., relationship with the partner, attachment style, mental health issues). Despite such advantages, there are also several limitations in that individual-level programs are time-consuming and costly. Moreover, individual-level programs lack the interactive component that is typical of group-based parenting interventions where parents share ideas with and offer support to one another (Holland & Holden, 2020).

Group-based parenting programs can be beneficial because they provide an opportunity for vicarious learning, modelling, and peer support (Gidron, 2013). Moreover, group-based interventions (particularly community-based ones) have a larger reach than individual-based programs and therefore are time- and cost-effective (e.g., Cunningham, Bremner, & Boyle, 1995). Group activities (e.g., role-play, problem-solving exercises, and group discussions) are often a core aspect of group-based programs. These aspects have been found to facilitate the acquisition of knowledge and skills and to

help parents cope with parenting challenges (Kane, Wood, & Barlow, 2007). One limitation of group-based interventions is around the management of issues related to confidentiality and disclosure during the group sessions. Moreover, some individuals might feel reluctant to share personal information that is sensitive (i.e., admitting to the use of child physical punishment) because of fear of being judged. Another important limitation includes the fact that some individuals might find difficult to share personal experiences due to their communication style and/or personality characteristics (e.g., quiet) and as a result, they might feel uncomfortable being in a group setting (Olofsson, Skoog, & Tillfors, 2016).

Hospital- or clinic-based programs can be both at the individual and group level, and they are accessible to parents either through referral or as outreach universal programs aimed at providing information to new parents and/or parents at risk. Because of the nature of the setting, these programs are delivered by specialized professionals (e.g., pediatricians, nurses, psychiatrists) in the health care setting (Dubowitz, 2020). These programs are based on a psychoeducational approach aimed at supporting parent learning and are short-term (e.g., one session). These programs have the advantage of being time-effective and inexpensive. However, their brief duration and focus on the presentation of information might not be effective in changing parenting behaviour.

Finally, multilevel parenting programs blend both individual- and group-based sessions and are organized in increasing levels of intensity and depth. The different levels are organized in a way that integrates fundamental aspects of parenting (e.g., child behaviour, attitudes, attachment) and combines a multitude of approaches and theoretical frameworks (e.g., behaviour management, cognitive training, attachment). Such programs are also designed to have levels delivered as separate components (e.g., psychoeducation about parenting and child development) or together in combination with other levels (e.g., psychoeducation and learning of specific parenting skills). Such a multilayered approach

constitutes the primary advantage of multilevel programs because it increases its efficiency and flexibility (e.g., the level given is based on the needs of the family), while also minimizing the cost of delivery (Prinz, 2020). For this thesis, group-based parenting programs were reviewed as this modality is relevant to the topic of my research on the evaluation of the group-based program called Positive Discipline in Everyday Parenting.

Among the most recent group-based programs addressing physical punishment are the following: Adults & Children Together Raising Safe Kids (ACT; Silva, 2007); Chicago Parenting Program (CPP; Gross et al., 2003); Mom's Empowerment Program (MEP; Graham-Bermann, 2012) and Positive Discipline in Everyday Parenting Program (PDEP; Durrant, 2007, 2013). The ACT is a psycho-educational program designed for parents of children between birth to 10 years of age, based on principles from social learning theory (Bandura, 1978). The program's main goal is to prevent violence in the community by teaching parents that aggression manifested in children is the product of violence experienced in the family and the community at large and that modelling healthy and caring behaviour toward others can help children develop positive behaviours (Silva, 2007). ACT involves nine group parenting sessions aimed at presenting research in child development, anger management techniques, strategies to reduce the influence of media as well as role-playing and modelling techniques to reduce aggression in children and youth. Motivational interviewing is also used to support parents' resistance to change and help them focus on their parenting goals. A randomized controlled trial (RCT) conducted with a U.S. community-based sample of 271 parents assigned to the experimental group (ACT) and wait-list control group indicated a significant decline of parent-reported harsh verbal and physical parenting from pre- to post-program and from post- to 3-month follow-up. There also was an increase in parent-reported nurturance from post-to 3-month follow-up for the experimental group, compared to the control group. It should be noted that the children's ages were not specified in this study. The study

findings provide evidence for the effectiveness of the ACT program in reducing harmful parenting and promoting positive child rearing practices (Portwood Lambert, Abrams, & Nelson, 2011).

CPP is a psycho-educational and behaviour management program specifically designed for parents of young children between two and five years. CPP's main goals are to help parents in setting age-appropriate expectations for their children, establishing consistent routines and build self-confidence, reducing the use of physical punishment by promoting positive parent-child interactions and promoting children's socio-emotional well-being by decreasing challenging behaviours. This program is organized into 11 weekly 2-hour sessions and a follow-up session, which runs three to four weeks following completion of the program. In the program, parents learn how discipline is tied to "teaching" children rather than to inflicting penalty and to consistency in implementing discipline in that it needs to occur contingent to the child's behaviour rather than depending on a parent's mood (Breitenstein, Gross, & Bettencourt, 2020). Pooled data from two clustered RCT studies (Breitenstein et al., 2012) comparing an experimental group receiving the CPP and a wait-list control group confirmed the effectiveness of the CPP program in reducing parent self-reported use of physical punishment as well as in increasing consistent discipline and parenting self-efficacy (both of which were also measured using self-reports). In these two studies, parents of young children (mean age 2.81 years) were recruited from daycare centres serving 2- to 4-year-old children in Chicago. The experimental group (relative to the control group) also reported positive changes in children's behaviour (i.e., decrease in both externalizing and internalizing) assessed by multiple informants (e.g., parents and teachers) and through observational data (Breitenstein et al., 2012). CPP is one of the few programs that can also be run entirely remotely through the use of an App "EzParent" (Breitenstein, Shane, Julion, & Gross, 2016). The ezParent program is composed of six self-administered 1-hour modules that reflect the topics presented in the original version of the program. A relatively recent implementation study comparing EzParent with group-based CPP in a sample of 79 parents (primarily

Latino and Black) of 2-5 year olds recruited at a pediatric clinic in Chicago found significant changes in parental warmth with comparable effect sizes for both groups (Breitenstein, Fogg, Ocampo, Acosta, & Gross, 2016). The EzParent program is currently undergoing rigorous testing through a large RCT involving 312 families (Breitenstein, Schoeny, Risser, & Johnson, 2016).

MEP is a brief community-based parenting support and behaviour management program for mothers of children aged 6-12 years, who have experienced intimate partner violence (IPV). IPV is associated with high levels of parenting stress (Graham-Bermann & Miller-Graff, 2015), which in turn is linked to a heightened risk of child maltreatment (Taylor, Guterman, Lee, & Rathouz, 2009). MEP involves 10 1-hour group sessions during which mothers work through their fears and discuss the impact of their experiences of family violence on parenting. Mothers are also taught alternative behaviour management techniques (e.g., constructive talk, time-out) with the intent of discouraging the use of punishment and promoting positive child-parent relationships. In a recent study, Grogan-Kaylor and colleagues (2019) examined the trajectories of physical punishment use in a sample of 113 mothers who experienced IPV and their young children aged 2-4 years. The mothers were randomly assigned to either an experimental group (MEP) or a wait-list control group. Parenting data were collected at baseline and post-program, as well as at 6- to 8-month follow-up. Results indicated that, compared to the wait-list control group, mothers in the MEP group reported a significant decline in the use of physical punishment both immediately after program completion and at follow-up (Grogan-Kaylor, Galano, Howell, Miller-Graff, & Graham-Bermann, 2019).

With respect to multilevel parenting programs aiming at reducing physical punishment, Triple P has become a widely-used approach around the world (Sanders, 2012). The Triple P program targets different age groups from infancy to adolescence and comprises different levels of specialization to address parents' specific needs. It differs from purely group-based programs in that it alternates between individual meetings with parents and group sessions, and it has a flexible format that allows

the delivery of one or more levels depending on the issues identified by the family. The Triple P levels include: Level 1 (media/communication) focuses on offering information to all parents by way of brochures, posters, billboards, and TV ads); Level 2 (brief parenting advice) involves one or several sessions (individual or group) aimed at providing support to parents who simply need to consult around minor challenging child behaviours (e.g., toilet training, bed-time routine). ; Level 3 (narrow-focus parenting support) involves several consultation sessions (individual or group) for parents of children displaying mild to moderate behavioural difficulties; Level 4 (broad-focus parenting support) is designed for parents of children with severe behavioural difficulties (e.g., oppositional defiant disorder, attention-deficit/hyperactivity disorder), and it can be delivered individually or with a blended approach (several group and individual sessions). ; and Level 5 provides intense individual support to families with complex needs (e.g., parent's mental health problems, family conflict, risk of child maltreatment; <https://www.triplep.net/glo-en/home/>).

To date, about 151 randomized trials have confirmed the effectiveness of Triple P (Sanders, 2020). In a recent meta-analysis of 101 studies examining different versions of the Triple P program, results indicated small to large effect sizes (.22 - .58) for changes in child emotional and behavioural adjustment, parental relationship, and parenting satisfaction and efficacy. Moreover, significant effects (.20-.46) were also found from post-program to follow-up on all these outcomes (Sanders, Kirby, Tellegen, & Day, 2014). However, research claims about the effectiveness of Triple P in reducing children's behavioural problems have been criticized for relying on trials that used small samples (i.e., fewer than 35 parents per group; Coyne & Kwakkenbos, 2013) and for program developers having undisclosed financial conflicts of interest (Wilson et al., 2012). Furthermore, it is still unclear whether Triple P is effective in reducing the endorsement and use of "mild" forms of physical punishment (e.g., spanking), as well as non-physical (emotional) punishment. For instance, in a small-scale evaluation of Triple P in Canada that involved 27 parents of 2-6 year olds, results indicated a significant decrease in

grabbing/shaking, but other forms of physical punishment (e.g, spanking, slapping) remained unchanged from pre- to post-program. Moreover, parents showed an increase in the use of non-physical punishment (e.g., removal of privileges), and there were no significant changes in the use of positive disciplinary strategies (Gonzalez, Ateah, Durrant, & Feldgaier, 2019).

In sum, several parenting programs exist whose focus is to reduce the use of child physical punishment. However, all the parenting programs reviewed above have a strong emphasis on behaviour management techniques that rely on the modification of children's conduct and that tend to problematize challenging child behaviour. Moreover, a behaviour management approach focuses on children's immediate compliance, with less attention to long-term parenting goals (i.e., raising children as caring and empathic adults) and to the impact of punitive measures (e.g., time-out, planned ignoring) on the quality of the child-parent relationship (Durrant et al., 2017; Holden, Grogan-Kaylor, Durrant, & Gershoff, 2017;). In this regard, the PDEP program diverges from traditional behaviour modification techniques in that it helps parents reflect on their disciplinary choices, establish age-appropriate developmental norms, and improve their efficacy in dealing with conflictual situations (Durrant, 2007, 2013).

Taking Action to Eliminate Physical Punishment: The PDEP Program

In support of emerging global initiatives to promote children's well-being, Save the Children together with Dr. Joan Durrant, a developmental psychologist at the University of Manitoba, initiated a project aimed at eliminating the use of child physical and emotional punishment. As part of this project, the Positive Discipline in Everyday Parenting (PDEP; Durrant 2007, 2013) program was developed as a primary prevention strategy to reduce child physical and non-physical punishment and promote positive discipline among parents. In general terms, PDEP is a group-based parenting program that aims to enhance parents' general knowledge of child and youth development and promote the use of

disciplinary measures that focus on non-violent conflict resolution and sensitive responding (Durrant 2007, 2013).

The PDEP program is based on several theories, namely the Social Information Processing theory (SIP; Milner, 2003) and the Theory of Planned Behaviour (TPB; Ajzen, 2012). SIP suggests that pre-existing cognitive schemata regarding the self and the child influence parenting behaviour. Particularly, parents who hold biased views of themselves by exhibiting a lack of confidence in their parenting role (i.e., low self-efficacy) and poor understanding of typical child behaviour are more likely to be physically abusive (Milner, 2003). TPB is a social cognitive theory of persuasion and behaviour change. According to this theory, individuals are motivated to behave in a particular way if they: 1) hold positive *attitudes* toward that behaviour; 2) feel *social pressure* to perform the behaviour; and 3) feel *competent* to perform the behaviour. PDEP is designed to target these three cognitive components to reduce physical and emotional punishment and increase positive disciplinary practices. PDEP targets each of the TPB components by 1) increasing parents' knowledge of children's emotional, behavioural, and brain development from birth to adolescence; 2) engaging parents in group-based problem-solving activities; and 3) providing skill-building practice activities. The objective is to gradually shift parents' responses to children's behaviour from anger and punishment to communication and problem-solving by increasing their knowledge, empathy, insight, and emotion regulation (Durrant, 2007; 2013).

Both SIP and TBP theories explain the conceptual model of the PDEP program, which is presented in Figure 1. This model is based on two main components of the program, namely the cognitive/affective processes and action goals. The cognitive and affective processes component focuses on increasing parents' knowledge of child development and understanding the reasons underlying children's behaviour. It also aims to increase parents' self-awareness (particularly concerning their emotional response) and reflectivity. Cognitive and affective processes components aim to help parents shift attitudes about physical and emotional punishment and its use. On the other

side, the action goals focus on helping parents improve their self-regulation skills and modulate their emotional response through intellectual processes, thereby helping them implement practice positive disciplinary practices. The action goal component aims to help parents improve their problem-solving abilities and emotion regulation, as well as reduce parenting stress.

PDEP is based on eight weekly group sessions that target the three following components: 1) reduce approval of physical and emotional punishment (shift attitudes); 2) modify negative attributions of typical child behaviours due to false beliefs with respect the child's developmental norms (shift subjective norms); and 3) strengthen parenting self-efficacy (increase perceived sense of competence in the use of positive disciplinary practices). Parents engage in small group activities and exercises aimed at normalizing parent-child conflict as well as redefining the concept of child "misconduct" and "developmentally appropriate behaviour." The program activities are sequenced to gradually increase parents' ability to take their child's perspective, thereby reducing frustration and anger (Durrant, 2007, 2013). Parents also learn about the brain's stress response and how it can escalate the conflict to result in punishment, coercion, and violence. Parents work on generating methods of regulating their own emotions, which they practice by way of guided exercises. Throughout the program, parents gradually learn a problem-solving framework that they can apply in a wide range of parent-child conflict situations (Durrant, 2007, 2013).

As an initial step, parents consider their long-term parenting goals as a way to refocus their attention on understanding the impact of their expectations regarding their child's future development on parenting behaviour. Second, parents learn about the importance of providing warmth (physical and emotional safety) and structure (information and guidance) for their child as a way to prevent and also resolve conflicts constructively. Third, parents learn in detail about the various stages of development from birth to adolescence and explore critical aspects of development that characterize each stage. Finally, as part of problem-solving activities, parents practice applying this framework to a range of

hypothetical scenarios and to situations from their own experience. The expected outcomes of this process are a reduction in parents' approval of physical and other punishments, better understanding of children's feelings and behaviour, greater sense of competence in their parenting role, a greater likelihood of problem-solving rather than punishment in response to conflict, and a lower level of parenting-related stress.

PDEP addresses several important parenting issues, including attention to children's needs for security, warmth, and structure. Further, it targets both emotional and cognitive processes that contribute to the use of physical and emotional punishment, including parental approval of harsh discipline, negative parental attributions of children's behaviour, and parenting self-efficacy. From this perspective, PDEP is an innovative alternative to other parenting programs as it adopts a multifaceted approach aimed at changing parental attitudes and beliefs regarding the use of negative child disciplinary practices (Durrant 2007, 2013).

Moreover, PDEP is different from the other programs mentioned above in that it does not focus on behaviour management, which relies on the use of consequences to modify children's behaviour. In contrast, PDEP is based on children's rights principles – the right to protection, the right to dignity, and the right to be heard. The objective is to shift parents' focus from punishment and control to problem-solving and mentorship. In contrast, PDEP does not prescribe rewards or punishments but places emphasis on helping parents understand the social, emotional, and cognitive aspects of children's behaviour so that they acquire a more empathic approach toward their children.

Furthermore, PDEP's approach to problem-solving is based strongly on strengthening trust and attachment and helping parents understand the process of development and self-regulation. Parents learn about the brain's response to stress and how they can help themselves and their children learn how to manage stress constructively. Most importantly, PDEP does not prescribe time-out or any other form of punishment (i.e., planned ignoring, withdrawal of privileges) but rather, it focuses on

strengthening parents' understanding of developmental processes and of their child's rights to be free from all forms of violence. Thus, PDEP shifts away from behaviour management strategies and instead uses inductive techniques to help parents reflect on their long-term parenting goals the reasons behind their child's challenging behaviours, and problem-solving skills (Durrant et al., 2014).

As part of PDEP's evaluation, Durrant et al. (2014) designed a questionnaire composed of four scales targeting the main aspects of parenting addressed through the program. To date, this is the only measurement instrument that has been used to quantitatively assess the program's outcomes. The scales are as follows: 1) *Approval of Physical Punishment* measures the degree to which parents are in favour of using physical punishment; 2) *Approval of Non-Physical Punishment* measures the degree to which parents are in favour of using emotional punishment (i.e., time-out, grounding, shaming); 3) *Subjective Norms* measures parental attributions of children's behaviours (with particular attention to developmentally-appropriate behaviours); and 4) *Parenting Self-Efficacy* measures parents' sense of self-efficacy and competence in their parenting role, as well as their confidence in constructively resolving conflict with children. Durrant et al. (2014) implemented PDEP across 14 Canadian cities and collected data on 321 parents who completed the program. Pre- and post-program data indicated that, following the 8-week program, parents were significantly less supportive of punitive parenting strategies, less likely to attribute typical child behaviour to misbehaviour/non-compliance, and more likely to believe in their parenting role (i.e., higher parenting self-efficacy).

Although these results are encouraging, they need to be viewed as preliminary for several reasons. First, the psychometric characteristics of the PDEP questionnaire are currently unknown. Moreover, Durrant et al. (2014) reported that the reliability of the Parenting Self-Efficacy scale is in the low range ($\alpha = .49-.52$). Therefore, it would seem that a detailed psychometric analysis is required to increase confidence in the primary questionnaire used to evaluate the PDEP program. In this dissertation, my first two studies involved a series of item-level analyses to determine the psychometric

properties of the PDEP questionnaire and then to apply all the necessary modifications to achieve a more reliable and valid measurement instrument. An additional limitation of the PDEP program is that its evaluation has relied on pre- to post-program data without the inclusion of a comparison group (Durrant et al., 2014). This hampers definitive conclusions about the efficacy of the PDEP program. Thus, my dissertation included a third study that made use of a quasi-experimental design with a wait-list control group to examine program outcomes.

Dissertation Objectives

As mentioned, my dissertation consisted of three separate studies. Study 1 evaluated the psychometric qualities of the scales included in the current version of the PDEP questionnaire by way of item-level analyses conducted using the Item Response Theory (IRT) method. Pending the results, items were modified, and a refined version of the questionnaire was administered to two expert samples (e.g., one mainly composed of researchers and professionals and the other composed of parents from the community) to assess content validity. Following the content validity analyses, the final revised version of the PDEP questionnaire was created. Study 2 was conducted to assess the factor structure of the scales of the new revised version of the PDEP questionnaire via IRT confirmatory analyses by using data from a sample of 400 parents recruited through an online platform. Moreover, reliability of the PDEP questionnaire was assessed in terms of both internal consistency (i.e., Cronbach's alpha) and test-retest reliability. For the latter, one-fourth of the total sample ($n = 100$) completed the same questionnaire four weeks after the initial administration. Data were also used to assess the convergent validity of the final revised version of the questionnaire. For this purpose, the various scales of the PDEP questionnaire were correlated with similar constructs to determine the degree of association among the selected measures. Finally, Study 3 examined the efficacy of the PDEP program in decreasing parents' approval of punishment and subjective norms (e.g., negative attributions of typical child behaviours) and in improving parenting self-efficacy, as well as other constructs of interest

including parental emotion regulation, parenting stress, problem-solving, and parenting practices. This evaluation was carried out by way of a quasi-experimental design that included a wait-list control group.

These three studies have important implications for improving the accuracy of the measures used to evaluate PDEP outcomes. The program evaluation served to provide more solid information about PDEP's impact in reducing the approval of child physical and non-physical punishment, as well as in increasing parents' ability to manage parenting stress, regulate their emotions, problem-solve, and engage in positive parenting practices. At a broader level, PDEP's goals are well-aligned with international efforts to reduce any form of violence against children and with growing awareness of the negative developmental impacts of physical punishment, alongside growing recognition of children's human rights through the United Nations Convention on the Rights of the Child. At a national level, evaluation of the PDEP program is well-aligned with current federal government initiatives. Specifically, the Truth and Reconciliation report called for a repeal of Canada's Section 43 law which allows for the use of physical violence against children, and the federal government has committed to addressing all calls to action in the report.

Study 1: Evaluating and Modifying the Current Positive Discipline in Everyday Parenting (PDEP) Questionnaire**Abstract**

The purpose of this first study was to gather information on the psychometric properties of the questionnaire currently being used to assess PDEP outcomes and assess its content validity. First, the responses of 1,047 Canadian parents who participated in the PDEP program between 2013 and 2015 were analyzed to determine the general psychometric qualities of the current questionnaire. Item Response Theory (IRT) analyses were performed to detect the items that demonstrated good measurement qualities (i.e., high discrimination) as well as redundant and poor fitting items. In the development of the revised version of the PDEP questionnaire, additional items were added to increase the coverage of scales included in the tool. Subsequently, the content validity was examined to assess the extent to which the questionnaire scales were in alignment with the theoretical principles underlying the PDEP program, and additional modifications were made following this analysis. For this purpose, two expert samples ($n = 8$ professionals in the field of child psychology and $n = 10$ parents) were recruited to assess the extent to which the revised scales were in alignment with the theoretical principles underlying the PDEP program. Overall, items within each scale obtained favourable ratings from the two expert samples, who provided several suggestions on how to modify the items to better align them with the constructs of interest. As a result, several items were removed, and additional modifications were made. In the end, a final version of the PDEP questionnaire was created.

Introduction

One critical component of a program's evaluation is the use of measurement instruments that adequately assess the program's goals and reliably capture important program aspects (Lichiello & Turnock, 1999). This first section describes the questionnaire that is currently in use to evaluate PDEP outcomes. The questionnaire was examined in its content as well as its alignment with the PDEP's underlying theoretical framework. In the second section, a brief overview of the measurement theory is presented along with a summary of the analytic strategies that were used in this study to assess the psychometrics of the PDEP questionnaire.

The PDEP Program: Objectives and Theoretical Framework

The PDEP program (Durrant, 2007, 2016) was developed with the complementary goals of eliminating the use of physical and emotional punishment of children while increasing the use of positive disciplinary strategies. It is a universal group program for parents of children newborn to 18 years that focuses on helping parents shift away from rewards and punishments toward problem-solving and communication (Durrant et al., 2014). The three objectives of PDEP are to: 1) reduce the approval of physical and non-physical punishment (shift attitudes); 2) modify negative attributions of typical child behaviours due to false beliefs regarding the child developmental norms (shift subjective norms); and 3) strengthen parenting self-efficacy (increase perceived behavioural control). With respect to the first objective, PDEP is designed to change parental attitudes toward any form of punishment (i.e., physical and emotional) by providing information on its long-term developmental risks and on the developmental benefits of trust, attachment, and communication within the parent-child relationship. For the second objective, the program aims to shift parents' subjective norms, allowing them to reflect on their perception of children's behaviour and encouraging them to rely on positive disciplinary techniques. For the third objective, PDEP focuses on strengthening parents' confidence in generating and implementing non-punitive solutions to conflict with children (Durrant, 2007; 2013).

These objectives are in line with the theory of planned behaviour (TPB; Ajzen, 2012) which provides an integrative conceptual framework to motivate behavioural change and promote positive actions. According to TPB, intentions are assumed to directly precede the performance of a behaviour. Moreover, one's intention to engage in a certain behaviour is affected by their behavioural beliefs, normative beliefs, and perceived behavioural control. *Behavioural beliefs* refer to the perceived consequences of a behaviour, which leads to the development of favourable or unfavourable attitudes toward a behaviour. *Normative beliefs* refer to social behavioural expectations and the perceived social pressure to conform to such expectations. Normative beliefs aid in the prediction of subjective norms, which are beliefs about whether people of importance to a person think (i.e., children) they should engage in the behaviour. Finally, *control beliefs* represent the perceived presence of impediments or facilitators that influence the performance of a behaviour. This set of accessible control beliefs, in turn, determines perceived behavioural control, which refers to a person's perceptions of his/her ability to perform a given behaviour (Ajzen, 2012). According to the TPB theory, change in the use of physical and emotional punishment would require the transformation of beliefs in three areas: 1) attitudes (the new behaviour of *positive discipline* is viewed more favourably than the use of punitive disciplinary strategies); 2) subjective norms (the individual feels increased social pressure to adopt new disciplinary strategies); and 3) perceived competence (the individual feels able to replace the old behaviour with the new behaviour; Durrant et al., 2014).

PDEP is also rooted in the Social Information Processing theory (SIP; Milner, 2003), which postulates four stages of parent-child conflict which serve to understand the dynamics of physical and emotional punishment use. In Stage 1, parents generally attend to the situation and interpret the events based on empirical evidence. In the opposite case, parents who struggle with discipline might misinterpret their child's emotional states and behaviour. In Stage 2, parents generate attributions for the child's behaviour. Parents who use punitive disciplinary strategies tend to attribute typical child

behaviour (e.g. tantrums) to intentional defiance, rather than attributing the same behaviour to developmental reasons (e.g., lack of understanding). In Stage 3, parents select a response from those available in their repertoire. Punitive and violent parents tend to have fewer options in their repertoires and more difficulty generating responses that are flexible or creative. In the final stage, parents enact the chosen response and monitors its effects. Parents' ability to monitor their actions is influenced by their affective states; a high level of negative arousal will interfere with their ability to attend to the child's emotional response and to inhibit their own escalating aggression.

Both theoretical principles of TBP and SIP are incorporated in the PDEP conceptual model (please refer to Figure 1), which is centred around two main components that together explain the processes through which parenting behaviours might become established, namely *cognitive affective processes* and *action goals*. Cognitive and affective processes, such as the enhancement of both parental knowledge of child development and parents' level of accuracy in the interpretation of children's behaviour, might help explain changes in attitudes toward and use of punitive disciplinary strategies. In other words, parents who have adequate knowledge of child development and understand the causes of a child's challenging behaviour are less likely to hold favourable attitudes toward the use of punishment and more likely to employ positive parenting strategies to handle conflictual situations with their children. Similarly, shifts toward insightfulness, better perspective-taking, and increased reflectivity are expected to influence parents' decisions concerning the use of positive parenting strategies. The second component of the model is action goals, which refers to regulatory processes that have an impact on parenting. Particularly, parents' increased ability to regulate their emotions in response to children's challenging behaviour, along with their perceived confidence in using positive disciplinary strategies, might help normalize parenting experiences by reducing stress and encouraging parents to become independent problem-solvers.

The PDEP Questionnaire

Durrant et al. (2014) developed the PDEP questionnaire to measure common outcomes expected to result from participation in the PDEP program. These scales, namely Approval of Physical Punishment, Approval of Non-Physical Punishment, Subjective Norms, and Parenting Self-Efficacy, were designed in alignment with the three objectives of PDEP: 1) reducing parental approval of physical and non-physical punishment; 2) decreasing the use of negative attributions of typical child behaviours (i.e., shifting subjective norms); and 3) strengthening parenting self-efficacy with respect to the use of positive disciplinary strategies. Moreover, these four scales map onto both the TPB framework and PDEP conceptual model (particularly with respect to cognitive and affective processes), and, hence, are ultimately expected to decrease parental approval of punitive parenting strategies and increase positive discipline by shifting subjective norms and enhancing parenting self-efficacy. However, several issues concerning the PDEP questionnaire need to be addressed. First, the psychometric and structural properties of the questionnaire have not yet been thoroughly examined. Second, the Parenting Self-Efficacy scale has been found to have low internal consistency, thereby raising concerns about its reliability (Durrant et al., 2014). This study addressed these issues by conducting an item-level analysis of the current version of the PDEP questionnaire to assess the various scales' psychometric characteristics and determine if any changes may be needed. Moreover, a content-based analysis using two expert samples was undertaken to examine the alignment of each measurement scale with the construct of interest, and a revised version of the questionnaire was subsequently produced.

Fundamentals of Measurement Theory: Reliability and Validity

Measurement consists of a systematic set of rules used to define the scaling properties of objects, individuals, groups, and events, thereby allowing for their classification into similar or different categories based on quantifiable characteristics (Nunnally, 1978). The measurement of

psychological phenomena is often complicated by the intangible nature of their constituent elements (i.e., percepts, mental states, attitudes). However, the existence of psychological events can be inferred by measuring observable properties, such as items on self-reported questionnaires. These are referred to as manifest indicators of the latent variable and are used as a proxy measure of the underlying psychological construct. Manifest indicators do not constitute a perfect measure of latent constructs but contain random measurement error. Such error represents a measure of the discrepancy between a true latent score and observed scores on the indicator variables (Byrne, 2013). In other words, it is the random variability in the underlying construct that remains unexplained after accounting for the relationship between the manifest indicators and the latent variable. Test taker's psychological states and contextual characteristics all constitute sources of measurement error that affect the accuracy and precision of a measurement instrument (John et al., 2007).

In psychometric assessment, the accuracy of a test is determined by the degree to which it is a reliable and valid measure of the latent construct. Reliability is generally defined as the ability of a measurement instrument to achieve “consistency of scores across instances of the testing procedures” (American Educational Research Association, 2014, p.33). Reliability is often assessed in terms of internal consistency, which is calculated through a Cronbach's alpha coefficient (Cronbach, 1951). Reliability is also assessed in terms of the consistency of an instrument over time (i.e., test-retest reliability), which is the correlation between scores on the same measure on two different occasions (Carmines & Zeller, 1979). Both internal consistency and test-retest reliability are used in Classical Test Theory (CTT; Novick, 1966). In modern test theories, such as Item Response Theory (IRT; Lord, 1980), the concept of reliability extends to measures of both test and item information that are used as proxies for the true reliability (details of these two methods are presented in the next section).

Validity is generally defined as “the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (American Educational Research Association,

2014, p.11). Thus, it is a somewhat more complex concept that involves a series of inferential procedures to collect evidence in support of the adequacy and soundness of a measurement instrument. There are different ways of assessing the validity of a measurement instrument. In this study, evidence about validity was gathered based on the content of the measurement instrument (i.e., content validity). The feedback of content experts on the measure of interest is often used to assess the correspondence between the content of the measurement instrument and the constructs being measured (Haynes, Richard, & Kubany, 1995).

Statistical Methods for Evaluating Measurement Scales: Classical Test Theory and Item Response Theory

Classical Test Theory (CTT) includes factor analytic techniques that are widely used for the psychometric assessment of measurement scales. CTT assumes that an observed score (X) is the sum of the true score (T) plus random error (E), where the person's true score is defined as the average score obtained after having repeatedly administered parallel tests (i.e., measures for which the error variance is equal) to the same sample over an infinite number of times. Hence, CTT is a linear model based on relatively simple mathematical procedures of parameter estimation, where the total scale score is predicted by the additive combination of independent variables (i.e., factor indicators; Lord & Novick, 1968). Moreover, CTT models are less conceptually demanding and have less restrictive assumptions compared to other modern methods (Hambleton & Jones, 1993).

Despite these advantages, CTT models come with several major drawbacks. In traditional CTT, items are considered to contribute equally to the scale total score (DeVellis, 2003). This is an inaccurate assumption because each item has its own characteristics (e.g., item discrimination and difficulty) that contribute to its performance at different levels of the latent trait score. In addition, CTT estimates are sample-specific in the sense that they are based on the comparison of the total scale score to a normative sample. Similarly, CTT estimates rely on the summed scores of a specific group of

respondents, from whom data are collected, hence limiting the generalizability of findings to a different sample (Embretson & Reise, 2000). Another limitation is that the true total score is test-dependent, so it is constrained to the specific set of items on a scale. In other words, adding or subtracting an item from a measurement scale will change the true score scale, thereby resulting in a different psychometric scale. Inherent to the concept of true score is reliability assessment. CTT's reliability index, commonly referred to as Cronbach's alpha coefficient is the mean of all split-half reliabilities (Cronbach, 1951). Despite the wide use of this reliability index, Cronbach's alpha does not give any information on how individuals at different levels of the latent trait respond to the scale items. Moreover, Cronbach's alpha is not a stable characteristic as it may be affected by the removal or addition of items (Reise & Henson, 2003). For instance, the magnitude of the coefficient alpha is dependent on the number of items in a scale. Generally, reliability will increase if the items that are added to the scale are homogenous, although these same items may be providing redundant information and, hence, would not contribute to the coverage of the construct that the scale is intended to measure. Thus, a measurement tool might achieve good internal consistency (i.e., high coefficient alpha) despite content saturation caused by too many items and by their intrinsic overlap (John et al., 2007).

Item response theory (IRT) is a viable alternative to CTT that can be used to help improve test construction and evaluate the quality of individual test items. While CTT works at the level of the true score, IRT measures latent traits. IRT models also specify a nonlinear relationship between item responses and the latent trait; this can be represented by a probabilistic function that models the interaction between these two parameters (Lord, 1980). The basic formula for IRT modeling is the following: Odds of Item Success = Item Ability $\times$ Item Difficulty (Wright, 1967). IRT assumes that the probability of endorsing a specific item (or item response) can be explained by an underlying latent trait (also called theta and denoted by θ). The item response is a function of both the item difficulty (or item location, denoted by b) and an individual's ability level (or item discrimination, denoted by a). In

IRT, the probability of endorsing an item, given a certain level of a latent trait, is expressed by the item response function (IRF). Item difficulty is defined as the location on the latent variable continuum where a respondent, who is chosen at random from a homogenous subpopulation, has a 50% probability of endorsing an item (Lord, 1980). Negative difficulty values are associated with “easy” items, which are items more frequently endorsed by the respondents with a low score on the latent trait. Items that are more difficult are associated with positive difficulty values. Only respondents with high scores on the underlying latent trait endorse difficult items.

The IRT model assumes that a respondent has a true location along a continuous latent trait dimension. Because both respondents and items are located on the same continuum, it is possible to make comparisons on how a typical person might answer an item. For instance, respondents who are more likely to hold positive attitudes toward physical punishment would have a higher probability of endorsing an item reflecting greater approval of physical punishment (frequently referred to as a more difficult item in IRT parlance). Item discrimination is related to the slope of the IRF, which reflects how well the item discriminates among individuals located at different points on the latent trait continuum. The steeper the IRF curve, the larger the item discrimination parameter and the better the item discriminates individuals with high levels of the latent trait from individuals with lower levels of the latent trait. Item discrimination is also related to the probability of response; the proportion of respondents endorsing a category response increases in a less gradual fashion as the IRF slope becomes steeper. Optimally, scale items should be able to discriminate between different levels of the latent trait. In other words, a scale should be composed of items with high discrimination values associated with both high and low levels of the latent trait (De Ayala, 2013).

One major advantage of IRT is that the estimates of the item responses are population-general or group invariant and, hence, are generalizable to any individual with similar latent trait levels (Lord, 1980). IRT approaches also provide a reliability measure for each item at different levels of the

underlying trait so that each item is associated with information parameters. Item information functions provide data regarding the measurement precision that an item provides at all levels of the latent trait. Moreover, item information is inversely related to the standard error of the estimate. Thus, as item information increases, there is a smaller degree of uncertainty about the person's location (De Ayala, 2013). In this sense, item information can be interpreted as a measure of the reliability of an item at different levels of the latent trait. In terms of limitations, IRT models are computationally more demanding and more difficult to interpret than factor analysis models. Moreover, IRT models usually require large sample sizes (i.e., 500 or more) to provide reliable estimates (Hulin, Drasgow, & Parsons, 1983). In sum, CTT and IRT methods are approaches used in the psychometric assessment of measurement instruments. Each model comes with advantages and disadvantages. In this study, IRT methods were used to assess the psychometric characteristics and structure of the PDEP questionnaire.

Study Objectives

The objectives of Study 1 were to examine the psychometric properties of the PDEP questionnaire and revise it, accordingly, based on IRT analyses and content validity analyses. The first objective of this study was to assess the psychometric characteristics of the current version of the PDEP questionnaire. This was addressed by conducting an item-level analysis of the PDEP questionnaire using existing data from parents who participated in the PDEP program from 2013 to 2015. For this purpose, a series of IRT analyses were carried out to examine the psychometric properties of the four scales of the PDEP questionnaire, namely Approval of Physical Punishment, Approval of Non-Physical Punishment, Subjective Norms, and Parenting Self-Efficacy. Based on the results of these analyses, refinements to the PDEP measure were deemed necessary (see results below). The second objective was to conduct a content-based analysis on the validity of the revised PDEP questionnaire and specifically, to evaluate whether the PDEP scale items were consistent with the constructs that they are intended to measure. For this purpose, two expert panels (e.g., professionals and researchers in the area

of child development as well as parents) were consulted and asked to rate the degree to which each of the items aligned the definition of the construct of interest.

Methods

Sample and Procedures

For the IRT component, secondary data available from 1,047 parents who participated in the PDEP program between 2013 and 2015 across various provinces of Canada was used. Dr. Durrant from the University of Manitoba was the principal investigator on this project, and permission to use the dataset was first sought from her. Upon her approval, an ethics application for the use of secondary data was submitted. Ethics approval was obtained from the Office of Research Ethics and Integrity of the University of Ottawa before the commencement of this study (REB # H11-15-09, approved November 30, 2015; Appendix A).

The sample was initially composed of 1,134 parents who attended the PDEP program between 2013 and 2015 across various provinces in Canada where the program was offered. These parents completed the PDEP questionnaire before and after the program. However, for the purpose of the current study, only pre-PDEP data were used. Among the 1,134 parents, 55 (4.8%) were removed from analyses because they were missing > 50% of the data on the PDEP questionnaire. An additional 3% ($n = 32$) were removed because they did not have any children. Thus, the final sample was comprised of 1,047 parents. For this final sample, the parents were predominantly female (84%). In terms of the province of residence, 41.2% of the participants were from Ontario, 38.4% from Manitoba, 9.5% from Alberta, 7.5% from Saskatchewan, and 3.4% from New Brunswick. Most parents (47.4%) ranged in age from 31-40 years; 28.9% were between 21-30 years; 20.4% were over 40 years, and 3.3% were under 20 years of age. Regarding the number of children, 37.6% of the parents had two children, 31.4% had one child, 18.1% had three children, and 12.9% had four or more children. Concerning education, 35.8% of the parents graduated from college or university, 16.2% obtained a high school degree, 14.3%

took college or university courses, 20% indicated not having completed high school, 7.3% completed a post-graduate degree, and 6.3% took post-graduate university courses.

Following the IRT analyses, several changes to the PDEP questionnaire were made in consultation with my supervisor and two individuals involved with the PDEP program (Dr. Durrant and Dr. Stewart-Tufescu from the University of Manitoba). The next step in this first study was to then gather information about the content validity of this revised version and make any additional modifications if deemed necessary. For the content validity of the revised PDEP questionnaire, two expert samples were recruited, namely professionals (e.g., psychologists, researchers, professors) and parents. Ethics approval was obtained from the Office of Research Ethics and Integrity of the University of Ottawa before the commencement of this study (REB # H-02-18-352, approved March 3, 2018; Appendix B).

A list of professionals from Canada and the U.S. with expertise in child development, parenting research, program evaluation, and psychometric assessment was developed in consultation with the thesis supervisor. This list initially included 12 individuals who were contacted by email to participate as content experts for the revised PDEP questionnaire. Eight professionals responded to the initial email and agreed to participate by completing the consent form (Appendix C). The sample was predominately female ($n = 7$; 87.5%), with a mean age of 44.9 years ($SD = 12.2$). The majority described their ethnic/racial background as White ($n = 7$; 87.5%), with the remaining participant self-identified as Latin-American (12.5%). Participants indicated that they had a high ($n = 5$; 62.5%) to very high ($n = 3$; 37.5%) level of knowledge about child development. Participants indicated that they had medium ($n = 1$; 12.5%), high ($n = 3$; 37.5%), or very high ($n = 4$; 50%) levels of knowledge about parenting research. Finally, participants indicated that they had medium ($n = 4$; 50%), high ($n = 3$; 37.5%), or very high ($n = 1$; 12.5%) levels of knowledge about measurement development.

The second expert sample was composed of 10 parents who participated in a focus group. Parents were recruited through two community centres in Ottawa by way of an employee who provided parents with a study notice. Study notices explained the purpose of the study and outlined inclusion criteria (e.g., parent of at least one child under age 18 years, ability to read and speak English; see Appendix E). Interested parents contacted the employee at the community centres who in turn contacted the researcher (i.e., the writer of this thesis) to decide on some of the potential dates for the focus groups. Once the date was set, interested parents were invited to join the focus groups. Two focus groups ($n = 5$ each) were held, one at each of the two participating centres. This sample size of 10 is consistent with similar research employing focus groups as part of a content validity assessment (Barnes & Adamson-Macedo, 2007; Hsu, Chan, & Lubomski, 2017). Group sizes also aligned with what is commonly cited in the literature, namely a minimum of four and a maximum of 12 participants per group (Kitzinger, 1995).

The parent focus groups lasted about two hours. In terms of structure, an overview of the study was initially presented and then parents completed a consent form (see Appendix F). After completing the socio-demographic questionnaire and content validity assessment forms, the parents took part in the focus group. Appendix G presents a list of questions used to guide the focus group. The focus group was not audiotaped in order to increase the comfort level among parents; however, field notes were taken as parents answered the questions.

The focus group sample was composed exclusively of mothers, whose average age was 35.2 ($SD = 4.6$) years. The majority of participants (75%; $n = 7$) were married, 16.7% ($n = 2$) were living with a partner, and 8.3% ($n = 1$) was single. One third of the participants indicated that their ethnic/racial background was Arab/Middle Eastern (33.3%; $n = 3$), followed by White (16.7%; $N = 2$), Latin American (16.7%; $n = 2$), South Asian (16.7%; $n = 2$), West Asian (8.3%; $n = 1$), and Black (8.3%; $n = 1$). The majority of mothers completed post-secondary education, with approximately

33.3% ($n = 3$) having completed a college degree, 33.3% ($n = 3$) having completed a bachelor's degree, and 25.1% a master's degree. One mother (8.3%) obtained a high school degree. In terms of partner education ($n = 9$), 22.2% ($n = 2$) obtained a high school degree, 11.1% ($n = 1$) a college degree, 33.4% ($n = 3$) a bachelor's degree, 22.2% ($n = 2$) a master's degree, and 11.1% ($n = 1$) a doctoral degree. With respect to employment status, 66.7 % of participants were not employed outside the home while 33.3% were employed. All partners were reported to be employed outside the home. Household income ranged from \$20,000-\$49,999 (40%; $n = 4$), \$50,000-\$89,999 (20%; $n = 2$), and \$90,000-109,999 (40%; $n = 4$). The majority spoke a language other than English at home (83.4%; $n = 9$) including Arabic (41.8%; $n = 4$), a language from South Asia (i.e., Telegu, Tami; 16.7%, $n = 2$), Farsi (8.3%; $n = 1$), Somali (8.3%; $n = 1$), and Spanish (8.3%; $n = 1$). Participants indicated that they had very high (25%), high (50%), and medium (25%) English reading comprehension skills. Participants indicated that they had very high (8.3%), high (41.7%), and medium (50%) English writing skills. Only one participant indicated having prior knowledge of the PDEP program; however, she reported that she had not attended the program.

Measures

Sociodemographic questionnaire. For the content validity component, the sample of experts answered several socio-demographic questions (i.e., age, gender, ethnic/racial background), as well as questions concerning their knowledge about child development, parenting research, and measurement development (see Appendix H). For the sample of parents, participants were asked several socio-demographic questions (i.e., age, gender, ethnic/racial background), as well as questions about their knowledge of the English language (e.g., reading comprehension and writing skills; see Appendix I).

The PDEP questionnaire (Durrant et al., 2014). The original version consists of four scales, namely Approval of Physical Punishment, Approval of Non-Physical Punishment, Subjective Norms, and Parenting Self-Efficacy. Approval of Physical Punishment is a 5-item scale designed to measure

parental attitudes toward physical punishment. Each item (e.g., Spanking is fine as long as the parent is not angry) is rated along a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*), with higher scores indicating greater approval of physical punishment. This scale was found to have acceptable internal consistency in the current sample of 1,1047 parents ($\alpha = .79$). Approval of Non-Physical Punishment is a 4-item scale designed to measure parental attitudes toward punishment that is not physical in nature (i.e., emotional punishment). Each item (e.g., If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay inside until he changes it) is rated along a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*), with higher scores indicating greater approval of non-physical physical punishment. This scale had poor internal consistency in the current sample of 1,1047 parents ($\alpha = .50$). Subjective Norms is a 5-item scale used to assess parents' negative attributions of typical child behaviours due to their false beliefs regarding the child's developmental norms. Participants are asked to indicate their agreement with statements attributing typical parent-child conflict to 'misbehaviour/non-compliance' on the part of the child (e.g., Young children who say "No!" to their parents are being defiant). Parents respond to each item on a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*), with higher scores indicating a stronger belief that the typical child's behaviour is attributable to 'misbehaviour/non-compliance'. This scale had questionable internal consistency in the current sample 1,1047 parents ($\alpha = .66$). Parenting Self-Efficacy is a 4-item scale designed to measure the degree to which parents feel competent and confident in their parenting abilities. Each item (e.g., I have the skills to be a good parent) is rated along a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*), with higher scores indicating greater parenting self-efficacy. This scale had poor internal consistency in the current sample 1,1047 parents ($\alpha = .58$).

Content validity form. I created the form for professionals and for parents following the recommendations of Haynes, Richard, and Kubany (1995). Each item was rated on two main dimensions, namely representativeness (i.e., how reflective the item is of the underlying construct)

using a 3-point scale from 1 (*slightly representative*) to 3 (*highly representative*) and clarity (i.e., how clear the item is) using a 3-point scale ranging from 1 (*low clarity*) to 3 (*high clarity*). The two content validity forms are presented in Appendix J.

Data Analyses

This first description pertains to the analyses conducted for the IRT examination of the original PDEP questionnaire.

Data screening procedures. Prior to analyses, PDEP questionnaire items were examined for the accuracy of data entry, missing values, and fit between their distributions and the assumptions of multivariate analysis. All screening procedures were conducted in SPSS 25.

Procedures for dealing with missing values. A Missing Value Analysis (MVA) determined the percentage of missing values for each of the questionnaire items and the pattern of missingness. The Little's test was used to determine if data were missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR; Little & Rubin, 1987). Determining the pattern of missingness is important when deciding which procedure to use to address the issue. With MAR data, the propensity of missing values depends on the observed data but not the missing data. In the case of MCAR data, there is no systematic relationship between the missingness of the data and any values, whether observed or missing (Schafer & Graham, 2002). Finally, in MNAR, missing data are dependent on the missing values themselves (i.e., the unobserved data). MNAR data are not 'non-ignorable' because the missing data mechanism itself should be modelled when dealing with missing values. MCAR and MAR are both considered 'ignorable', so it is not necessary to include any information about the missing data itself when handling the missing data (Graham, 2009).

In this study, listwise deletion of cases was used to deal with the problem of missing data for cases that had missing observations on more than 50% of the items within a particular scale. Listwise deletion does not introduce bias if the data are MCAR (Graham, 2009; Tabachnick & Fidell, 2007).

Results from the Little MCAR test were not significant ($\chi^2 = 6.32, p = > .10$), indicating that the data were missing completely at random. Of the initial sample of 1,139 parents, 134 were deleted because they had > 50% missing data on the PDEP questionnaire. The final sample of 1,047 parents had 5-8% missing data on items in the PDEP questionnaire (672 parents had complete data), that were imputed using the Expectation-Maximization (EM) method. In the EM algorithm, the two Expectation (E) and Maximization (M) steps are iterated through multiple times using maximum likelihood (ML). In the E-step, the distribution for the missing data is calculated based on the known values for the observed data and estimates of the sums of squares and cross-products (SSCP) are generated. In the M-step, the unknown parameters are estimated using all the other variables as predictors in a regression model, and the SSCP matrix obtained in the E-step is then used to estimate a new covariance matrix along with the associated regression coefficients. This two-step procedure is repeated for several iterations until the difference between the estimated covariance matrices from step-to-step falls below some specified convergence criterion (Little & Rubin, 1987). In SPSS, the convergence criterion is set as .0001. The EM method is known to produce unbiased estimates when compared to other imputation methods such as mean substitution (Rubin, Witkiewitz, Andre, & Reilly, 2007).

Normality. Screening variables for normality is an important step of any multivariate analysis. The normality assumption was assessed before running any analyses on the data through skewness and kurtosis values. For large samples, either an absolute skew value larger than 2 or an absolute kurtosis larger than 7 may be used as reference values for determining substantial non-normality (West, Finch, & Carrant, 1995). Graphical methods, such as histograms and normal probability plots, were also used in the current study (Tabachnik & Fidell, 2007). Univariate normality results supported the normality assumption for the majority of the items as the skewness and kurtosis values were below the cut-off points (± 2 for skewness and ± 7 for kurtosis; see Table 1). Graphical inspection of the various histograms and normal probability plots also supported normality for the majority of the items. The

only exception was Item SN6 of the Subjective Norms scale (e.g., “Babies cry in the middle of the night to make their parents angry”) which showed signs of severe non-normality as the associated skew and kurtosis values exceeded the respective cut-offs (Skew = 4.00; Kurtosis = 17.31). Attempts were made to render the item less skewed by using simple both root square and log transformations, even though such transformations generally are not used because they change the original response categories and make interpretation difficult. However, whenever the item was entered in the model, the program either did not converge (i.e., error message) or did not provide a viable solution (i.e., output with appropriate statistics was not produced). For these reasons, this item was removed from further analyses.

Testing IRT assumptions: Unidimensionality and local independence. For each of the PDEP scales, unidimensionality and local dependence were tested prior to performing the IRT analyses. Unidimensionality refers to the statistical dependence among items (observations on the manifest variable), which is a function of a single latent trait (Osterlind, 1983). The assumption of unidimensionality can be also be examined by comparing the ratio of the first with the second eigenvalue in exploratory factor analysis. The ratio of the first to second eigenvalues of the polychoric matrix greater than two is indicative of essential unidimensionality (Hambleton, Swaminathan, & Rogers, 1991; Lord, 1980). Alternatively, one can examine the amount of variance explained by the first factor. Reckase (1979) recommended that, to obtain stable item parameters, the first factor should account for at least 20% of the variance to obtain stable item parameters. The second important assumption of IRT models is local dependence. Local dependence refers to the fact that the probability of a randomly-chosen individual in a homogenous group of individuals responding incorrectly to an item is not affected by the response of the same individual to another item on the same latent trait scale (Lord & Novick, 1968). The LD- χ^2 statistic is used to evaluate local item dependencies. The statistics are computed by comparing the observed two-way cross-tabulation table to the expected cross-

tabulation table for each item pair. Values exceeding 10 indicate a violation of local dependency (Chen & Thiessen, 1997).

IRT analyses: The graded response model. To determine the psychometric properties of the PDEP questionnaire, a series of exploratory IRT models were fit to the data. The choice of this analytic method was justified on the basis of theoretical grounds. The PDEP questionnaire is composed of four scales that are assumed to represent four distinct latent constructs. However, the psychometric properties of the scales have not yet been analyzed; for this reason, an exploratory approach was deemed appropriate. IRTPRO 4.20 (Cai, Thissen, & du Toit, 2011) was used to conduct these analyses. Maximum Likelihood (ML) is the default for the estimation method in IRTPRO (Cai, Thissen, & du Toit, 2011). Samejima's (1969) Graded Response Model (GRM) was used as the main response function model for the IRT analyses. GRM is appropriate for polytomous data, such as Likert-type items with an underlying response continuum. Moreover, in GRM, the response scale does not have to be based on the same ratings (e.g., items can be measured on different Likert scales). GRM models relate item parameters and characteristics of the respondents (latent trait) to the probability of endorsing multiple-response categories.

GRM assesses the performance of a psychometric scale by means of item discrimination and category thresholds. Item discrimination coefficients determine the strength of the relationship between each item with the underlying latent construct and are equivalent to unstandardized factor loadings in a factor analysis. The category threshold and is defined as the respondent's trait level (i.e., item score) associated with a 50% probability of responding to a specific response category or higher. For any polytomous item j , there are $k-1$ thresholds (b_j), where k is equal to the number of response categories. (ICC; De Ayala, 2013). Item information curves (IICs) are also used to assess the contribution of each item to the overall estimation of the latent trait. IICs provide the most direct index of the precision of measurement or reliability of individual items.

Assessing model-data fit. Following an assessment of the model dimensionality, an assessment of model-level fit was performed. Model fit assessment pertains to the assessment of the overall fit of the model as well as the fit of the single items. This involves the assessment of the goodness-of-fit statistics that provide information concerning discrepancies in responses from the respective parameters. Hence, the Root Mean Square Error of Approximation (RMSEA; Steiger & Lind, 1980) was used to assess the fit of the model. Hu and Bentler (1999) recommend that the values of RMSEA be lower than 0.06 for a close fit of the statistical model. The Akaike information criterion (AIC; Akaike, 1974), Bayesian information criterion (BIC; Schwartz, 1978) and M_2 (Maydeu-Olivares & Joe 2005; Maydeu-Olivares & Joe, 2006) were used to compare competing non-nested models. Lower values of these indices indicate a better fit. These statistics have high sensitivity for model misfit and high reliability when used in combination with other statistics such as the RMSEA (Xu, Paek, & Xia, 2017).

Assessing item fit. A number of parameters and statistics were also examined at the item level. Item retention was based on concurrent consideration of all the available indexes. First, quantitative indices, such as the discrimination and thresholds parameters, were inspected. Discrimination parameters are equivalent to factor loadings and represent the strength of the relationship between the item and the latent trait. Discrimination can be classified as follows: 0.01-0.24 is 'none'; 0.25-0.64 is 'low'; 0.65-1.34 is 'moderate'; 1.35-1.69 is 'high'; and > 1.7 is 'perfect' (Baker, 2001). Threshold parameters (b, difficulty parameters) that span the range of -4 to + 4 are usually interpreted in terms of item difficulty/location with respect to a given response category. With respect to the IICs, no accepted quantitative standard exists to judge the obtained item and test information values. Rather, visual inspection of the item information functions is performed. Greater information is associated with greater measurement precision. Information is inversely related to the standard error of the estimate; therefore, greater information will result in a smaller standard error associated with the estimated latent

trait score. Typically, narrower and higher curves imply higher accuracy, whereas shorter and wider curves imply lower accuracy. Generally, very low height IICs may indicate poor fit of the item and as such could be used to make judgments regarding the removal of any specific item (De Ayala, 2013).

This second description pertains to the analyses conducted for the content validity examination of the PDEP questionnaire using the sample of professionals and the sample of parents.

Content validity analyses. For the content validity assessment, mean representativeness and mean clarity for individual items were examined. The comments made by professionals and parents were examined together with the focus group field notes using a note-based analysis, which may be used when the purpose of a study is narrowly defined (e.g., pilot-testing items; Krueger, 2014). In terms of analytical strategy, the notes were first read entirely to understand the main ideas. Then, they were reread, and all codes were generated by assigning similar comments to the same code. Also, comments that diverged from the majority of participants were noted and included in the final report.

Results

This first set of results pertains to the IRT analyses that were conducted on the original PDEP questionnaire.

Unidimensional Analysis

Results of the principal component analysis supported a primary dimension for the majority of the PDEP scales. The first three eigenvalues for each scale were as follows: 2.82, 0.82, and 0.56 for Approval of Physical Punishment; 1.65, 0.93, and 0.82 for Approval of Non-Physical Punishment; 2.17, 0.83, and 0.75 for Subjective Norms; and 1.78, .89, and .74 for Parenting Self-Efficacy. Lord's condition for essential unidimensionality, as demonstrated by the greater than 2 ratio of the first to second eigenvalue (Hambleton, Swaminathan, & Rogers, 1991; Lord, 1980), was met by all the scales with the exception of Approval of Non-Physical Punishment. The ratio of the first to second eigenvalue was 3.43 for Approval of Physical Punishment, 1.77 for Approval of Non-Physical Punishment, 2.61

for Subjective Norms, and 2.01 for Parenting Self-Efficacy. However, the first factor of each of the scales accounted for more than 20% of the variance, providing support for the unidimensionality of all the four scales (Reckase, 1979). The variance accounted for by the primary dimension for the four scales was 56.4% for Approval of Physical Punishment, 41.2% for Approval of Non-Physical Punishment, 43.4% for Subjective Norms, and 44.5% for Parenting Self-Efficacy.

IRT Results for the Approval of Physical Punishment Scale

Results of the IRT analyses for the Approval of Physical Punishment scale are displayed in Figure 2 and Tables 2, 3, and 4. Model fit statistics for the original questionnaire indicated an acceptable fit, as demonstrated by the RMSEA values below the recommended cut-off ($< .06$; see Table 2). Discrimination and category threshold values are presented in Table 3. The majority of the discrimination parameters or item slopes (i.e., a values) were above 1.70 indicating perfect fit, except for item APP4 (e.g., If parents don't use punishment, their children will be spoiled) which presented a moderate discrimination value ($a = 0.87$). The thresholds (i.e., difficulty parameters, b values) were spread appropriately within each response category. Examining the threshold statistics for each item revealed that the items had an ascending trend from response categories *strongly disagree* (threshold b_1) to *strongly agree* (threshold b_5) and covered a moderate range of the latent trait (overall range from -0.04 to 2.28). All the $LD-\chi^2$ values for the items were acceptable (< 10), thereby excluding local dependence (i.e., unwanted shared variance between pairs of items within a scale; see Table 4). Inspection of the Item Information Curves (IICs) helped to evaluate reliability at the item level (See Figure 2). The IIC of item APP2 (e.g., Spanking is fine as long as the parent is not angry) was noticeably higher than all the other items. The lowest IIC was detected for item APP4 (e.g., If parents don't use punishment, their children will be spoiled), indicating high levels of measurement error for this particular item compared to other items on the same scale.

IRT Results for Approval of Non-Physical Punishment Scale

Results of the IRT analyses for the Approval of Non-Physical Punishment scale are displayed in Figure 3 and Tables 2, 5, and 6. Model fit statistics indicated a good fit, as demonstrated by the RMSEA values below the recommended cut-off ($< .06$; See Table 2). Discrimination and category threshold values are presented in Table 5. Discrimination values for individual items were above 1.70, (which indicates perfect fit) with the exception of item ANP3 (e.g., If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay inside until he changes it) which presented a moderate value ($a = 0.89$) and ANP4 (e.g., If children break the rules, their parents should take away privileges), which presented a low discrimination value ($a = 0.43$). Examining the statistics for each item revealed that the items had an ascending trend from response categories strongly disagree (threshold b1) to strongly agree (threshold b5), covered a reasonably wide range of the latent trait (range from -3.27 to 6.25), and were in the correct order. An exception was ANP1 (e.g., If a 14-year-old is failing in school, his parent should make him do extra chores) which showed a positive value at the lower threshold level ($b1 = 0.22$), thereby indicating low sensitivity in detecting low levels of the latent trait. Inspection of the Item Information Curves (IICs) helps to evaluate reliability at the item level (See Figure 2). All the LD- χ^2 values for the items were acceptable (< 10), thereby excluding local dependence (see Table 6). The IIC of item ANP1 (e.g., If a 14-year-old is failing in school, his parent should make him do extra chores) was noticeably higher than all the other items. The lowest IIC was detected for item ANP4, (e.g., If children break the rules, their parents should take away privileges) indicating high levels of measurement error for this particular item compared to other items on the same scale.

IRT results for Subjective Norms Scale

The results of the IRT analyses for the Subjective Norms scale are displayed in Figure 4 and in Tables 2, 7, and 8. Model fit statistics indicated a good fit, as demonstrated by the RMSEA values below the recommended cut-off ($< .06$; See Table 2). Table 7 presents the discrimination and category

thresholds for all the items. Discrimination values ranged from 0.78 to 1.93 indicating a moderate to strong relationship to the latent trait. Item SN2 (e.g., If children have tantrums, they are probably spoiled) demonstrated higher slope parameters, suggesting that this item has higher discriminatory power than the other items and provides more information about different levels of the latent trait (See Table 2). The category thresholds were spread appropriately within each response category. For the majority of the items, the range of thresholds (overall range from -1.01 to 2.15) along the latent trait was moderate and in the correct order, implying that most items could measure a reasonable range of the latent trait. An exception was SN5 (e.g., If an 8-year-old uses bad words in front of his parents, this is a sign of disrespect.) which demonstrated high sensitivity for low threshold values ($b_1 = -1.01$) coupled with low sensitivity for high thresholds values ($b_5 = 0.98$). Thus, this item was able to accurately represent low scores of the latent trait but not higher scores. All the LD- χ^2 values for the items were acceptable (< 10), thereby excluding local dependence (see Table 8). Figure 4 presents information curves for all the items. The IIC for item SN2 (e.g., If children have tantrums, they are probably spoiled) was noticeably higher than all the other items, indicating that it provided more information about different levels of the latent trait. The lowest IIC was for item SN5 (e.g., If an 8-year old uses bad words in front of her/his parents, this is a sign of disrespect). This indicates high levels of measurement error for this item compared to other items on the same scale.

IRT Results for the Parenting Self-Efficacy Scale

Results of the IRT analyses for the Parenting Self-Efficacy scale are displayed in Figure 5 and in Tables 2, 9, and 10. Model fit statistics indicated a good fit as demonstrated by RMSEA values below the recommended cut-off ($< .06$; See Table 2). Discrimination values ranged from 1.08 to 1.36, indicating a moderate relationship to the latent trait (see Table 9). The category thresholds were spread appropriately within each response category. Within each item, the range of thresholds along the latent trait was wide (overall range from -4.44 to 4.87) for most items and in the correct order, implying that

most items could measure a reasonably wide range of the latent trait. LD- χ^2 values for the items were acceptable (< 10), except for Item SN1 (e.g., Most people are better parents than I am) with both SN2 (e.g., I can solve most of my parenting challenges) and SN3 (e.g., I have the skills to be a good parent), thereby indicating unwanted shared variance between each pair of items (e.g., SN1-SN2 and SN1-SN3; see Table 10). Figure 5 presents information curves for all the items. The IIC for item SN4 (e.g., As a parent, I often feel like I just don't know what to do) was noticeably higher than all the other items. The lowest IICs were detected for item SN2 (e.g., I can solve most of my parenting challenges). However, this item seemed to perform relatively well (i.e., information curve of the approximately same height) compared to other scale items (including SN3).

Summary of Preliminary IRT Findings and Initial Modifications to the PDEP Questionnaire

To consider the IRT results and possible modifications to the original PDEP questionnaire, a team of four researchers, including myself, Dr. Romano (supervisor), Dr. Durrant (PDEP program developer), and Dr. Stewart-Tufescu (post-doctoral student at the University of Manitoba) met to make a number of edits. The results of the IRT analyses were discussed with the goal of creating a scale that had statistically good-fitting items (at least relative to the preliminary IRT results) and that included a variety of items to offer good coverage of the construct being assessed in each scale of the questionnaire.

Findings from the IRT analyses suggested that the four scales of the PDEP questionnaire performed moderately well and demonstrated an acceptable fit to the model. However, several issues emerged. For the Approval of Physical Punishment scale, several of the current items were reworded to make them more relevant to the specific age group (See Table 11). IRT results showed that one item showed low information/high measurement error (Item APP4: “If parents don't use punishment, their children will be spoiled”) as indicated by a flat ICC curve. Hence, after consulting with the research team, a unanimous decision was made to change slightly the item (along with other items such as

APP1, APP2 and APP3) to specify both spanking and slapping as forms of physical punishment “*If parents don’t spank or slap their children, they will be spoiled*”. Moreover, eight items were added to include acts of physical punishment toward older children and more subtle aspects of the use of physical punishment: 1) *It’s ok to slap a teenager’s face if she uses bad words in front of her parents*; 2) *Young children who aren’t spanked will get into trouble when they are teenagers*; 3) *Spanking teaches children right from wrong*; 4) *If a 4-year-old won’t leave a store, his parent should pull his arm and drag him out*; 5) *It’s ok to shake a 3-year-old to make her stop screaming*; 6) *If a 2-year-old bites another child, it’s ok to bite him to teach him how it feels*; 7) *Children who are spanked learn to respect their parents*; and 8) *It’s ok to slap a 2-year-old’s hand for touching dangerous objects*. In sum, the revised scale now had 13 items in total – 5 that were retained from the original (with minor word changes) and 8 that were added (See Table 11).

For the Approval of Non-Physical Punishment Scale, a problematic item was ANP4 (e.g., *If children break the rules, their parents should take away privileges*) as indicated by an almost flat information curve. However, given the short length of the scale, the item was retained for the second round of analyses and new items were added to increase coverage of the latent construct. A total of six items were added to the Approval of Non-Physical Punishment Scale: 1) *If a 3-year-old won’t eat, his parent should make him stay at the table until he finishes his food*; 2) *If two sisters won’t stop fighting, the parent should threaten to take away their phones*; 3) *A good way to teach a 13-year-old to dress properly is to embarrass her in front of her friends*; 4) *If a 6-year-old won’t get dressed, his parent should threaten to take him out in his underwear*; 5) *A 10-year-old who hurts another child should be forced to apologize*; and 6) *If a 3-year-old won’t share his toys, he should be made to sit on a chair for 3 minutes*. In sum, the revised scale now had 10 items in total – 4 that were retained from the original and 6 that were added (see Table 11).

Turning to the Subjective Norms scale, low information value was obtained for SN5 (e.g., If an 8-year-old uses bad words in front of her/his parents, this is a sign of disrespect), indicating low reliability of this item. However, given the short length of the scale (and the moderate discrimination value), this item was kept for the second set of analyses. Moreover, the wording of some of the items (e.g., SN1: Two-year-olds who say “no!” are being defiant) was change slightly to make the items more relevant to a specific age-group. In addition, five items were added to increase the coverage of the latent structure (and include items relative to different age groups): 1) *A 16-year-old who questions her parents’ religious beliefs is being disrespectful*; 2) *A 5-year-old child who won’t go to sleep is being stubborn*; 3) *A 6-year-old child who won’t show affection to his grandmother is rude*; 4) *Boys should learn not to cry by the time they are 12 years old*; and 5) *Teenagers who live at home should always follow their parents’ rules, even if they don’t agree*. In sum, the revised scale now had 10 items in total – 5 that were retained from the original (with minor word changes) and 5 that were added (see Table 11).

Finally, the Parenting Self-Efficacy scale was reevaluated, and all the items were reworded to reflect not exclusively parents’ parenting self-efficacy in general but also self-efficacy in dealing with challenging child behaviours and avoiding conflict with the child, which is one of the main objectives of the PDEP program. The item “*Most parents are better parents than I am*” was dropped because was deemed too broad. Six additional items were added: 1) *I can control my anger in challenging situations with this child*; 2) *I feel badly about how I handle challenging situations with this child*; 3) *I feel hopeful that I can improve my relationship with this child*; 4) *In challenging situations, I end up shouting at this child*; 5) *I believe that I will be able to handle future challenges as this child gets older*; and 6) *I feel hopeless about my relationship with this child*. In sum, the revised scale now had 9 items in total – 3 that were retained from the original (with minor word changes), 1 that was removed, and 6 that were added (See Table 11).

Content Validity

This second set of results pertains to the content validity analyses conducted on the PDEP questionnaire following the modifications that were made based on the IRT analyses and the consultation with the research team and program developer.

PDEP Item Ratings for Experts and Parents

Table 12 presents the results of the experts' content validity assessment of the revised PDEP questionnaire. The mean representativeness of the Approval of Physical Punishment items was 2.8 ($SD = 0.3$), ranging from 2.5 to 3.0, and thereby indicating that items were overall highly representative. The mean clarity across all the items was 2.7 ($SD = 0.5$) and ranged from 2.2 to 3.0, thereby indicating overall moderate clarity. Experts rated the majority of items as having high clarity and being appropriate in the use of terms in reference to physical punishment. Experts commented on a few items. In reference to item APP3 (i.e., Parents should have the right to decide whether to spank or slap their children), two experts were concerned that parental rights may not necessarily imply that they endorse the use of child physical punishment. Moreover, an expert suggested removing the term "drag out" from item APP9 (e.g., If a 4-year-old won't leave the store, his parent should pull his arm and drag him out). In addition, others did not agree with the use of the word "dangerous" to define behaviour in reference to item APP5 (e.g., It's ok to spank a 5-year-old's bottom if she does something dangerous) and APP13 (e.g., It's ok to slap a 2-year-old's hand for touching dangerous objects).

The mean representativeness of the Approval of Non-Physical Punishment items was 2.7 ($SD = 0.3$), ranging from 2.2 to 3.0, thereby indicating that items were overall moderately representative. Experts rated item ANP10 (e.g., If a 3-year-old won't share his toys, he should be made to sit on a chair for 3 minutes) unanimously highly representative. The mean clarity was 2.6 ($SD = 0.5$), ranging from 2.1 to 3.0, thereby indicating that items were overall moderately clear. Experts commented on a few items. In reference to item ANP2 "*If a 12-year-old gets into trouble at school, her parent should punish*

her before she has a chance to give excuses”, three experts pointed out that the term “punish” was too general and as such could be interpreted as “physical punishment.” Similarly, an expert indicated that ANP4 “*If children break the rules, their parents should take away privileges*” was too broad and needed to be respecified. Others were concerned that the wording in item ANP8 “*If a 5-year-old won’t get dressed, his parent should threaten to take him out in his underwear*” and ANP9 “*A 10-year-old who hurts another child should be forced to apologize*” may not be appropriate and needed to be respecified. However, no specific suggestions were offered in reference to the use of alternate wording. Moreover, an expert indicated that item ANP3 “*If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay until he changes it*”, among other items (not specified) referring to teenagers, was odd.

The mean representativeness of Subjective Norms items was 2.7 ($SD = 0.3$), ranging from 2.4 to 2.9, thereby indicating that items were moderately representative. Experts rated the majority of the items as being highly representative of the construct of subjective norms. The mean clarity was 2.7 ($SD = 0.5$), ranging from 2.5 to 2.9 and indicating moderate clarity of these items. One expert commented that item SN4 “*A 13-year-old who does not want to be seen with his mother should be ashamed of himself*” was not very clear. However, no specific suggestions were made in reference to the use of alternate wording.

The mean representativeness of Parenting Self-Efficacy items was above 2.8 ($SD = 0.5$), ranging from 2.5 to 2.9 and indicating high representativeness of the items. Experts rated the majority of the items as being highly representative of the construct of parenting self-efficacy. The mean clarity was 2.5 ($SD = 0.2$) ranging from 2.1 to 2.9, thereby indicating moderate clarity of these items. One expert indicated that item SE5 “*I feel badly about I handle challenging situations with this child*” as being not representative of the construct of interest. Moreover, two experts commented on item SE6 “*I feel hopeful that I can improve my relationship with this child*” and indicated that this item was

misaligned to the construct of reference. They suggested rewording the item to better reflect parenting self-efficacy (e.g., There are things I can do to improve my relationship with this child). Others were not convinced that item SE9 “*I feel hopeless about my relationship to this child*” tapped into the construct of self-efficacy.

Parents’ mean ratings of the representativeness and clarity of each item are presented in Table 13. The mean representativeness of Approval of Physical Punishment items was 2.5 ($SD = 0.3$), ranging from 2.2 to 2.7 indicating overall moderate representativeness of the items. The mean clarity was 2.7 ($SD = 0.5$) ranging from 2.5 to 2.8, indicating moderate to high clarity of the items. Parents reported several concerns with item wording (i.e., the use of more than one term to refer to physical punishment and the use of words such as “swat” or “slap” in item APP1 and APP3). In line with the experts’ findings, some parents did not agree with the use of the word “dangerous” to define behaviour in reference to item APP13 (e.g., It’s ok to slap a 2-year-old’s hand for touching dangerous objects). Similarly, a parent suggested removing the term “drag out” from item APP9 (e.g., If a 4-year-old won’t leave the store, his parent should pull his arm and drag him out). Moreover, two parents did not agree with the use of “he/she” in APP6 “*It’s ok to slap a teenager’s face if she uses bad words in front of her parents*” and suggested the use of “they.”

The mean representativeness of Approval of Non-Physical Punishment items was 2.6 ($SD = 0.4$), ranging from 2.4 to 2.8 and indicating overall moderate representativeness. The mean clarity was 2.7 ($SD = 0.4$), ranging from 2.2 to 2.7 and indicating overall moderate clarity. Parents reported several concerns with item wording (i.e., use of more specific terms or changing some of the scenarios). Some parents indicated that the term “extra chores” in item ANP1 “*If a 14-year old is failing in school, his parent should make him do extra chores*” was not clear and too general. In line with what reported by the experts, some parents pointed out that the term “punish” in ANP2 “*If a 12-year-old gets into trouble at school, her parent should punish her before she has a chance to give excuses*”, was too broad

and needed to be modified. Moreover, a parent indicated that the term “hairstyle” in item ANP3 *“If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay inside until he changes it”*, was not clear and that “stay inside” should have been changed. However, no specific suggestions were offered in reference to the use of alternate wording. A parent suggested that for ANP4 *“If children break the rules, their parents should take away privileges”* was opportune to specify the type of favourite activities that the parent would take away (i.e., television, electronics). In addition, some parents suggested to change the term “sisters” in item ANP6 *“If two sisters won’t stop fighting, the parent should threaten to take away their phones”* to “siblings”. Others were concerned that items ANP7 *“A good way to teach a 13-year old to dress properly is to embarrass her in front of her friend”* and ANP8 *“If a 5-year-old won’t get dressed, his parent should threaten to take him out in his underwear”* did not indicate non-physical discipline but rather humiliation. Another parent suggested to respecify item ANP10 *“If a 3-year old won’t share his toys, he should be made to sit on his chair for 3 minutes”* by add alternative forms of punishment.

The mean representativeness of Subjective Norms items was 2.5 ($SD = 0.6$), ranging from 2.4 to 2.7 and indicating moderate representativeness of the items. The mean clarity was 2.8 ($SD = 0.4$), ranging from 2.7 to 2.9 and indicating high clarity. One parent reported not being familiar with some of the terminology, such as “tantrum” and “defiant” in item SN1 *“Two-year-olds who say “no!” are being defiant”* and SN2 *“Usually, 3-year-old have tantrums because they are spoiled”*, respectively. Moreover, some parents suggested several changes in the terms used to describe the scenarios, such as the use of gender-neutral language. One parent commented that term “affection” should be replaced with “love” in item SN8 *“A 6-year old who won’t show affection to his grandmother is being rude”* and that “rude” was not very clear, it should have been replaced by “mad”. Moreover, two parents indicated that item SN9 *“Boys should not cry by the time are 12 years old”* was inappropriate and should be removed.

The mean representativeness of Parenting Self-Efficacy items was 2.8 ($SD = 0.6$), ranging from 2.6 to 3.0 indicating overall moderate representativeness. The majority of items were rated as highly representative, with one item being unanimously rated as highly representative (Item SN1: *I can solve challenging situations without punishing this child*). The mean clarity was 2.9 ($SD = 0.4$) ranging from 2.50 to 3.0 and indicating overall high clarity. A few parents indicated that SE5 “*I feel badly about how I handle challenging situations with this child*” should have been changed to reflect the parent feeling “guilty” or “sad” rather than bad. One parent commented on item SE8 “*I believe that I will be able to handle future challenges as this child gets older*” should be respecified as “when the child grows up”. Others were not certain about the use of the term “hopeless” or “helpless” in item SE9 “*I feel hopeless about my relationship to this child.*”.

Parent Focus Group Findings

When asked whether there were any questions that seemed confusing, a number of parents indicated that the questionnaire was generally easy to understand. However, they also noted that certain terms (i.e., defiant, tantrum) may not be familiar to some parents, particularly new immigrants. Moreover, two parents were unclear about the meaning of several verbs (e.g., spank and swat) and indicated that adding a definition would help to make the items less confusing.

Parents were also asked about the language used to formulate the question and if it was easy or difficult to understand. A number of parents indicated that the language was generally easy to understand. One parent expressed concern over the appropriateness of the word “dangerous” in Item 5 of the Approval of Physical Punishment scale (i.e., It’s ok to spank a 5-year-old’s bottom if she does something dangerous), noting that a number of children’s actions could be considered dangerous to their safety. This parent suggested describing the behaviour in simple terms and providing specific examples to contextualize what behaviour may be considered dangerous. Several parents indicated that two items on the Approval of Non-Physical Punishment scale, namely Item 7 (i.e., A good way to teach

a 13-year-old to dress properly is to embarrass her in front of her friends) and Item 8 (i.e., If a 6-year-old won't get dressed, his parent should threaten to take him out in his underwear), imply humiliation and should be reworded to match the definition of non-physical punishment (i.e., taking away privileges). Another parent expressed concerns regarding the cultural relevance of an Item 8 of the Subjective Norms scale (i.e., A 6-year-old child who won't show affection to his grandmother is being rude). She noted that not all cultures place importance on a child showing affection to family members or initiating physical contact with them. However, the majority of other parents seemed to agree that this item was relevant (and culturally appropriate). Finally, several parents suggested the use of gender-neutral language and proposed removing Item 9 (e.g., Boys should learn not to cry by the time they are 12 years old) from the Subjective Norms scale.

Summary of Content Validity Analysis Results

Overall, the content experts and parents provided evidence that the items adequately represented the various PDEP questionnaire scales and that the items were clearly stated. The two participant groups also made a number of suggestions related primarily to changes in wording. The four members of the research team reviewed the questions in light of the feedback provided by the expert samples and made several changes to the questionnaire, as outlined in Table 14. Particular attention was paid to items that received negative feedback from both professionals and parents.

For the Approval of Physical Punishment scale, the language in item APP6 *"It's ok to slap a teenager's face if she uses bad words in front of her parents"* was flagged as problematic (i.e., for the term "bad words" being too specific). This item was changed to *"if she insults her parents"* to be more general in nature. The term *"drag him out"* in Item APP9 *"If a 4-year old won't leave the store, his parent should pull his arm and drag him out"* was considered inappropriate by several of the experts and parents. The term was deleted, and the item was modified to read *"pull him out by his arm."* Finally, the term *"dangerous"* in item APP13 *"It's ok to slap a 2-year old's hand for touching*

something dangerous” was highlighted by both professionals and parents as problematic because of the lack of contextual information. This item was changed to *“touching an object he/she was told not to touch.”*

For the Approval of Non-Physical Punishment scale, several parents indicated that the term *“extra chores”* in item ANP1 *“If a 14-year old is failing in school, his parent should make him do extra chores”* was not clear and too general. The item was made more specific by adding the following phrase at the end of the items: *“to teach him to work harder.”* Both professionals and parents indicated that the term *“punish”* in item ANP2 *“If a 12-year-old gets in trouble at school her parent should punish her before she has a chance to give excuses”* was too general and could be misinterpreted as physical punishment. This term was changed to *“scold”* to make it more consistent with the construct of non-physical punishment. The term *“hairstyle”* in ANP3 *“If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay inside until he changes”* was not convincing for several parents; professionals also highlighted this item as being odd. The term was changed to *“clothing.”* Both professionals and parents indicated that item ANP4 *“If children break the rules, their parents should take away privileges”* was too general and could instead include children’s favourite activities (e.g., electronics, sleepover). The item was changed as follows: *“should take away favourite activities, like watching TV or seeing friends.”* Several parents disagreed with Item ANP6 *“If two sisters won’t stop fighting, the parent should threaten to take away their phone”* as sibling rivalry is not exclusive to one sex. Moreover, it is not common for young children to own phones until they are young teenagers. In light of these suggestions, the item was reworded as *“two teenage siblings.”* The majority of parents (and several professionals) agreed that item ANP8 *“If a 6-year-old won’t get dressed, his parent should threaten to take him out in his underwear”* implied humiliation and should be changed to reflect more inductive forms of punishment. As such, the item was changed to *“make him stay in his room until he changes his mind.”* A parent did not agree with the term *“forced to”* in ANP9 *“A 10-year old who hurts*

another child should be forced to apologize.” The research team changed the item slightly to specify a 7-year old child to reflect a situation commonly involving younger children.

With respect to the Subjective Norms scale, the research team members agreed to substitute the term “are rude” in SN3 “*Four-year-olds who interrupt adults are rude*” to “*should know better.*” One parent expressed difficulties in understanding some of the terminology (e.g., defiant in SN1, tantrum in SN2). It was not clear if the lack of clarity was because of the parent’s English proficiency or because they were a newcomer to Canada. Since almost all other parents did not raise concerns, the research team decided to keep these items. The majority of parents did not agree with Item SN9 “*Boys should learn not to cry by the time they are 12 years old*” because of the gender stereotype. Hence, the item was deleted.

For the Parenting Self-Efficacy scale, Item SE6 “*I feel hopeful that I can improve my relationship with my child*” was criticized by several professionals as not being related to self-efficacy. The item was changed to “*I feel confident.*” A similar concern was raised for SE9 “*I feel hopeless about my relationship this this child.*” Since the research team believed that a key aspect of parenting self-efficacy is associated with a parent’s personal feelings and perceptions in regard to the relationship with their child, a decision was made to keep the item as is. A parent suggested changing the term “end up” in item SE7 “*In challenging situations, I end up shouting at this child.*” The term was substituted with the following phrase “*I am able to stop myself from shouting at this child.*”

Discussion

This study assessed the psychometric properties of the original PDEP questionnaire using IRT methodology and then made several modifications following examination of the results by a team of four researchers, including myself. This modified version was then administered to two samples (i.e., professionals and parents) for content validity assessment. The items were examined in terms of representativeness and clarity. As well, the two samples offered suggestions about the wording of items

and their meaning. Following the content validity assessment, the research team reviewed the comments made by the two expert samples and made changes accordingly. In the decision-making process, attempts were made to include all major suggestions for modification (particularly those that were highlighted by both professionals and parents). However, there were several instances where the decision was made to keep the item intact. Changes included: 1) removal of terms that were deemed too harsh; 2) the addition of specificity and context to several items; and 3) (in some instances) the removal of gendered language. Following changes based on the findings of the content validity assessment, the final version of the PDEP questionnaire was developed.

Strengths and limitations

One of major strengths of this study was the use of both quantitative and qualitative techniques to examine the psychometric characteristics of the PDEP. In addition, these techniques were used in a sequential manner, beginning with the IRT analyses and associated modifications and then followed by the content validity assessment (professional and parent feedback) and associated modifications. Several limitations, however, need to be mentioned. First, only a small number of fathers participated in this study, namely 16% in the preliminary IRT analyses based on existing data and none in the content validity based on parent focus groups. As such, feedback on the PDEP questionnaire primarily represents the perspective of mothers. However, the fact that the sample from this study was composed almost exclusively of mothers, maps onto the majority of parents who are likely to participate in the PDEP parenting groups. For instance, Durrant et al., 2014 reported that 87% of participants (279 out of 321) were mothers. Similarly, in Study 3 of this dissertation, about 84% of participants (69 out of 83) were mothers. Nonetheless, future research should attempt to validate (and perhaps adapt) these scales in father samples. Second, there was little socio-demographic diversity in both the sample of professionals (almost all female and White) and the sample of parents (almost all non-White and female). Moreover, it was not possible to track if two partners (i.e., dyads/couples) concurrently took

part in the study. This can raise issues regarding the dependence of the data. Future studies should attempt to measure and control for this possibility.

Study 2: Examining the Psychometric Properties of the Revised Positive Discipline in Everyday Parenting Questionnaire**Abstract**

This study assessed the psychometric properties of the revised Positive Discipline in Everyday Parenting (PDEP) questionnaire, internal consistency, test-retest reliability, as well as its construct validity (i.e., convergent). The PDEP questionnaire was administered to a sample of 400 Canadian parents recruited through an online platform. Other scales assessing similar constructs as the one measured by the PDEP questionnaire were also administered to assess convergent validity. A series of confirmatory Item Response Theory (IRT) analyses were conducted separately for each of the questionnaire's four scales to assess their structural properties. Overall, the scales demonstrated a good fit, but several items were removed because of low information and discrimination values. Following the removal of these problematic items, the IRT models achieved a better fit. Internal consistency of the four scales (in their final version after the confirmatory IRT analyses) ranged from acceptable to excellent. With respect to test-retest reliability, a subsample of 100 parents completed the PDEP questionnaire four weeks following the first administration. Correlations between the two scores were statistically significant and in the expected direction. Subsequently, convergent validity analyses were conducted to assess the alignment of each scale with its respective construct of interest. Correlations among the measures were statistically significant and in the expected direction. In conclusion, findings indicated good psychometric properties of the PDEP scales, as well as strong evidence of internal consistency, test-retest reliability, and convergent validity.

In the first study, Item Response Theory (IRT) analysis was conducted on the original questionnaire developed for the PDEP program, following which some modifications were made in consultation with members of a research team. This version was then subjected to a content validity assessment with experts in the field of parenting research and child development, as well as with parents who had not completed the PDEP program. These findings were further incorporated into the PDEP questionnaire to arrive at a revised version. This second study had two objectives. The first objective was to conduct an assessment of the psychometric properties of the revised version of the PDEP questionnaire using IRT methodology. This method allows for a detailed item-level analysis based on examination of item discrimination and difficulty and information function (details of this method were discussed previously in this thesis when describing Study 1 procedures).

The second objective was to gather information on the reliability and validity of the revised PDEP questionnaire, after incorporating changes from the IRT analyses. For reliability, the internal consistency of the four subscales was calculated for the sample of 400 parents. Moreover, the revised version of the PDEP questionnaire was administered after 4-weeks from the initial study to a sub-sample of 100 parents from the initial pool of 400 parents, that agreed to complete the PDEP questionnaire a second time for test-retest reliability assessment. Convergent validity was performed to determine the association of the revised PDEP questionnaire with other existing measures of related constructs. Convergent validity relies on gathering convergent evidence to determine the extent to which a specific measure designed to assess the construct of interest is associated with a measure of a related construct (Simms & Watson, 2007). With respect to parental attitudes toward physical punishment, Specifically, the Approval of Physical Punishment and Approval of Non-Physical Punishment scales from the PDEP questionnaire were both correlated with the Attitudes Toward Spanking scale (ATS; Holden et al., 1995). The PDEP Subjective Norms scale was correlated with the Parent Cognition scale (PCS; Snarr, Slep, & Grande, 2009), and the PDEP Parenting Self-Efficacy

scale was correlated with the Domain-Specific Discipline Self-Efficacy subscale (Coleman & Karraker, 2000). These scales were selected based on their conceptual relatedness with the PDEP scales and their well-established psychometric qualities.

The PDEP questionnaire was designed by the program developer (Dr. Durrant) as the primary assessment tool for program outcomes. The rationale for creating the questionnaire was the need for a measure that targeted the set of parental cognitions identified as some of the main mechanisms underlying harsh parental behaviours. Generally, parental cognitions can be classified as follows: 1) *evaluative cognitions*, such as attitudes toward child rearing which are characterized by pre-existing beliefs regarding the use of specific parenting practices; 2) *analytical cognition* including “conflict-promoting attributions” that are centred around attributing blame to the child for challenging behaviour, and 3) *self-efficacy cognitions* that involve the parents’ perception of their sense of competence in their parenting role and their disciplinary style (Bugental & Johnston, 2000).

Attitudes are defined as “an individual’s predisposition, reaction to, or affective evaluation of the supposed facts about an object or situation” (Holden & Edwards, 1989, p.89). In reference to parenting, attitudes reflect a favourable or unfavourable opinion regarding a specific parental style or set of strategies. Parenting attitudes are influenced by parents’ knowledge and expectations toward their children, which are also dependent on socio-cultural traditions as well as parents' own experiences and values (Holden & Buck, 2002). Research indicates that attitudes toward parenting, whether measured using self-reports or using experimental paradigms; have shown to exert a significant impact on the choice of disciplinary styles (Rodriguez, Smith, & Silva, 2016). For instance, in the context of child abuse, findings from a study conducted on a random sample of 450 U.S. families indicate that more favourable attitudes toward parental aggression were associated with an increased likelihood of physical and psychological maltreatment, and exerted a larger effect on these outcomes compared to other predictors including anger expression, overreactive discipline, childhood history of aggression

and child-responsible attributions (Slep & O’Leary, 2007). In the context of physical punishment, attitudes toward spanking constitute a consistent predictor of its use (Holden et al., 1995; Vittrup, Holden, & Buck, 2006).

Interestingly, current research on child rearing attitudes has highlighted the presence of a gap in the measurement of this construct, partially due to the limited availability of valid and reliable measures (Holden & Buck, 2002). In a review of 83 measurement instruments used in the assessment of child rearing attitudes, Holden and Edwards (1989) identified a series of methodological problems regarding the nature and structure of the items, including the use of vague terms, misuse of the third person, irrelevance with respect to the construct of interest, use of “double-barreled” questions. Moreover, a large number of the measurement instruments reviewed by the authors lacked reliability data, with 45% of the measurement instruments reporting internal consistency and only 25% reporting test-retest reliability.

A commonly used instrument to assess attitudes toward physical punishment is the Attitudes Toward Spanking (ATS; Holden, Coleman, & Schmidt, 1995) scale. This 10-item scale is designed to measure parents’ approval of spanking, which is the most common form of physical punishment used with young children. Despite its popularity, this measure has its limitations in that only focuses on spanking and excludes other common forms of physical punishment such as slapping (i.e., hitting any part of a child’s body with the open palm) which is commonly used with older children and adolescents (Afifi et al., 2019). Another widely used instrument to measure attitudes toward of physical punishment is the Value of Corporal Punishment Subscale (10 items) of the Adult Adolescent Parenting Inventory – Version 2 (AAPI-2; Bavolek, & Keene RG, 2001; Conners, Whiteside-Mansell, Deere, Ledet, & Edwards, 2006) which measures parents’ beliefs in the use of physical punishment as a disciplinary method. However, the same as the ATS, the items of the Value of Corporal Punishment scale focus

exclusively on spanking. Moreover, neither the ATS nor the AAPI-2 subscale considers other aspects of non-physically punitive discipline (including the use of time-out). To my knowledge, no existing measure is available, to date, to measure attitudes toward non-physical (emotional) punishment. The PDEP questionnaire is unique in that it includes two scales: 1) Approval of Physical Punishment, which measures parents' approval of the use of spanking and other forms of physical punishment (e.g., slapping, grabbing) and 2) Approval of Non-Physical Punishment, which measures parents' approval of the use of emotional punishment (e.g., time-out, removal of privileges, shaming).

With respect to attribution to children's behaviour, negative appraisal of children's behaviour (i.e., seeing children challenging behaviour as intentional) has been associated with punitive and harsh discipline (Lorber, O'Leary, & Kendziora, 2003; MacKinnon-Lewis, Lamb, Hattie, & Bardaran, 2001). Among self-reports of parental attribution, the Child Vignettes (Plotkin, 1983) employs 18 brief hypothetical scenarios in which children display challenging behaviours. Participants are told to imagine that the child is their own and rate the vignettes. This measure is used to assess the degree to which parents attribute negative intentionality to child misbehaviour (e.g., Negative Attribution rating) and the degree to which they would punish for the misbehavior (e.g., Punishment rating). These vignettes focus on children between infancy and toddlerhood, and a few in early elementary school-aged years (Plotkin, 1983).

Another self-report scale that is used to measure negative attributions is, the Parent Opinion Questionnaire (POQ; Twentyman, Plotkin, Dodge, & Rohrbeck, 1981). This measure is composed of 80 items measuring unrealistic expectations of children including 24 items targeting parents' poor judgment of children developmental abilities (e.g., It's reasonable to think that most 5-year-olds can cross a busy street and buy groceries at a corner store). In an initial validation of scale conducted with a small sample of parents ($n = 30$), the scale was subdivided into six subscales: Self-Care, Family Responsibility and Care of Siblings, Help and Affection to Parents, Leaving Children Alone, Proper

Behavior and Feelings, and Punishment (Azar & Rohrbeck, 1986). However, a more recent study conducted with a larger sample of 315 parents failed to replicate the results reported by Azar and Rohrbeck (1986) and reported an uninterpretable factor solution of the POQ because of the low factor loadings and poor reliabilities (range .31-.56). In this latter study, maltreating parents were compared with a group of parents with no history of maltreatment on the item scores of two measurement instruments, namely the POQ and Child vignettes, to determine the degree of endorsement of the items of the two attribution scales as well as their relationship of with observational ratings of parenting behaviour. With respect to the POQ, the majority of the items were endorsed by less than 5% of the abusive parents and the endorsement among non-abusive parents was extremely low (with values near zero percent for the majority of the items), thereby indicating that the items did not hold good discrimination in the non-clinical sample. Moreover, no statistically significant differences in the endorsement of the Negative Attribution rating of the Child Vignettes were found between the two groups. In addition, the predictive validity of neither of the two measures was supported, as the two measures were unrelated to observed parenting behaviour (Haskett, Scott, Willoughby, Ahern, & Nears, 2006).

Another popular self-report measure of attribution to children's behaviour is the Parent Cognition Scale (PCS; Snarr, Slep, & Grande, 2009). This measure distinguishes from two main subfactors: 1) child-responsible attributions which refer to intentionality of behaviour from the respondents' child and 2) parent-causal attributions that refer to parents perceiving themselves as the problem. This measure has demonstrated good psychometric qualities both in terms of reliability and construct validity in a sample of 450 mothers and fathers of young children (mean age 5.44; Snarr, Slep, & Grande, 2009). However, PCS focuses exclusively on the negative perceptions of the target child, and hence its application is limited to the specific individual. On the other hand, for the Subjective Norms of the PDEP questionnaire was created to assess parents' negative attributions of

typical children (and adolescents) behaviours due to their false beliefs regarding developmental norms. Thus, the Subjective Norms scale includes items for all major age groups from infants to middle childhood to adolescence and is not constrained to the target child. Another important characteristic of PDEP questionnaire scale items is the use of “rights-respecting language”. For instance, the use of terms such as “misbehaviour”, “difficult child” or any other statement that would imply children being inherently bad is replaced with more neutral language such as “challenging behaviour” or “children facing difficulties”.

Self-efficacy is a construct first introduced by Bandura (1982) and refers to the individual’s beliefs concerning his or her ability to perform a task or perform a necessary behaviour in a competent and effective fashion. According to the author, self-efficacy plays a central role in the way individuals perceive situations, approach different tasks and execute specific actions in certain situations (Bandura 1982). Parenting self-efficacy beliefs is a critical determinant of parenting behaviour, and research suggests the presence of both direct and indirect effects of this construct on parenting practices (Jones & Pritz, 2005). Generally, parents who score higher on parenting self-efficacy are more likely to engage in positive parenting strategies (i.e., increased collaborative activities, involvement), which in turn relate to an increased likelihood of children’s positive academic and social–psychological outcomes (Ardelt & Eccles, 2001).

With respect to existing measures of parenting self-efficacy, a large number of measurement instruments are dealing with domain-general self-efficacy which focuses on global assessment of self-efficacy and is considered independent from the individual’s self-perceived ability in a specific parenting task (Abidin 1986, Campsis et al. 1986, Gross & Rocissano 1988, Wells-Parker et al. 1990, Teti & Gelfand 1991, Dumka et al. 1996, Coleman & Karraker 1997). In a recent review of 76 studies, 34 self-report measures of parenting self-efficacy were identified (Wittkowski et al., 2017), 22 of

which measured domain-specific self-efficacy, representing parents' sense of competence regarding the use of specific disciplinary practices. However, the majority of the domain-specific measures identified by the authors were designed for parents of young children below 6 years of age, with the exception of the Self-efficacy for Parenting Tasks Indexes (SEPTI; Colaman & Karraker, 2000) which targets parenting self-efficacy in the context of the use of disciplinary practices for children between the ages of 6 and 12 years. The SEPTI is composed of five subscales designed to assess parents' sense of competence pertaining to specific parenting tasks including 1) supporting the child's academic achievement (Achievement); 2) socialization with peers (Recreation), 3) providing structure and discipline (Discipline), (d) providing emotional nurturance (Nurturance), and (e) helping the child maintaining adequate physical health (Heath; Coleman, & Karraker, 2000). However, this scale did not receive a high ranking compared to other domain-specific scales, particularly in terms of content validity (i.e., receiving a score of 1 – indicating doubtful methodology) and test-retest reliability (i.e., receiving a score of 0 – indicating no information available on the replicability of the score; Wittkowski et al., 2017). Moreover, the Discipline subscale of the SEPTI assesses perceived self-efficacy with respect to disciplinary measures but does not make any reference to the use of physical punishment and or managing conflict with children. The Parenting Self-efficacy scale of the PDEP questionnaire was designed to measure parenting self-efficacy in general but also self-efficacy in dealing with challenging child behaviours and avoiding conflict with the child, which is one of the main objectives of the PDEP program. Moreover, the items are not constrained to a specific age group and, therefore, can be applied to both young children and young adolescents.

Study Hypotheses

For the IRT confirmatory analyses, it was expected that the four PDEP scales (Approval of Physical Punishment, Approval of Non-Physical Punishment, Subjective Norms, and Parenting Self-Efficacy) demonstrated a good fit, as indicated by goodness-of-fit statistics in the appropriate range and

moderate to high discrimination values for the questionnaire items. With respect to internal consistency, it was hypothesized that, at a minimum, each of the four scales will achieve acceptable reliability (.70 or higher). For test-retest reliability, it was hypothesized that a significant and very strong positive correlation between the first and second administration (4-week follow-up) of the PDEP questionnaire for each of the four scales will be found.

For the convergent validity assessment, the three main hypotheses include: 1) A significant positive correlation between the Attitude Toward Spanking total score and both Approval of Physical Punishment and Approval of Non-Physical Punishment was expected; 2) A significant positive correlation between the Child-Responsible Attributions of the Parental Cognition Scale subscales and the PDEP Subjective Norms scale was expected; and 3) A significant positive correlation between the Discipline Self-Efficacy subscale of the Self-Efficacy for Parenting Task and the PDEP Parenting Self-Efficacy scale was expected.

Method

Procedures and Sample

Parents ($n = 400$) were recruited through community organizations across Canada, including Parents' Lifeline of Eastern Ontario; Maternal, Newborn, and Child Health Promotion Network; and Ontario Early Ontario Child and Family Centres. Where needed, permission to post the link to the online study on the websites of these organizations was obtained. The study information was also posted on social media platforms (e.g., Facebook, Twitter, Reddit) to explain the purpose of the study, outline inclusion criteria (i.e., having a least one child who is between the ages of birth and 18 years; having access to a computer, being able to read and write English) and provide the study web-link. As an incentive to participate, parents were invited to enter a draw for one of four \$40 gift cards.

For the test-retest reliability assessment, a subsample of 100 parents from the 400 participants who took part in the convergent validity study agreed to completed two assessments: baseline and a re-

test assessment approximately 4 weeks later. All of the 400 participants who participated in convergent validity phase of the project were initially informed of the possibility of being invited to participate in a test-retest validity assessment about 4 weeks later. All these parents were then contacted via email after 4 weeks from the initial assessment. The recruitment email (Appendix K) included a brief message with a link to the consent form and online questionnaire (Appendix L). The study was approved by the Office of Research Ethics and Integrity of the University of Ottawa (REB # H-02-18-352, approved March 3, 2018; see Appendix B).

Table 15 shows that the sample of 400 parents was predominantly female (93.8%) and White (92.8%). The participants' mean age was 35.18 years ($SD = 4.6$), and the majority of participants (68.5%) were married. The number of children in the home ranged from 1-5, with a mean of 2.04 ($SD = 0.9$) and a mean child age of 7.1 years ($SD = 4.9$). The majority of parents had obtained a post-secondary degree (81.6%). For employment, the majority (62.8%) were employed outside the home. Household income was distributed as follows: Less than \$20,000 (10.3%); \$30,000-\$49,999 (14.5%); \$50,000-\$89,999 (33.6%); \$90,000-119,999 (21.3%); \$120,000-150,000 (10.3%); and over \$150,000 (10%). The majority of the participants were recruited through social media (98.9%) with a small percentage recruited through friends (0.8%) and community centres (0.3%).

Table 16 presents the demographic characteristics of the subsample of 100 parents who participated in the test-retest reliability portion of the study. The subsample was predominantly female (94%) and of White (92.8%). The participants' mean age was 36.3 years ($SD = 7.3$), and the majority of participants (67%) were married. The number of children in the home ranged from 1-5, with a mean of 1.9 ($SD = 0.8$) and a mean child age of 8.1 years ($SD = 6.7$). The majority of parents had obtained a post-secondary degree (86.0%). For employment, the majority (63%) were employed outside the home. Household income was distributed as follows: Less than \$29,999 (13%), \$20,000-\$49,999 (17%);

\$50,000-\$89,999 (34%); and \$90,000-119,999 (14%), \$120,000-\$150,000 (6%) and over \$150,000 (16%).

Measures

Socio-demographic questionnaire. Participants answered several socio-demographic questions (i.e., age, gender, ethnic/racial background, income), as well as questions concerning their recruitment (e.g., community centres, friends, social media).

Revised PDEP questionnaire. The revised PDEP questionnaire includes four scales. The Approval of Physical Punishment is composed of 13 items designed to measure parental attitudes toward physical punishment. Each item (e.g., It's ok to slap a 2-year-old's hand for touching an object he was told not to touch) is rated along a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*), with higher scores indicating greater approval of physical punishment. The Approval of Non-Physical Punishment scale is composed of 10 items designed to measure parental attitudes toward non-physical forms of punishment (i.e., emotional punishment). Each item (e.g., If two teenage siblings won't stop fighting, their parent should threaten to take something away, such as their phones) is rated along a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*), with higher scores indicating greater approval of non-physical punishment.

The Subjective Norms scale is composed of nine items designed to measure parents' beliefs regarding children's behaviours. Participants were asked to indicate their agreement with statements attributing typical child behaviour to 'misbehaviour/non-compliance' on the part of the child (e.g., A 6-year-old who won't show affection to his grandmother is being rude). Parents respond to each item on a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*), with higher scores indicating a stronger belief that typical child behaviour is due to 'misbehaviour/non-compliance'.

Finally, the Parenting Self-Efficacy scale is composed of nine items and is designed to measure the degree to which parents feel competent and confident in their ability to manage challenging parenting situation without incurring their anger and use punishment. Each item (e.g., I can control my anger in challenging situations with this child) is rated along a 5-point Likert scale from 1 (*never*) to 5 (*always*), with higher scores indicating greater parenting self-efficacy.

Attitudes Toward Spanking Scale (ATS; Holden et al., 1995). This scale consists of 10 statements (e.g., Sometimes a spank is the best way to get my child to listen) that are rated on a 7-point Likert scale from 1 (*strongly disagree*) to 7 (*strongly agree*). It is intended to measure parents' attitudes toward physical spanking, with a higher score indicating a more positive attitude. The ATS has demonstrated good internal consistency in a sample of college-educated mothers ($\alpha = .81-.91$; Holden et al., 1995). The ATS demonstrated good internal consistency in the current sample of 400 parents ($\alpha = .87$).

Parent Cognition Scale (PCS; Snarr, Slep, & Grande, 2009). The PCS is presented in Appendix M. This scale consists of 30 items designed to assess the degree to which parents endorse child-responsible or parent-causal attributions for child challenging behaviour. Respondents are asked to think about their child's challenging behaviour over the past two months and rate possible causes for such behaviour on a 6-point scale from 1 (*always true*) to 6 (*never true*). Items are reverse-scored so that higher scores indicate greater endorsement of the attributions (either child-responsible or parent-causal). There are 9 child-responsible attributions (e.g., My child tries to get my goat or push my buttons), 7 parent-causal attributions (e.g., I'm not patient), and 14 distractor items that are not used in the scoring (e.g., My child is in a stage). For this study, the Child-Responsible Attribution subscale was used to assess convergent validity of the Subjective Norms scale of the PDEP questionnaire. The Child-Responsible subscale has been found to possess good internal consistency in a community sample of

couples ($\alpha = .81$; Snarr et al., 2009). In the current sample of 400 parents, both the Child-Responsible ($\alpha = .90$) had excellent reliability.

Self-Efficacy for Parenting Tasks Scale – Discipline subscale (SEPTI; Coleman & Karraker, 2000). The Discipline subscale of the SEPTI (see Appendix N) consists of eight items (e.g., I have more trouble with discipline than any other aspect of parenting) measured along a 6-point Likert-type scale ranging from 1 (*strongly agree*) to 6 (*strongly disagree*). It assesses perceived self-efficacy with respect to the use of disciplinary measures. The subscale has been found to have high internal consistency in a sample of mothers ($\alpha = .86$; Coleman & Karraker, 2000). The discipline subscale had a good internal consistency in the current sample of 400 parents ($\alpha = .86$).

Data Analyses

This first description pertains to the analyses conducted for the IRT analyses of the revised PDEP questionnaire.

Data screening for the IRT analysis. Data screening procedures were conducted to examine data entry accuracy, missing values, and assumptions of multivariate analysis. All the screening procedures were conducted in SPSS 25. A Missing Value Analysis (MVA) was performed to determine the percentage of missing values for each of the questionnaire items and the pattern of missingness. However, no missing data were recorded for this sample of parents. Items were also screened for normality through skewness and kurtosis values, with either an absolute skew value larger than 2 or an absolute kurtosis larger than 7 indicating substantial non-normality (West, Finch, & Currant, 1996).

Testing IRT assumptions. For each of the four PDEP scales, unidimensionality and local dependence were examined. Unidimensionality was assessed by comparing the ratio of the first with the second eigenvalue in principal component analysis, with values greater than 2 indicating unidimensionality (Hambleton, Swaminathan, & Rogers, 1991; Lord, 1980). Moreover, the amount of variance explained by the first factor should be at least 20% of the total variance (Reckase, 1979).

Local dependence (i.e., dependence among items) was assessed using the LD- χ^2 statistic, with values exceeding 10 indicating a violation of local dependence (Chen & Thiessen, 1997).

IRT model assessment. Following the preliminary analyses, the IRT method was used to examine the psychometric characteristics of the revised PDEP questionnaire by way of IRTPRO 4.20 (Cai, Thissen, & du Toit, 2011). Maximum Likelihood (ML) is the default for the estimation method in IRTPRO 4.20. Samejima's (1969) Graded Response Model (GRM) was used as the main response function model for the IRT analyses. The overall fit of the IRT models was assessed using the RMSEA (Steiger & Lind, 1980), with a cut off value of 0.06 for a close fit of the data with the statistical model. The Akaike information criterion (AIC; Akaike, 1974), Bayesian information criterion (BIC; Schwartz, 1978), and M_2 (Maydeu-Olivares & Joe 2005; Maydeu-Olivares & Joe, 2006) were used to compare competing non-nested models. Lower values of these indices indicate a better fit. These statistics have shown high sensitivity for model misfit and high reliability in combination with other statistics such as the RMSEA (Xu, Paek, & Xia, 2017).

At the item level, quantitative indices, such as the discrimination and threshold parameters, were inspected. Discrimination parameters (a) are equivalent to factor loadings and represent the strength of the relationship between the item and the latent trait. Discrimination can be classified as follows: 0.01-0.24 is 'none'; 0.25-0.64 is 'low'; 0.65-1.34 is 'moderate'; 1.35-1.69 is 'high'; and > 1.7 is 'perfect' (Baker, 2001). In addition, visual inspection of the item and test information functions are used to assess the reliability of individual items. The narrower and higher the item information curve (IIC), the higher the accuracy of an item. Conversely, a shorter and wider IIC indicates lower accuracy.

This second description pertains to the analyses conducted for the validity and reliability analyses of the revised PDEP questionnaire.

Internal consistency, test-retest reliability, and convergent validity. Prior to data analyses, means and standard deviations were obtained for the total scores of all study variables. Data screening

procedures were conducted to examine data entry accuracy, missing values, and assumptions of multivariate analysis. A Missing Value Analysis (MVA) was performed to determine the percentage of missing values for each of the questionnaire items and the pattern of missingness. However, no missing data were recorded for this sample of parents. Scale scores were also screened for normality through skewness and kurtosis values.

Cronbach's alpha was used to assess the internal consistency of each scale. As a general rule, Cronbach's values of .70 or above are considered acceptable (Nunnally, 1978). In this study, cut-offs determined as follows: “ $\alpha > .9$ – Excellent, $\alpha > .8$ – Good, $\alpha > .7$ – Acceptable, $\alpha > .6$ – Questionable, $\alpha > .5$ – Poor, and $\alpha < .5$ – Unacceptable” (George & Mallery, 2003, p. 231). Test-retest reliability refers to the stability of a test score over time and is measured by administering the same test to the same sample on two different occasions (Drost, 2011). Subsequently, the scores obtained on the two occasions are correlated and the correlation is used to discern whether the measure achieves a stability over time. In order to obtain reliable estimates, it is important that the time-lag between the two measurement periods is neither short nor exceedingly long (Drost, 2011). In this study, the test-retest reliability was assessed by determining the association of the revised PDEP questionnaire measured at Time 1 and the same questionnaire measured 4-weeks later. Pearson correlation coefficients were used to test the convergent validity hypotheses. Cut-off values for interpreting the strength of a correlation between two constructs are the following: $r = 0-0.19$ is a very weak correlation; $r = 0.20-0.39$ is a weak correlation; $r = 0.40-0.59$ is a moderate correlation; $r = 0.60-0.79$ is a strong correlation; and $r = 0.80-1$ is a very strong correlation (Mueller, 1995). All data analyses were conducted in SPSS 25.

Results

This first set of results pertains to the IRT analyses that conducted on the revised version of the PDEP questionnaire.

Data Screening

There were no missing data and no major departures from multivariate normality (see Table 17), except for two items from the Approval of Physical Punishment scale, namely APP6 “*It’s ok to slap a teenager’s face if she insults her parent*” (skew = 3.58; kurtosis = 14.08) and APP10 “*It’s ok to shake a 3-year-old to make her stop screaming*” (skew = 6.34; kurtosis = 50.72). These two items had very little variability across response categories (87-94% of parents strongly disagreed with these items). For the Approval of Non-Physical Punishment scale, ANP7 “*A good way to teach a 13-year-old to dress properly is to embarrass her in front of her friend*” (skew = 3.39; kurtosis = 13.97) had low variability across response categories (85% of parents strongly disagreed with this item). Attempts were made to render these items less skewed by using simple transformation (i.e., collapsing some of the low frequency categories) and more complex methods (i.e., transformations). However, whenever these items were entered in the IRT model, the program either did not converge (i.e., error message) or did not provide a viable solution (i.e., output with appropriate statistics was not produced). For these reasons, these three items were removed from further analyses.

Testing the Assumptions of Unidimensionality and Local Dependence

Results of the principal component analysis supported a primary dimension for the four PDEP scales. The first three eigenvalues for each scale were as follows: 6.05, .98, and .83 for Approval of Physical Punishment; 2.81, 1.32, and 0.91 for Approval of Non-Physical Punishment; 2.71, 1.28, and 0.91 for Subjective Norms; and 4.37, 0.69, and 0.64 for Parenting Self-Efficacy. Lord’s condition for essential unidimensionality, as demonstrated by the greater than two ratio of the first to second eigenvalue (Hambleton, Swaminathan, & Rogers, 1991; Lord, 1980), was met by all the scales. Specifically, the ratio of the first to second eigenvalue was 6.17 for Approval of Physical Punishment, 2.12 for Approval of Non-Physical Punishment, 3.42 for Subjective Norms, and 6.33 for Parenting Self-Efficacy. Moreover, the first dimension of each of the scales accounted for more than 20% of the

variance, providing support for the unidimensional nature of all four scales (Reckase, 1979). The variance accounted for by the primary dimension of the four scales was 55.03% for Approval of Physical Punishment, 28.03% for Approval of Non-Physical Punishment, 30.05% for Subjective Norms, and 48.54% for Parenting Self-Efficacy.

The results presented next relate to the IRT analyses conducted on the revised PDEP questionnaire, with the additional removal of items that violated multivariate normality. These included items APP6 *“It’s ok to slap a teenager’s face if she insults her parent”*, APP10 *“It’s ok to shake a 3-year-old to make her stop screaming”*, and ANP7 *“A good way to teach a 13-year-old to dress properly is to embarrass her in front of her friend.”*

IRT Results for the Approval of Physical Punishment Scale

For the revised version of the Approval of Physical Punishment scale (11 items), the model fit statistics indicated a better fit compared to that of the 5-item version (previously presented in Table 2), as indicated by the lower AIC, BIC, and RMSEA values (See Table 18). As shown in Table 19, the majority of the discrimination values for individual items were above 1.70, which indicates perfect fit of the items (Baker, 2001). The only exception was item APP9 (e.g., If a 4-year old won’t leave a store, his parent should pull him out by his arm), which presented a moderate discrimination value ($a = 1.04$). Moreover, the thresholds were spread appropriately within each response category and had an ascending trend across the response categories from strongly disagree (threshold b1) to strongly agree (threshold b5). The highest discrimination value was recorded for item APP2 (e.g., Spanking or slapping is fine as long as the parent is not angry; $a = 5.33$), indicating a strong relation of this item with the latent trait. The spread of category thresholds (-1.25 to 7.91) across different items seemed reasonably broad, implying a broad amplitude of coverage in measuring the latent trait. The IICs of APP2 (e.g., Spanking or slapping is fine as long as the parent is not angry) and APP8 (e.g., Spanking teaches children right from wrong) were noticeably higher than all the other items, suggesting high

reliability for these two items (see Figure 6). The lowest IIC was detected for item APP9 (e.g., If a 4-year old won't leave a store, his parent should pull him out by his arm), indicating high levels of measurement error for this particular item compared to other scale items. This item was deleted, and the model was rerun. The model fit of the 10-item model showed an improvement as indicated by lower AIC, BIC and M^2 values (see Table 18). The majority of the LD- χ^2 values for the items were acceptable (< 10), thereby excluding local dependence (see Table 20). However, item APP3 (e.g., Parents should have the right to decide whether to spank or slap their children) seemed to share unwanted variance with most of the scale items (range LD- χ^2 from 14.1-29.3). Hence, item APP3 was removed and the IRT model was rerun. The reduced model with nine items showed a slightly better fit compared to both the model with 11 items and the model with 10 items, as indicated by lower goodness-of-fit statistics such as the AIC, BIC, M_2 , and RSMSEA (see Table 18).

IRT Results for the Approval of Non-Physical Punishment Scale

For the revised version of the Approval of Non-Physical Punishment scale (nine items), model fit statistics indicated a good fit of the model compared to the model fit of the 4-item version, as demonstrated by lower AIC, BIC and M^2 values and RMSEA values below the recommended cut-off ($< .06$; See Table 18). As shown in Table 21, the majority of discrimination values (i.e., a , slope parameters) for individual items were moderate and ranged between 0.66 and 1.65 (Baker, 2001), implying a moderate level of similarity among items in measuring the latent trait. The thresholds within each response category were reasonably broad and in the correct order, implying that a good range of the latent trait could be measured by most items. Examining the statistics for each item revealed an ascending trend from response categories strongly disagree (threshold b1) to strongly agree (threshold b5) and covered a wide range of the latent trait (range from -4.01 to 4.72). The highest discrimination value was obtained for ANP10 ($a = 1.65$; If a 3-year-old won't share his toys, he should be made to sit on a chair for 3 minutes), indicating a strong relationship of this item with the latent construct.

Moreover, the IIC of ANP10 was noticeably higher than all the other items, indicating high reliability of this item (see Figure 7). Conversely, ANP9 (e.g., A 7-year old who hurts another child should be forced to apologize) received the lowest discrimination value, indicating a weak relationship with the latent trait when compared to other items on the same scale. The lowest IIC was also detected for ANP9, which signifies a high level of measurement error for this item (see Figure 7). The majority of LD- χ^2 values for the items were acceptable (< 10), with the exception of ANP4 (i.e., If a 10-year-old breaks the rules, her parents should take away favourite activities, like watching TV or seeing friends; see Table 22). Indeed, item ANP4 shared unwanted variance with ANP3 (i.e., If a 16-year-old wears clothing that his parent disapproves of, he should have to stay inside until he changes it), ANP6 (e.g., If two teenage siblings won't stop fighting, their parent should threaten to take something away, such as their phones), and ANP9 (e.g., A 7-year old who hurts another child should be forced to apologize) as indicated by LD- χ^2 values above the cut-off (LD- χ^2 range 24.9-11.00; see Table 22).

In order to address the problem of local dependence between the items, ANP4 (i.e., If a 10-year-old breaks the rules, her parents should take away favourite activities, like watching TV or seeing friends) was eliminated from the model. The removal of ANP4, along with the elimination of ANP9 (which had low discrimination and information), improved the fit of the model (now composed of seven items), as demonstrated by lower goodness-of-fit statistics that included the AIC, BIC, M_2 , and RSMSEA (see Table 18).

IRT Results for Subjective Norms Scale

The model fit statistics of the Subjective Norms scale (nine items) indicated a good fit of the model compared to the 5-item version, as demonstrated by lower AIC and BIC values and an RMSEA value below the recommended cut-off ($< .06$; See Table 18). Discrimination values (i.e., slope for individual items) ranged from 1.00 to 2.05, indicating a moderate to strong relationship to the latent trait (see Table 23). Particularly, SN8 (e.g., A 6-year-old who won't show affection to his grandmother

is being rude) demonstrated a higher slope parameter ($a = 2.05$), suggesting that this item has a stronger relationship with the latent trait and provides more information about different levels of the latent trait.

All the LD- χ^2 values for the items were acceptable (< 10), thereby excluding local dependence (see Table 24). The category thresholds were spread appropriately within each response category. Within each item, the range of thresholds along the latent trait was broad (overall range from -1.76 to 7.03) for most items and in the correct order, implying that most items could measure a reasonable range of the latent trait. Figures 8 presents information curves for all the items. The IIC for SN8 (e.g., A 6-year-old who won't show affection to his grandmother is being rude) was slightly higher than all the other items, indicating high reliability of this item. However, all IIC values were approximately in the same range with no particularly shallow curves, thereby suggesting moderate reliability for all the items of the scale.

IRT Results for the Parenting Self-Efficacy Scale

For the revised version of Parenting Self-Efficacy scale (nine items), model fit statistics indicated a good fit compared to the 4-item version, as demonstrated by lower AIC and BIC values and a RMSEA value below the recommended cut-off ($< .06$; See Table 18). The majority of discrimination values were above 1.70, thereby indicating a nearly perfect fit of the items. The only exception was item SE6 (e.g., I feel confident that I can improve my relationship with this child) which presented a moderate discrimination value ($a = 1.17$; See Table 25), indicating a somewhat weaker relationship of this item with the latent construct when compared to other scale items. The highest discrimination value was recorded for SE9 (e.g., I feel hopeless about my relationship with this child; $a = 2.35$) and SN3 (e.g., I feel like I just don't know what to do with this child; $a = 2.31$), indicating a stronger relationship of these items with the latent trait. The category thresholds were spread appropriately within each response category. Within each item, the range of thresholds along the latent trait was reasonably broad (overall range from -8.02 to 9.22) for most items and in the correct order, implying

that most items could measure a broad range of the latent trait. LD- χ^2 values for the items were acceptable (< 10), thereby supporting local independence (see Table 26). Figure 9 presents information curves for all the items. Moreover, the IICs for both SE9 and SN3 were the highest and indicated high reliability for these two items. The lowest IIC was detected for SE6 (e.g., I feel confident that I can improve my relationship with this child), indicating high measurement error. The removal of SE6 improved the fit of the model (now composed of eight items), as demonstrated by lower goodness-of-fit statistics for the AIC, BIC, M_2 , and RSMSEA (see Table 18).

Summary of the IRT Results

Overall, the four PDEP scales from the revised questionnaire based on Study 1 findings performed moderately well and demonstrated an acceptable fit. In order to further maximize the precision of the scales, IRT analyses were conducted so that poorly performing items were removed, and adjusted models were subsequently re-tested for goodness-of-fit. With respect to the Approval of Physical Punishment scale, item APP9 (low discrimination value and high measurement error) and APP3 (violation of local dependence) were removed. These two deletions led to improved model fit. The final version of the scale was composed of nine items. For Approval of Non-Physical Punishment, ANP9 (i.e., low discrimination and high measurement error) and ANP4 (i.e., violation of local independence) were removed. The deletion of these two items improved the model fit. The final version of the scale was composed of seven items. Turning to the Subjective Norms scale, no changes were made based on the performance of the items, so the scale retained its nine items. Finally, item SE6 of the Parenting Self-Efficacy scale was flagged as problematic as it showed low information values and high measurement error. The removal of this item improved the fit of the model. The final version of the scale was composed of eight items.

Reliability and Validity Results

This second set of results relates to the classical reliability and validity analyses conducted on revised version of the PDEP questionnaire after the confirmatory IRT analyses.

Preliminary and Descriptive Analyses

There were no missing data and no major departures from multivariate normality. Scale distributions were neither skewed nor kurtotic. Table 27 presents the means and standard deviations of all study variables. The majority of mean scores were moderate as indicated by a score near their respective midpoints. Means of the Discipline Self-Efficacy ($M = 33.3$) and revised PDEP Parenting Self-Efficacy ($M = 27.3$) scales were slightly higher than their respective midpoints.

Internal Consistency

With respect to internal consistency, Cronbach's coefficients ranged from .74 to .92 for the current sample of 400 parents. The 9-item version of the Approval of Physical Punishment scale was found to have excellent internal consistency ($\alpha = .92$). The 7-item version of the Approval of Non-Physical Punishment scale was found to have acceptable internal consistency ($\alpha = .74$). The 9-items version of the Subjective Norms scale showed good internal consistency ($\alpha = .82$). Finally, the 8-item version of the Parenting Self-Efficacy showed good internal consistency ($\alpha = .86$).

With respect to the test-retest reliability assessment, correlations for the revised PDEP scales were very strong from the first to second administration four weeks later and ranged from 0.81 to 0.94. The highest correlation was for the Approval of Physical Punishment ($r = 0.94, p < .001$) scale, followed by Approval of Non-Physical Punishment ($r = 0.85, p < .001$), Subjective Norms ($r = 0.83, p < .001$), and Parenting Self-Efficacy ($r = 0.81, p < .001$). The correlations were all very strong indicating excellent reliability of the scales over time.

Convergent Validity

Turning to the convergent validity, there was a very strong and significant positive correlation between Attitudes Toward Spanking and the revised PDEP Approval of Physical Punishment scale ($r =$

0.92; $p < .001$), indicating that greater parental approval of spanking was associated with greater approval of physical punishment. Moreover, there was a moderate and significant positive correlation between Attitudes Toward Spanking and the revised PDEP Approval of Non-Physical Punishment scale ($r = 0.47$; $p < .001$), thereby indicating that parents who held more favourable attitudes toward spanking were more likely to approve non-physical forms of punishment, such as emotional punishment. There was a small positive correlation between Child-Responsible Attributions subscale of the Parental Cognition Scale and the revised PDEP Subjective Norms scale ($r = 0.14$; $p = .004$), indicating that parents who tended to attribute challenging child behaviours to children's own characteristics also had more negative general beliefs regarding normative child behaviours. However, the association between these two variables was very weak. Finally, Discipline Self-Efficacy subscale scores were significantly and positively related to Parenting Self-Efficacy scores of the revised PDEP questionnaire ($r = 0.62$; $p < .001$). This finding indicates that parents who reported greater effectiveness in implementing their parenting strategies (regardless of whether they are punitive or not) also reported greater effectiveness in engaging in non-punitive disciplinary strategies.

Discussion

The first objective of this study was to conduct a confirmatory IRT analysis on the revised version of the PDEP questionnaire obtained in Study 1. IRT was employed to examine the psychometric characteristics of the PDEP questionnaire using a sample of 400 parents and to find the best set of items representing their corresponding latent traits. As hypothesized, the results provided evidence of good psychometric qualities for the PDEP questionnaire. However, the analyses highlighted several problematic items that were removed. Following the removal of these items, the final version of the PDEP questionnaire was confirmed.

Using this final version of the PDEP questionnaire, the second study objective was to conduct a reliability (i.e., internal consistency and test-retest) and validity (i.e., convergent) assessment of the scales. In line with the initial hypothesis, most of the scales (i.e., three out of four) were found to have good to excellent internal consistency, indicating fairly good agreement among items within the scales. Moreover, the test-retest results supported the initial hypothesis as all four PDEP scales were reliable in a sample of 100 parents over a 4-week period. Test-retest reliability coefficients ranged from good (.81 for Parenting Self-Efficacy) to excellent (.94 for Approval of Physical Punishment), providing strong evidence in support of the PDEP questionnaire's temporal consistency and invariance over time.

Furthermore, based on the sample of 400 parents, the hypotheses around convergent validity generally were supported. In line with the first hypothesis, the Attitudes Toward Spanking (ATS) score was significantly and positively associated with both Approval of Physical Punishment and Approval of Non-Physical Punishment from the PDEP questionnaire. The Approval of Physical Punishment scale, therefore, taps into aspects of parenting that are closely related to facets of the ATS scale; indeed, many of the items on the two scales are similar in that they tap into the approval of the use of spanking. This close alignment between the two measures presumably accounts for the high correlation between the two scales. Interestingly, ATS scores were positively related to Approval of Non-Physical Punishment, thereby indicating that attitudes regarding the use of spanking are linked with attitudes about the use of emotional punishment. This is a novel finding in that it is the first study indicating a link between attitudes toward the use of "mild" forms of physical punishment, such as spanking, and approval of non-physical punishment in the form of emotional punishment. A plausible explanation for this relationship is that parents who hold more favourable attitudes toward physical punishment may be more likely to hold favourable attitudes toward punishment strategies more generally (e.g., removal of activities or privileges, time-out). However, this is the first study to examine the relationship between

approval of non-physical (emotional punishment) and attitudes toward physical punishment. Further research is needed to corroborate these findings.

In line with the second hypothesis, the Child-Responsible Attributions subscale of the Parent Cognition Scale (PCS) was positively and significantly related to the PDEP Subjective Norms scale. This indicates the presence of a link between parental beliefs about age-appropriate behaviours and attributions relative to the intentionality of the child (i.e., the child as the cause of challenging behaviour). However, despite the significant correlation, the small magnitude of the coefficient indicates a weak alignment between the two constructs. However, it is important to note that items on the Child-Responsible Attributions subscale tap into a parents' negative attributions to their own child, whereas the Subjective Norms measures parents' beliefs to measure parents' beliefs regarding children's behaviours at different ages and not related only to their own child. This difference might explain the weak relationship between these two measures.

Finally, in line with the third hypothesis, the Discipline Self-Efficacy subscale was significantly and positively related to Parenting Self-Efficacy, thereby indicating the presence of a strong alignment between these two constructs. It is important to note that, despite the relationship between the two scales in assessing one's general sense of competence in the parenting role, the items of the Parenting Self-Efficacy scale from the PDEP questionnaire were more specific to the use of punishment and the ability to manage challenging parenting situations constructively (i.e., without anger or the use of punishment). In contrast, the Discipline Self-Efficacy scale items focuses on self-efficacy on the global sense of discipline (as setting rules and being in charge of decisions around parenting) with no direct reference to the use of punishment or managing conflict with children.

In sum, overall study results indicated that the PDEP questionnaire (in its final version) possesses good psychometric properties. In particular, the four scales of the questionnaire showed adequate to excellent internal consistency and test-retest reliability over a 4-week period, and good

convergent validity. These findings are important since the original PDEP questionnaire had never undergone such a rigorous examination to evaluate its reliability and validity.

Limitations and Future Directions

A major strength of this study was the combination of Classic Test Theory (CCT) and Item Response Theory (IRT) methodologies to assess the psychometric characteristics of the PDEP scales. Generally, the use of Cronbach's alpha has been criticized as being population-dependent (i.e., coefficient changes depending on the sample of test-takers) and subject to possible inflations due to the increasing number of items in a given scale (Reise & Henson, 2003). On the other hand, the IRT approach is not only population-independent (i.e., consistent estimates across different populations) but it also provides a reliability measure for each item (e.g., item information curve) at different levels of the underlying trait so that each item is associated with information parameters (De Ayala, 2013). Thus, the use of IRT in addition to the traditional reliability assessment (i.e., Cronbach alpha) constitutes an advantage in the interpretation the reliability of a measurement instrument in that it provides valuable information at various stages of the scale development.

In parallel with study strengths, there also were several limitations, beginning with the fact that the sample of parents was almost exclusively composed of mothers. It is possible that the characteristics of questionnaire items (i.e., discrimination parameters, thresholds) could change depending on the sex of the parent. Future research should attempt to examine gender differences in IRT-based methods using differential item functioning with a sample composed of an approximately equal number of mothers and fathers. Moreover, the majority of parents who participated in this study were White and had obtained a high level of education. These socio-demographic characteristics are not reflective of the Canadian population as a whole. Despite IRT parameter estimates being considered population-independent (hence, they should provide fairly stable estimates across different

populations), it would be useful to compare the current results with Canadian-based research with more diverse parent populations in terms of gender, ethnicity, and education, to name a few.

A second limitation is that participants were not asked if their partner (if applicable) concurrently took part in the study. The fact that two partners (i.e., dyads/couples) might have participated in the study can raise issues regarding the dependence of the data. Future studies should attempt to measure and control for this possibility. Third, shared method variance is an important limitation of this study, which focused exclusively on parent self-report. Moreover, self-report measures are subject to social desirability biases although it is important to note that all study responses were provided by parents in an anonymous manner. Hence, it would be important to expand the current results by including explicit measures of attitudes and attributions (e.g., observational/analogue measures).

Study 3: A Quasi-Experimental Evaluation of the Positive Discipline in Everyday Parenting Program

Abstract

The purpose of this work was to conduct a quasi-experimental study to evaluate outcomes of the PDEP program whose primary aim is to reduce parental use of physical and emotional punishment and promote non-violent positive discipline. PDEP facilitators from community-based agencies in Ottawa and surrounding regions delivered the program to 83 parents who were assigned to either the PDEP experimental group ($n = 47$) or a PDEP wait-list control group ($n = 36$). All parents completed self-report questionnaires at times corresponding to the experimental group's pre-program, post-program, and 1-month follow-up. At post-program, parents in the PDEP experimental group reported significantly less approval of physical and non-physical punishment, greater number of positive child attributions, greater parenting self-efficacy, greater use of non-physical punishment, lower parenting stress, and greater supportive and proactive parenting, compared with parents in the wait-list group. For approval of physical and non-physical punishment, subjective norms, parenting stress, impulsivity, use of non-physical punishment, and use of supportive parenting, significant improvements for the experimental group (relative to the wait-list parents) were reported at 1-month follow-up. These findings highlight some of the benefits of the PDEP program, particularly as they pertain to simultaneously reducing harsh parenting practices (and their precursors) and promoting positive parenting behaviours. This study provides evidence for the successful implementation of the PDEP program and its effectiveness in improving various aspects of parenting, including cognitive/affective processes (e.g., attitudes and attributions, self-efficacy) and action goals (e.g., stress reduction, emotion regulation).

Physical Punishment: A Threat to Children's Mental Health and Well-Being

Accumulating research evidence supports the negative impact of physical punishment on children across a variety of developmental domains that include behavioural, emotional, moral, and cognitive outcomes (Gershoff, 2010; Gershoff, 2002; Gershoff & Grogan-Kaylor, 2016). A meta-analysis of 75 studies that focused exclusively on spanking yielded 79 unique significant effect sizes; of these, 99% indicated associations between spanking and negative child outcomes including aggression, antisocial behaviour, externalizing and internalizing behaviour problems, negative relationships with parents, lower moral internalization, and slower cognitive development (Gershoff & Grogan-Kaylor, 2016). The effect sizes did not vary based on whether studies were longitudinal, experimental, or cross-sectional in design. Furthermore, the relationships between spanking and child outcomes were similar in size to those found between more severe physical abuse and the same child outcomes.

In a recent longitudinal study on a community-based sample of 3,632 U.S. mothers, Ward and colleagues (2019) examined the relationship between maternal spanking (which is the most common form of physical punishment) and externalizing behaviours obtained through self-reports when the child was one, three, and five years of age. After controlling for demographic factors, child attachment style, and maternal mental health, results indicated that maternal spanking was related to child externalizing behaviours across the first five years of life, and there were no significant differences for ethnic background (i.e., White, African American, American Indian). In a similar study conducted with a sample of 2,705 U.S. families, Grogan-Kaylor and colleagues (2020) explored the relationship between maternal spanking when the child was three years of age and child externalizing behaviours at five years. The authors also examined the effects of neighbourhood risk as a moderator of this relationship. In line with the previous findings, exposure to spanking in early childhood was related to increased child externalizing behaviours two years later. Moreover, such changes did not differ across

levels of neighbourhood risk (Grogan-Kaylor et al., 2020). Finally, a recent study of 1,686 Canadian children from the Quebec Longitudinal Study of Child Development indicated that both boys and girls who experienced physical punishment (i.e., spanking and/or hitting) at 41 months were more likely to exhibit increased externalizing behaviours and conduct problems at age five (Piché, Huynh, Clément, & Durrant, 2017). Altogether, these various findings indicate that spanking can have deleterious effects on children's behaviours regardless of individual-level characteristics, such as ethnicity and gender, or broader community-level factors.

Although most research has focused on externalizing behaviours, physical punishment has also been linked with internalizing behaviours. In a recent study on a nationally representative sample of 4,732 U.K. families, children who experienced physical punishment at three years of age were found to report greater emotional problems at age 11 (Rajyaguru, Moran, Cordero, & Pearson, 2019). Similar findings have been reported by a prospective cross-lagged study that examined the reciprocal influences of parental discipline and child behaviour on the future occurrence of children's externalizing and internalizing behaviours. Specifically, in a U.S. sample of 594 African American and Hispanic families, Coley, Kull, and Carrano (2014) found that greater parental endorsement of physical punishment at four years of age was associated with a significant increase in externalizing and internalizing behaviours at age nine, after controlling for previous levels of behavioural problems. Moreover, children's behavioural problems at age four were not predictive of either concurrent or later physical punishment used by parents, thereby disconfirming the hypothesis of child-elicited effects on parental behaviour.

Physical punishment may also exert negative effects on children's cognitive functioning. Berlin et al. (2009) examined the impact of spanking and verbal punishment at one, two, and three years of age on cognitive development in a sample of 2,375 U.S. children from low-income families. After controlling for demographics, spanking at all ages was related to lower cognitive performance. In a

similar U.S. study, Straus and Paschall (2009) found that 1,250 2-4-year olds who experienced high-frequency spanking (defined as $\geq$ three episodes per week) exhibited significant lower cognitive performance at both 2-4 and 5-9 years when compared with children who experienced less frequent or no spanking, even after controlling for baseline cognitive abilities. Finally, in a large comprehensive study, MacKenzie, Nicklas, Waldfogel, and Brooks-Gunn (2013) examined the relationship between physical punishment at three years and cognitive outcomes at five and nine years of age. Findings indicated the presence of a significant effect of physical punishment on children's vocabulary skills, even after controlling for family-related risk factors and children's general verbal abilities. Moreover, the negative effect of physical punishment on verbal skills was constant across gender, ethnicity, and levels of parental warmth so these variables had no moderating influences.

The relationship between physical punishment and well-being extends beyond childhood and into the adult years. A nationally-representative U.S. study of 34,000 adults revealed that, after controlling for socio-demographics and family dysfunction (e.g., family member with alcohol/drug problems, family member with a history of mental illness), experiences of harsh physical punishment in childhood (defined as the use of force that goes beyond slapping and spanking but does not qualify as severe physical abuse) were associated with an increased risk for mood and anxiety disorders, alcohol/drug abuse, and personality disorders in adulthood (Afifi, Mota, Dasiewicz, MacMillan, & Sareen, 2012). In more recent research conducted with a U.S. sample of 8,613 adults, Afifi et al. (2017) found that childhood spanking was related to increased odds of serious mental health issues including suicide attempts, heavy drinking, and use of street drugs, after accounting for experiences of emotional and physical abuse. Based on these findings, the authors proposed that childhood physical punishment should be classified as an adverse childhood experience.

Physical Punishment and Family Violence

Research indicates that physical punishment can escalate to physical abuse, particularly in instances of increasing frequency and severity (Zolotor, Theodore, Chang, Berkoff, & Runyan, 2008). Physical punishment has also been linked with child welfare involvement. In a community-based U.S. sample of 2,788 families, spanking at the age of one year was associated with a 33% increased probability of subsequent child welfare involvement between the ages of one and five, even after accounting for other risk factors (e.g., maternal depression; Lee, Grogan-Kaylor, & Berger, 2014). Data from the Canadian Incidence Study of Reported Child Abuse and Neglect (CIS) conducted in 2008 suggest that, among 18,687 substantiated cases of physical abuse, about 74% were considered by the investigating child welfare worker to have occurred within the context of parental use of physical punishment (Jud & Trocmé, 2012). These results are in line with previous findings from an analysis of 102,866 substantiated cases of child maltreatment investigated in 2003 in which 75% were classified as incidents of physical punishment (Durrant et al., 2006). Finally, a retrospective study of 351 university students from Ontario found a statistically significant association between self-reported childhood physical punishment and physical abuse. In particular, individuals who reported experiencing physical punishment during childhood had a 60% increased risk of also experiencing physical abuse as children (Fréchette, Zoratti, & Romano, 2015).

Mechanisms Underlying the Impact of Physical Punishment on Children's Development

There are a variety of mechanisms that might help explain the link between physical punishment and developmental outcomes, particularly behavioural functioning. Harsh parenting practices (such as physical punishment) may lead to a child modelling aggressive behaviour as a way to manage conflict; over time, this may contribute to the development of oppositional and conduct problems (Deater-Deckard & Dodge, 1997). Physical punishment may also exert a cascading effect on children's behaviours, from aggression to the development of internalizing difficulties. Moreover, children who exhibit early aggressive behaviour may experience greater difficulties interacting with

peers and building friendships, which in turn may lead to withdrawal and isolation (Gromoske & Macguire-Jack, 2012).

Physical punishment could also negatively affect the quality of the child-parent relationship (Durrant & Ensom, 2020). For instance, children who are subject to physical punishment tend to associate the painful experience to the parent who is inflicting the pain and as such, may be more likely to avoid both the punitive action that causes physical pain as well as the parent. Such fear of experiencing pain might exacerbate the emotional distance and trust between the child and parent (Durrant & Ensom, 2020). In a similar vein, physical punishment has also been linked to poor attachment relationships. Coyl, Roggman, and Newland (2002) examined the relationship between disciplinary practices and attachment in a sample of 169 14-month-olds and their mothers attending a U.S. Head Start program. More frequent spanking was related to lower ratings of infant security as measured on the Attachment Q-set. There were similar findings in a study of 45 U.S. mothers and their preschoolers ($M = 4.5$ years), confirming the presence of significant direct inverse effects of physical punishment and attachment security, again as measured on the Attachment Q-set (Posada & Pratt, 2008). Moreover, there may be long lasting effects of physical punishment use on attachment relationships. Palmer and Hollin (2001) examined the effect of frequent physical punishment (i.e., how often the child has been spanked or hit by the parent) on adolescents' ratings of attachment security; results indicated a statistically negative association between physical punishment and self-reported attachment security in adolescence.

Use of Non-Physical Punishment: A Current Controversial Issue

Despite common agreement related to the negative consequences of physical punishment on children's well-being, claims about the detrimental effects of non-physical punishment have not yet reached a consensus among researchers (Holden, Grogan-Kaylor, Durrant, & Gershoff, 2017; Larzelere, Gunnoe, Roberts, & Ferguson 2017). In a recent article, Larzelere and colleagues (2017)

criticized “exclusively” positive parenting programs (EPP) that discourage the use of punishment (i.e., physical and non-physical) and focus on positive aspects of child-rearing by claiming that such an approach is fundamentally biased and that punitive strategies, if adopted appropriately, can be beneficial to children. To prove this point, the authors examined three major fallacies common to parenting research: (1) making causal claims about the effectiveness of EPP programs based primarily on correlations; (2) discounting causal evidence from randomized trials providing evidence of the effectiveness of the use of punishment with specific populations (e.g. children with behavioural problems); (3) basing conclusions on the assumption that that low use of punitive parenting strategies equate to zero usage and that only few studies have differentiated between spanked and non-spanked children; and (4) failing to discriminate the modality of employment of certain punitive techniques (i.e., “conditional spanking) and the circumstances in which they are used. Based on such claims, the authors suggest that current parenting research is biased in that it overgeneralizes scientific findings concerning the negative impact of punitive parenting on children’s outcomes, while neglecting the potential beneficial effects that such strategies might have on specific groups of children (e.g., children struggling with behavioural problems). Hence, they suggest that EPP programs should be scrutinized using more rigorous empirical testing and using different populations to reduce the risk of methodological fallacies.

In response to this critique, Holden and colleagues (2017) presented their counter-arguments by clarifying their definition of positive parenting and their position regarding current methodological fallacies in parenting research. The authors disagree with the EPP term used by Larzelere et al. (2017). They define “strong” positive parenting programs (as opposed to the term “EPP”, which is rather considered denigratory), those that shift away from the use of contingencies (i.e., non-physical punishments and rewards) and power-assertion techniques to reduce challenging behaviour, and instead favour a view of children deserving respect and love and devote increasing attention to relationship

building and conflict resolution strategies based on problem-solving. These parenting programs have the ultimate goal of helping children develop positive qualities, such as empathy, prosocial behaviour, and moral judgment, rather than focusing exclusively on short-term compliance and reduction of problematic behaviours. The authors also bring attention to some of the methodological fallacies noted by Larzelere et al. (2017) as an argument against positive parenting approaches by highlighting the growing evidence for the detrimental effects of physical punishment on children's well-being and the lack of consistent research on the positive effects of physical punishment on children's behaviour. Finally, Holden and colleagues (2017) complete their review by insisting on the importance of cultivating a culture of respect for children that begins in the family of origin and asserting the need for rigorous research into the effects of various forms of physical and non-physical punishment.

In the next section, some of the recent research on the effects of different types of non-physical punishment, including time-out, parental psychological control, and verbal aggression, will be presented.

Use of time-out. The most common type of non-physical punishment is time-out, which is a form of negative punishment that involves removing the child from the situation that caused the challenging behaviour and placing them in an environment that is “less” reinforcing for a set amount of time that is deemed reasonable (e.g., usually 2-5 minutes; Dadds & Tully, 2019). Time-out is presented as an effective strategy to deal with challenging behaviour by a number of popular evidence-based programs, including Triple P (Sanders, Kirby, Tellegen, & Day, 2014), Parent-Child Interaction Therapy (Querido, Bearss, & Eyberg, 2002), and Incredible Years (Webster-Stratton, 2001). Furthermore, this method has also been explicitly recommended by the American Academy of Pediatrics (2018) as a viable alternative to physical punishment.

Although there are claims about the effectiveness of time-out in addressing children's behaviour challenges (Kaminski & Claussen, 2017; Larzele et al., 2017), the use of time-out as an appropriate

(and effective) disciplinary strategy remains controversial. Results from an international study – involving China, India, Kenya, Philippines, and Thailand - with a sample of 292 families indicated that the frequency of use of time-outs was linked to greater internalizing symptoms among 8-12 years olds (Gershoff et al., 2010). Moreover, a recent review of 42 articles examining parenting programs using time-out revealed that this strategy is not taught in consistently across programs, therefore making it difficult to draw conclusions on the effectiveness of this parenting technique (Corralejo, Jensen, Greathouse, & Ward, 2018).

A recent RCT conducted using a sample of 320 children between four and nine years of age with behavioural difficulties examined the effects of two programs - an emotion coaching program (Tuning in to Kids; for more information, please see Havighurst & Harley, 2007) whose focus is to shift away from power assertion techniques (e.g., time-out) and rely on parental emotion self-regulation strategies and a behaviour management program (Triple P; Duncombe et al., 2016). These two programs were also compared against a wait-list control group. Results indicated that, compared to the wait-list control group, children in the emotion coaching group demonstrated a significant decrease in disruptive behaviours as reported by multiple informants (e.g., parents, teachers, and child). Moreover, both the behaviour management and emotion coaching programs had comparable effects on decreasing behavioural problems, therefore debunking the idea of the superior performance of behaviour management techniques such as time-outs (Duncombe et al., 2016). These findings also corroborated an earlier meta-analysis that compared behaviour and non-behaviour parenting programs and found that both approaches were equally effective in reducing child behavioural problems, improving parenting practices (i.e., increase positive parenting and decrease physical punishment), and changing parental perceptions (i.e., decreased parent-related stress and increased confidence in parenting; Lundhal, Risser, & Lovejoy, 2006). In addition, in a recent longitudinal study reporting no differences in the emotional and behavioural functioning of children whose parents used time-outs regularly relative to

those who did not engage in this parenting practice, Knight et al. (2020) concluded that there are neutral (non-detrimental) effects of time-outs on children's psychological well-being. However, the authors did not include positive outcomes in their study, therefore making it difficult to reach a definite conclusion on the presence of any beneficial effects of time-outs on children's behaviour. Furthermore, longitudinal studies on the prolonged positive effects of time-out are lacking, pointing out the need for additional research to determine the impact of this parenting technique on children's outcomes.

Parental psychological control. Besides time-out, other forms of non-physical punishment include shaming (i.e., criticism directed to the child), guilt-induction (i.e., the use of manipulative strategies enacted to induce guilt in the child), and love withdrawal (i.e., threatening to remove love, warmth, and attention from the child). Such behaviours are enacted by parents to pressure the child to comply with parental requests (Soenens & Vansteenkist, 2010; Shaefer, 1965). Such forms of psychological control are generally perceived as manipulative and intrusive, and they undermine the child's autonomy and self-esteem (Barber & Harmon, 2002). Parental psychological control has been associated with detrimental developmental outcomes in children and adolescents, and its long-term effects have been confirmed into young adulthood. For instance, in a U.S. study conducted with a clinical sample of 175 7-14-year olds, higher levels of self-reported maternal psychological control (e.g., guilt-induction, manipulation) were predictive of higher scores on parent-reported anxiety, depressive symptoms, and externalizing behaviours (Wei & Kendall, 2014). Similar findings were confirmed in another cross-sectional U.S. study conducted with a community-based sample of 585 adolescents examining the relationship between parental psychological control and problem behaviours. In this study (Symeou & Georgiou, 2017), paternal and maternal psychological control were the only significant predictors of increasing levels of both externalizing and internalizing problems after accounting for the effects of demographics and parental behavioural control (e.g., routine setting. Similarly, in an experimental study, 185 young adults (18-35 years) completed a series

of social stressor tasks that measured reactions to stressful performance-based conditions (e.g., job interview and arithmetic problems) and then a self-report measure of maternal control (Goger, Rozenman, & Gonzalez, 2020). Findings indicated that participants with high levels of perceived maternal control reported greater anxiety symptoms and poorer emotion regulation throughout the social stressor tasks, compared to participants with either average or low maternal psychological control. Furthermore, a recent U.S. study was conducted with 134 13-year-olds who were followed longitudinally into adulthood (age 32) to examine the long-lasting effects of parental psychological control on psycho-social adjustment. Adolescents who reported higher levels of parental control at age 13 demonstrated lower observed social support, lower likelihood of being involved in a romantic relationship, and lower academic achievement in adulthood (Loeb et al., 2020).

Parental verbal aggression. Parental verbal aggression includes measures of coercive control (e.g., yelling, shouting, insulting, and making verbal threats) that usually emerge as a result of parental poor judgment and impulsivity in response to stressful parenting situations (Straus & Field, 2003). A majority of studies examining the effect of punitive discipline on children have not distinguished verbal aggression from physical punishment, instead merging these two constructs into a broad dimension of “harsh parenting” (McKee et al., 2007). Verbal aggression can exert a negative impact on mental health outcomes beyond childhood and into adolescence. For instance, in a longitudinal U.S. study with a community-based sample of 976 families of young adolescents, mothers’ and fathers’ harsh verbal discipline at age 13 predicted an increase in adolescent conduct problems and depressive symptoms one year later (Wang & Kenny, 2014). Furthermore, a U.S. study conducted with a community sample of 2,518 emerging adults (18-25 years) found a significant negative impact of parental verbal aggression on measures of well-being (i.e., increased psychiatric symptoms) that was not mitigated by the presence of verbal affection, thereby indicating the independent effect of verbal aggression on mental health outcomes (Polcari, Rabi, Bolger, & Teicher, 2014). Generally, it has been reported that children

with Oppositional Defiant Disorder (ODD) tend to experience greater verbal aggression from parents compared with children who do not have behavioural challenges (Kolko, Dorn, Bukstein, & Burke, 2008). Similarly, in a study with 2,450 clinically referred girls (5-16 years) with ODD symptoms, both parent and child self-report ratings of parental verbal aggression predicted increases in ODD symptoms after accounting for demographics, prior behavioural problems and physical punishment (Derella, Burke, Stepp, & Hipwell, 2020).

Determinants of Physical Punishment Use

Attitudes toward physical punishment. There is growing research on the variables that impact parental use of physical punishment and that may, therefore, be a focus of efforts aimed at eliminating physical punishment use. Attitudes of physical punishment play a key role in determining the type of strategies parents are likely to use (Holden et al., 1995; Vittrup, Holden, & Buck, 2006). Ateah and Durrant (2005) identified four variables that were the most predictive of parental physical punishment use, namely 1) approval of physical punishment use, 2) attributions about a child's transgression, 3) perceived seriousness of a child's transgression, and 4) anger in response to the transgression. Together, these four variables (which, except for parental anger, are classifiable as parental cognitions) accounted for 54% of the variance in parents' physical punishment use. Approval of physical punishment emerged as the strongest distal predictor, which included other variables such as a mother's own childhood history of physical punishment, her knowledge of child development, and her repertoire of positive disciplinary strategies (Ateah & Durrant, 2005). Moreover, Perron et al. (2014) conducted a study of 2,340 Canadian parents of children aged two to 12 years and found that, after controlling for child- and family-related factors, parents' favourable attitudes toward the use of physical punishment was the strongest predictor of physical punishment use. In addition, parents who held pro-spanking attitudes were nearly five times more likely to also endorse the use of physical punishment.

Findings from another Canadian study conducted with a representative sample of 1,000 families indicated that parents who experienced physical punishment in childhood also tended to believe that physical punishment is not likely to cause physical injuries, and they were more likely to be in favour of using physical punishment as a disciplinary strategy (Gagné et al., 2007). Finally, Romano, Bell, and Norian (2013) surveyed 212 community-based individuals (25% of whom were parents) to examine predictors of change in attitudes toward physical punishment. Individuals rated their degree of agreement with Section 43 of the Criminal Code, which gives parents the right to use “reasonable” force for disciplinary purposes. They were then provided with information about negative child effects of physical punishment and about alternative disciplinary practices. When individuals were then asked to reconsider Section 43 in light of this information, there was a significant shift in parental attitudes. Between 68.3%-70.1% of participants were in favour of ending Section 43, compared with 37.7% at the initial assessment. Such findings show that providing information on the potential harmful effects of physical punishment and on the availability of alternatives can influence attitudinal change toward physical punishment.

Attributions about children’s behaviours. Among other proximal factors that have been examined in relation to physical punishment use is parental hostile attributions of children’s behaviour, defined as the inferences that parents draw from children’s behaviour that can affect their responding to the child (Bugental, Johnston, New, & Silvester, 1998). Parental dysfunctional child-centred attributions preserve a negative view of children and hold them responsible for transgressions. For instance, parents might believe that children misbehave to make their parents mad; such attributions would most likely have a negative impact on how parents react to the child’s misbehaviour (Dix & Grusec, 1985). Such dysfunctional attributions have been shown to relate to harsh parenting (Smith & O’Leary, 1995; Wilson, Gardner, Burton, & Leung, 2007) and an increased likelihood of physical punishment use (Ateah & Durrant, 2005).

In one of the few experimental studies, Slep and O' Leary (1998) manipulated the dysfunctionality of parental attributions by providing parents with feedback scripts targeting a child's responsibility for their own behaviour. The group of parents receiving feedback high in child responsibility were told that the child was in control of their own behaviour and actions (i.e., children cry to get their own way). Direct observations of parenting behaviour took place in a testing room. Compared to the control group of parents who received low child-responsibility feedback, the experimental group manifested greater use of harsh parenting interacting with their children (i.e., over-reactive discipline, angry responses). Such findings highlight the role of parental attributions in guiding disciplinary practices.

In a more recent U.S. randomized intervention study of 731 mothers, negative relational schemata (i.e., beliefs that the child purposely acted out to provoke the parent) assessed through speech samples at baseline predicted an increase in coercive interactions assessed observationally through age 2-4 years, which in turn predicted increasing oppositional and aggressive behaviours at age 7-8 after controlling for baseline levels of behavioural problems, positive relational schemas, child gender, child ethnicity, and cumulative risk (Smith, Dishion, Shaw, & Wilson, 2015). These results were further corroborated by a recent investigation examining the role of socio-cognitive factors on the use of physical punishment in a U.S. sample of mothers and fathers (Rodriguez & Wittig, 2019). Self-report and observational measures of parent-child aggression and attributions of child behaviour were obtained when the child was an infant and were combined into composite scores that were then used as predictors of physical punishment in toddlerhood. Results showed a significant impact of maternal approval of parent-child aggression and negative attributions of child behaviour on later reported use of physical punishment (Rodriguez & Wittig, 2019).

Parenting self-efficacy. Parenting self-efficacy refers to beliefs concerning one's own abilities to manage the daily tasks of parenthood and a sense of competence in one's parenting role (Ardelt &

Eccles, 2001; Jones & Prinz, 2005). Low parenting self-efficacy has been related to frequent use of physical punishment (Khoury-Kassabri, Attar-Schwartz, & Zur, 2014; Rominov, Giallo, & Whelan, 2016). On the contrary, high levels of parenting self-efficacy have consistently been associated with competent and positive parenting practices (Hamovitch, Acri, & Bornheimer, 2019; Hill & Bush, 2001). For instance, results from a large-scale study with a sample of 2,477 Arab mothers indicate that lower maternal self-efficacy was related to an increased likelihood of using physical punishment with their 4-5 year old children; this, in turn, was related to greater child behavioural and emotional difficulties (Khoury-Kassabri et al., 2014). In a similar study conducted with a U.S. sample of 213 parents of children diagnosed with ODD (age 7-10), high levels of parenting self-efficacy were associated with high levels of positive parenting (e.g., praising the child for good behaviour) and inversely related to inconsistent discipline (Hamovitch et al., 2019). Finally, in a longitudinal sample of 3,741 Australian fathers, lower parenting self-efficacy reported after the birth of a child predicted higher levels of hostile parenting and lower parenting consistency when the child was 4–5 years; in turn, these were associated with emotional-behavioural problems at 8–9 years (Rominov, et al., 2016).

Parenting stress. Parenting stress can be defined as “a set of processes that lead to aversive psychological and physiological reactions arising from attempts to adapt to the demands of parenthood (Deater-Deckard, 2004, pg. 6).” Parenting stress appears to exert powerful influence on parenting behaviour, contributing to child-rearing difficulties. In a large-scale U.S. study of 2,508 mothers, Taylor et al. (2009) found that higher levels of maternal stress were associated with an increased risk of using physical and psychological aggression, as well as neglect. These effects were significant even after controlling for exposure to IPV, maternal depression and demographic variables (Taylor et al., 2009). Moreover, results from a recent longitudinal U.S. study of 1, 281 mothers indicated the presence of a significant relationship between increases in parenting stress from infancy to toddlerhood and increases in maternal aggression at age three in mothers classified at low and high risk of child

maltreatment (Prendergast & McPhee, 2020). However, there is a scarcity of research on the effects of parenting stress on the use of physical punishment among parents in North America. In lower income countries however, research shows a significant link between parenting stress and use of punitive child rearing strategies. For example, a relatively recent cross-sectional study on 683 Chinese parents of 3-6-year-old children found the presence of significant direct effects of parenting stress on harsh punishment (defined as the use of both psychological aggression and physical punishment; Liu & Wang, 2015). Similar findings were confirmed in a longitudinal study conducted with 159 of Turkish mothers of young children (mean age 7). In this sample, high levels of mother report of parenting stress obtained when the child was four years old were associated to greater use of physical punishment at six years, which in turn exerted negative indirect effects on child behaviour at age seven, leading to an increase in child aggression and a decrease in prosocial behaviour (Gülseven et al., 2018)

Problem-solving and disciplinary strategies. Parental problem-solving involves the ability to identify challenging situations, generate ideas on how to solve these challenges, design an action plan, and revise the plan based on child outcomes (Sanders, Turner, & Metzler, 2019). Generally, parents with strong problem-solving skills are capable of discerning different aspects of a problem that arise in a parenting situation, providing possible solutions, and gauging each option before adopting a definitive strategy. Conversely, parents with limited problem-solving abilities might have difficulties in responding adaptively to caregiving challenges and may be more likely to experience failure and to manifest greater aggression toward their child (Azar, Robinson, Hekimian, & Twentyman, 1984).

A growing body of research has investigated how problem-solving is related to parenting behaviour. In two separate studies, physically maltreating mothers exhibited poorer problem-solving abilities compared to non-abusive mothers (Azar et al., 1984; Azar, Stevenson & Johson, 2012). Moreover, low problem-solving ability related to problematic parent interactions characterized by lower maternal warmth and responsiveness as well as higher dyadic intrusiveness, disengagement, and

conflict (Azar, Stevenson, & John, 2012). In a recent U.S. study of 67 mothers of 3 to 5-year olds at risk of child maltreatment, Azar et al. (2016) examined the relationship between parental self-reported problem-solving skills and the frequency and severity of physically abusive incidents. In this at-risk sample, weaker problem-solving abilities (but not severity of physical abuse which it was related at a trend level with hostile attributions) were associated with greater frequency of physical abuse (Azar et al., 2016).

Parental emotion regulation. Parenting generally elicits a wide range of positive and negative emotions. The type of behaviour parents display in response to challenging child-rearing situations depends not only on the intensity of the behaviour displayed by children (i.e., child-elicited effects) but also one's emotion self-regulation abilities. Emotion self-regulation refers to strategies used to control, modulate, and modify emotional reactions in response to changing environmental demands (Gross & Thompson, 2007). Parental emotion regulation relates to different aspects of parenting, including monitoring and interpreting emotional cues, as well as controlling difficult emotions and modulating their expression during interactions with children (Fabrizio et al., 2015). A multi-method assessment study using self- and informant-reports, psychophysiological assessments, and behavioural observations with 119 mothers and their toddlers indicated that clusters of mothers characterized by emotion dysregulation (i.e., high level of "negative" emotions during parent-child interactions, increased vagal tone, and low self-reported emotion regulation) were more likely to exhibit maladaptive parenting patterns (i.e., low structure, low dyadic collaboration, and low warmth; Shaefer et al., 2018). In a recent experimental U.S. study with 190 parents of young adolescents (12-14 years) indicated that lower scores on objective measures of parents' emotion regulation (i.e., vagal activity and heart rate) during a conflict resolution task were related to adolescent-reported ratings of parental hostile behaviour at 1-year follow-up (Sturge-Apple et al., 2020). This study corroborates previous findings from a U.S. study with a sample of mothers and their toddlers during the Strange Situation

paradigm (i.e., laboratory task assessing parent-child separation anxiety). In this sample, mothers exhibiting lower emotional self-regulation (i.e. heightened arousal in response to distress elicited by the experimental condition) also displayed higher levels of harsh, hostile, and intrusive parenting (Sturge-Apple, Skibo, Rogosh, Ignajtovic, & Heinzelman, 2011).

The Positive Discipline in Everyday Parenting Program

In support of emerging global initiatives that promote children's rights, Save the Children (a non-governmental organization that delivers programs to improve children's lives around the globe) initiated a project aiming at eliminating child physical punishment. As part of this project, Joan Durrant developed the Positive Discipline in Everyday Parenting (PDEP; Durrant 2007, 2013) program. PDEP is a group-based parenting program aimed at enhancing general knowledge of child development and promoting the use of alternative disciplinary measures based on non-violent confrontation and sensitive responding. PDEP focuses on helping parents understand the developmental reasons behind children's behaviour so that parents can shift away from behaviour management strategies (i.e., rewards and punishments) toward communication and problem-solving. From this viewpoint, parent-child conflict is reduced through enhancing the relationship, strengthening trust and attachment, and building a strong foundation for communication throughout the child's development (Durrant 2007, 2013).

According to the conceptual model of the PDEP program (see Figure 1), the program targets two main components of parental behaviour: 1) cognitive/affective processes that includes attitudes, subjective norms, and parenting self-efficacy, and 2) action goals that refer to parenting stress and problem-solving skills. Over eight 2-hour sessions, participating parents take part in a series of interactive activities that help them understand their children's development from birth to adolescence. PDEP aims to help parents recognize the key developmental themes that underlie common parent-child conflicts so that they become less likely to attribute them to intentional "misbehaviour" on the part of the child. The program activities are sequenced to gradually increase parents' ability to take the child's

perspective, thereby reducing frustration and anger. Parents also learn about the brain's stress response and how it can escalate conflict to result in punishment, coercion, and violence (Durrant, 2007; 2013). Throughout the program, parents gradually learn a problem-solving framework that they can apply in a wide range of conflict situations. First, they think about their long-term parenting goals to refocus their attention from winning the conflict to increasing their understanding about the developmental reasons that could underlie the child's behaviour, as well as the reasons for their own behaviour. Second, parents think about how they can provide warmth (physical and emotional safety) and structure (information and guidance) to resolve the conflict constructively. Next, they also learn about child development, focusing on the important milestones and various stages of growth from birth to adolescence. They practice applying this framework to a range of hypothetical situations and to situations from their own experience. The expected outcomes of this process are a reduction in parents' approval of physical and other punishments, a decrease in negative child-centred attributions, increased self-efficacy in engaging in positive discipline, increased emotional regulation, a greater likelihood of problem-solving rather than punishment in response to conflict, and a lower level of parenting-related stress (Durrant, 2007; 2013).

PDEP is delivered on a non-profit basis through community agencies by trained program facilitators to groups of five to 20 parents. Facilitators are trained by either PDEP master trainers (five experts in child development, parent support, and group facilitation) or PDEP country trainers (experienced PDEP facilitators who have taken a 4-day advanced training delivered by a master trainer, followed by a mentorship period). The facilitator training focuses on both mastery of program content and group facilitation skills (Durrant, personal communication, September 19, 2016).

As a universal program, PDEP targets all parents regardless of socio-demographic background or child age. PDEP has grown rapidly in response to a high demand for facilitator training by agencies across Canada and around the world. Pre- and post-program data are routinely collected; findings from

321 parents living in 14 Canadian cities demonstrated that, over the course of the program, parents became significantly less supportive of physical punishment, less likely to attribute typical child behaviour to misbehaviour, and more likely to believe that they have the parenting skills they need. Moreover, 90% of the parents believed more strongly that parents should not use physical punishment and that PDEP could help them build stronger relationships with their children (Durrant et al., 2014). Moreover, the authors found that more than 90% of parents believed that PDEP would help them control their anger and understand their children's feelings, and 82% believed that they could solve most of their parenting challenges (Durrant et al., 2014). Similar results have been found among 516 parents who completed standardized pre-to-post program assessment (Durrant et al., 2018). Parents responded to several questions about the impact of the program on their parenting. The large majority of participants agreed that PDEP helped them use less physical punishment (80%), control their anger (80%), communicate better with their children (90%), build stronger relationships with their children (95%), understand children's development (91%), and understand children's feelings (90%; Durrant et al., 2018). In addition, a recent qualitative study on a small sample of Canadian parents ($n = 7$) who completed PDEP and program facilitators ($n = 4$) indicated an appreciation of the program's parenting approach. Parents also indicated that PDEP helped to improve their knowledge of child development, re-evaluate challenging parenting situations in light of the child's developmental stage, and reduce the likelihood of using physical punishment (Stenason, Moorman, & Romano, 2020).

Similar results have been confirmed in studies conducted with parents from different countries, including parents in Japan ($n = 77$; Mori, Mochizuki, & Stewart-Tufescu, 2016) and Kosovo ($n = 55$; Ademi-Shala, Hoxha, & Ateah, 2016). Following completion of the PDEP program, parents were less favourable toward the use of physical and emotional punishment. Also, they rated the PDEP program as being useful in helping establish a positive relationship with their children as well as in increasing their confidence in using positive discipline (Ademi-Shala et al., 2016; Mori et al., 2016).

In sum, the above-mentioned findings provide preliminary support for the positive impact of PDEP on parental attitudes and behaviours. Although the research base on PDEP outcomes is expanding, the studies that have been conducted to date have not made use of an experimental design that includes a control group and randomization of participants to either the experimental or control group. Without such controls, it cannot be concluded with certainty that any parenting changes are due solely to participation in the PDEP program. A second limitation of existing PDEP outcome research is the lack of follow-up data, without which it cannot be determined whether any changes at post-program are maintained over a longer period of time.

Research Questions and Hypotheses

To address the limitations of previous PDEP research, a quasi-experimental study was conducted, with the inclusion of a wait-list control. The first research question examined the impact of the PDEP program on parental attitudes toward the use of physical and emotional punishment, parental attributions of child behaviour, and parenting self-efficacy. Compared with the wait-list control group, it was hypothesized that parents who completed PDEP would exhibit improvements in cognitive/affective processes following completion of the program including: 1) lower levels of physical punishment approval; 2) lower levels of non-physical (emotional) punishment approval; 3) be less likely to view children's "misbehaviour" as intentional noncompliance and more likely to view it as reflecting developmental transitions; 4) higher perceptions of themselves as competent in providing positive discipline. It was also expected that the PDEP parents (compared to non-PDEP parents) would exhibit further improvement in cognitive/affective processes at the 1-month follow-up assessment.

The second research question was around whether participation in the PDEP program would result in better action goals (e.g., parenting stress level, problem-solving abilities, and emotion self-regulation). It was hypothesized that PDEP parents would report lower levels of parenting stress (i.e., in terms of frequency of parenting stress and intensity of stress related to parenting tasks and child

behaviour) at post-program, compared with parents in the wait-list control group. It was also expected that PDEP parents would exhibit further improvements at 1-month follow-up. Moreover, it was hypothesized that parents in the PDEP group (relative to non-PDEP parents) would report an increase in their tendency to approach problems with their children (rather than avoid them). In addition, it was hypothesized that parents who completed PDEP (relative to non-PDEP parents) would exhibit greater emotional self-regulation (e.g., lower impulsivity, greater awareness, and better strategies to deal with negative emotions) following their participation in parenting program and would exhibit further improvements in these areas at the 1-month follow-up assessment.

The third research question addressed whether the PDEP program would lead to changes in parenting behaviours. Compared to the wait-list control group, parents who completed PDEP were expected to report 1) decreased use of physical punishment, 2) decreased use of non-physical (e.g., emotional) punishment, and 3) greater use of positive parenting strategies at post-program. It was also expected that the PDEP parents (compared to non-PDEP parents) would exhibit further improvement in parenting practices at the 1-month follow-up assessment.

Method

Participants and Procedures

A sample of 83 parents were recruited from nine collaborating community-based agencies in Ottawa and surrounding regions between Fall 2018 and Winter 2019. The first column of Table 28 presents the demographic characteristics for the entire sample. Parents were predominantly female (84.1%) and White (73.5%). The participants' mean age was 36.3 years ($SD = 7.3$), and the majority (70.70%) were married. The number of children in the home ranged from 1-5, with a mean of 1.9 ($SD = 0.8$) and a mean child age of 2.0 years ($SD = 0.9$). In terms of education level, the majority of participants (87.8%) had obtained a post-secondary degree. With respect to partner's education, the majority (where applicable) had obtained a post-secondary degree (74.7%). For employment, the

majority of participants (63.5%) and their partners (91.5%) were employed outside the home. In terms of household income, half of the participants (50.8%) had incomes ranging from \$50,000 and \$119,999. The second and third columns of Table 28 present socio-demographic characteristics separately for the experimental ($n = 47$) and wait-list control ($n = 36$) groups. Across all variables, there were no statistically significant differences, indicating that the two groups were comparable in terms of socio-demographic characteristics.

In terms of arriving at the final sample, Figure 10 shows that there were initially 157 parents who expressed an interest in learning more about the research study. Of these, 43 (27.4%) were excluded for various reasons, most notably because they could not be reached after a number of contact attempts. Among the 114 parents who were contacted, 10 (8.8%) were excluded either because the child was too young to be included in the research study or they did not have children of their own. Of the 104 parents who agreed to participate in the study, 9 (8.6%) were excluded because they did not complete the pre-program questionnaire. Among the 90 eligible participants, 38 parents were assigned to the experimental group and 55 to the wait-list control (either through randomization or in compliance with a parent's preference). However, seven (7.8%) were not included in the study because they did not complete the questionnaire at either the post-PDEP data collection or the 1-month follow-up (i.e., there were only pre-program questionnaire data). These various exclusions resulted in a final sample of 83 parents. At post-PDEP, 78 (93.8%) parents had completed the questionnaire, and 76 (91.6%) had completed the questionnaire at 1-month follow-up. Of all participants, 71 (85.5%) had complete across all the three time points but the analyses included 83 parents as HLM uses the full maximum likelihood (see Data Analysis section). Of the final sample of 83 parents, 22 (26.5%) had been randomized to either the experimental ($n = 8$) or wait-list control ($n = 14$) group. The remaining parents were placed in the group of their preference, which was largely determined by the timing of the group (e.g., either fall or winter).

The agencies that were involved in this study were community-based resource centres that provide programming for young children and families. To comply with the experimental research design, each agency offered two PDEP groups, the first of which constituted the experimental group (Fall 2018) and the second of which was the wait-list control group (Winter 2019). After receiving ethics approval from the Research Ethics Board at the University of Ottawa (REB # H-05-18-630, approved June 2018, Appendix O), recruitment posters were provided to participating agencies for distribution to families who were considering participation in the PDEP program (see Appendix P).

Parents who were interested in learning more about the research study could either contact the researchers at the University of Ottawa directly or ask the PDEP facilitator to provide their contact information to the researchers. A research assistant then contacted parents by phone to provide additional information about the study and ensure that eligibility criteria were met (see Recruitment notice in Appendix Q). Eligibility criteria were the following: 1) being a parent of a 2-6-year old; 2) being able to comprehend English as the questionnaire was only available in this language; and 3) not participating in another parenting program while completing (or waiting to complete) PDEP. The research assistant also provided an overview of the research project, reviewed the limits to confidentiality, answered parents' questions, and gauged whether parents would agree to randomization procedures. Note that only a small number of parents ($n = 22$) agreed to be randomly assigned to either the PDEP experimental group or wait-list control group; the majority expressed a preference for a specific group. This randomization strategy was utilized at a within-agency level and not for parents across all participating agencies to maintain adequate sample size for the two groups that were offered at each agency.

For parents who agreed to participate in the research study, they were sent an e-mail inviting them to complete the pre-program questionnaire by clicking on a web-link. When parents clicked on the link, they were directed to a consent form (Appendix Q). They were instructed to read the consent

form and, if desired, to click on a box indicating their consent. At this point, parents were directed to the questionnaire that was available through the Qualtrics survey website. As a measure of identity protection, parents could not save their responses and return to the questionnaire at a later date; however, they were provided with an 8-hour window during which to complete the questionnaire. As an additional measure to protect anonymity, parents were not asked for any identifying information. Instead, they developed a 6-digit code based on their date of birth and the first two letters of their first name. Parents were instructed to provide the same code at each data collection (pre-program, post-program, and 1-month follow-up) so that their responses could be linked across time. Following completion of the online questionnaire, participants were directed to another window in order to provide their address (email or home) so that the research assistant could send a \$50 gift card. Once parents had completed the pre-program questionnaire, they were contacted by email to inform them of the PDEP group to which they were assigned, either by random assignment for those who agreed to this procedure or as confirmation of the specific group they requested.

For parents in the experimental group (Fall 2018), PDEP attendance across the eight program sessions was recorded by a trained research assistant at each participating agency. These trained research assistants also were present at all PDEP sessions to complete a fidelity checklist to assess facilitators' adherence to the core content that was to be covered across each of the PDEP sessions (Appendix R). Within one week following the completion of the first PDEP program (Fall 2018), all parents in both the experimental and wait-list control groups were contacted by email to complete the online questionnaire another time, following which they were provided with a \$50 gift card using the same procedure outlined above. Finally, one month after the end of the PDEP group, all parents (experimental and wait-list) were emailed and asked to complete the online questionnaire, following which they were provided with a \$50 gift card. Shortly following the 1-month follow-up data collection, parents in the wait-list control group began the PDEP program (Winter 2019).

Measures

All questionnaires (with the exception of socio-demographics) were completed at three data collection points, namely prior to the start of PDEP, post-PDEP, and 1-month follow-up.

Socio-demographics. Parents answered questions on the following: gender and age of the child on whom they were focusing in responding to the questionnaires; their own age and gender; number of children in the home; marital status; ethnicity; household income; employment status (their own and that of their partner, if applicable); and highest level of education (their own and that of their partner, if applicable). Parents were also asked to indicate whether they had attended other parenting programs and, if so, to identify the program among a selection of options. Note that this question about program attendance was asked at the three data collections (Appendix S).

Cognitive/affective processes. These measures were used to measure cognitive/affective processes described in the conceptual model of the PDEP program (see Figure 1).

PDEP questionnaire (revised). The revised version of the PDEP questionnaire (based on findings from Study 2 of my thesis) was used to examine parental approval of physical and emotional punishment, subjective norms, and parenting self-efficacy. The Approval of Physical Punishment scale includes nine items (e.g., Spanking is fine as long as the parent is not angry) rated along a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*). Total scores can range from 9-54, with higher scores indicating a stronger approval of use of physical punishment. In the current sample of 83 parents, this scale had good internal consistency at pre-program ($\alpha = .84$), post-program ($\alpha = .85$), and 1-month follow-up ($\alpha = .87$).

The Approval of Non-Physical Punishment scale includes seven items on parental attitudes toward non-physical forms of punishment (i.e., time-out, shaming). Each item (e.g., If a 3-year-old won't share his toys, he should be made to sit on a chair for 3 minutes) is rated along a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*). Total scores can range from 7-42, with higher

scores indicating a stronger endorsement of non-physical punishment. In the current sample of 83 parents, this scale was found to have acceptable internal consistency at pre-program ($\alpha = .71$), post-program ($\alpha = .70$), and 1-month follow-up ($\alpha = .71$).

The Subjective Norms scale includes nine items that measure parents' beliefs regarding children's behaviours. Participants were asked to indicate their agreement with statements attributing typical child behaviour to 'misbehaviour/non-compliance' on the part of the child (e.g., A 6-year-old who won't show affection to his grandmother is being rude). Parents respond to each item on a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*). Total scores can range from 9-54, with higher scores indicating a stronger belief that parent-child conflict is due to 'misbehaviour.' In the current sample of 83 parents, this scale had acceptable to good internal consistency at pre-program ($\alpha = .80$), post-program ($\alpha = .76$), and 1-month follow-up ($\alpha = .83$).

Finally, the Parenting Self-Efficacy scale includes eight items that measure the degree to which parents feel competent and confident in their parenting abilities, as well as in their ability to manage challenging situation without incurring their anger and use punishment. Each item (e.g., I can control my anger in challenging situations with this child) is rated along a 5-point Likert scale from 1 (*never*) to 5 (*always*). Total scores can range from 8-40, with higher scores indicating greater perceived self-efficacy. In the current sample of 83 parents, this scale had good internal consistency at pre-program ($\alpha = .87$), post-program ($\alpha = .90$), and 1-month follow-up ($\alpha = .84$).

Parental Emotion Regulation. The Difficulties in Emotion Regulation Scale - Short Form (DERS-SF; Kaufman, et al., 2016) is an 18-item emotion regulation measure. Each item is rated on a 5-point scale ranging from 1 (*almost never 0-10%*) to 5 (*almost always 91-100%*), with higher scores representing higher frequency of use of a particular self-regulation strategy. The measure is composed of six subscales: 1) Impulse (3 items; e.g. When I'm upset, I lose control over my behaviours) assesses difficulties in controlling "negative" emotions; 2) Strategies (3 items; e.g. When I'm upset, I believe

that I will remain that way for a long time) measures difficulties in finding strategies to regulate “negative” emotions; 3) Non-acceptance (3 items; e.g., When I’m upset, I feel guilty for feeling that way) assesses a person’s sense of guilt in response to their experience with “negative” emotions; 4) Goals (3 items; When I’m upset, I have difficulty concentrating) assesses difficulties in focusing while experiencing “negative” emotions; 5) Clarity (3 items; e.g., I am confused about how I feel) measures a person’s sense of clarity with respect to their emotions; and 6) Awareness (3 items; e.g. When I’m upset, I acknowledge my emotions) measures a person’s ability to acknowledge their emotions. The subscales have shown good to excellent internal consistency in two adult samples (range $\alpha = .79-.91$; Kaufman et al., 2016). In this study, the Impulse, Strategies, and Awareness subscales were used because they are aligned with the PDEP program goals (Appendix T). The score within each scale can range from three to 15. In the current sample, the Impulse (pre-PDEP $\alpha = .92$, post-PDEP $\alpha = .86$, and 1-month follow-up $\alpha = .82$), Awareness (pre-PDEP $\alpha = .89$, post-PDEP $\alpha = .88$, and 1-month follow-up $\alpha = .74$), and Strategies (pre-PDEP $\alpha = .80$, post-PDEP $\alpha = .73$, and 1-month follow-up $\alpha = .92$) were found to have acceptable to excellent internal consistency at all three time points.

Action goals. These measures were used to examine action goals described in the conceptual model of the PDEP program.

Parenting stress. The Parenting Daily Hassles Scale (Crnic & Greenberg, 1990) asks parents to rate the frequency of 20 everyday events associated with parenting preschoolers, as well as how much impact (i.e., intensity) each of these events has had on the parent over the past month (Appendix U). Each frequency item (e.g., Difficulties in leaving kids for a night out or at school or day care) is rated along a 4-point scale from 1 (*rarely*) to 4 (*constantly*). Total scores can range from 20 to 80, with higher scores indicating greater occurrence of the challenging parenting event. The frequency scale demonstrated good internal consistency in the original sample of mothers ($\alpha = .81$; Crnic & Greenberg,

1990). The internal consistency of the frequency scale in the current sample of 83 parents was adequate at pre-program ($\alpha = .67$), post-program ($\alpha = .83$), and 1-month follow-up ($\alpha = .72$).

A subset of 15 out of the 20 items composed two subscales: Parenting Task and Challenging Behaviours. Five of the 20 items were excluded because they served as distractors. Each item is rated along a 5-point scale from 1 (*low*) to 5 (*high*). The Parenting Task is composed of eight items depicting parenting tasks that could potentially generate stress (e.g., Continually cleaning up kids' messes; Difficulties getting kids ready for outings on time). The total score can range from eight to 40, with higher scores indicating a higher intensity of stress generated by these parenting tasks. This subscale demonstrated good internal consistency in the original sample of mothers ($\alpha = .86$; Crnic & Greenberg, 1990). In the current sample of 83 parents, the Parenting Task was found to have adequate internal consistency at pre-program ($\alpha = .75$), post-program ($\alpha = .76$), and 1-month follow-up ($\alpha = .72$). The Challenging Behaviour subscale is composed of seven items used to assess the stress generated by a child's challenging behaviour (e.g., Kids don't listen, won't do what they are asked without being nagged; Kids are difficult to manage in public places). Total scores can range from seven to 35, with higher scores indicating a higher intensity of stress arising from child behaviours. The Challenging Behaviour subscale score demonstrated good internal consistency in the original sample of mothers ($\alpha = .86$; Crnic & Greenberg, 1990). In the current sample of 83 parents, the subscale had adequate to good internal consistency at pre-program ($\alpha = .72$), post-program ($\alpha = .84$), and 1-month follow-up ($\alpha = .76$).

Problem-solving. The Personal Problem-Solving Inventory (PSI; Heppner & Petersen, 1982) is a 32-item measure assessing individuals' personal evaluation of their problem-solving abilities. Each item is rated on a 6-point Likert scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). The measure is divided in three main subscales, namely: 1) Confidence is composed of 11 items (e.g., I trust my ability to solve new and difficult problems) assessing the individuals' trust in their own

problem-solving abilities; 2) Approach/Avoidance is composed of 16 items (e.g., When I have a problem, I consistently examine my feelings to find out what is going on in the situation) measuring an individual's tendency to rely on approach (i.e., actively seeking for a solution of the problem) or avoidance (i.e., escape, avert a problem) strategies in managing problems; 3) Personal Control is composed of 5 items (e.g., I make snap judgments and later regret them) measuring self-control abilities. In this study, the Approach/Avoidance subscale was used in that it fits the PDEP program goal of actively using problem-solving strategies to resolve conflict (Appendix V). Total scores can range from 16-96, with a higher score indicating a greater tendency to approach a problem to find a solution. This subscale demonstrated good reliability in the original sample of undergraduate students ($\alpha = .84$; Heppner & Petersen, 1982). In the current sample of 83 parents, this scale showed good internal consistency at pre-program ($\alpha = .80$), post-program ($\alpha = .82$), and 1-month follow-up ($\alpha = .79$).

Parental Behaviour. The next two measures were employed to examine different aspects of parenting behaviours.

Parenting strategies. This scale was designed by Dr. Joan Durrant in collaboration with members of her research team to evaluate parents' use of physical punishment and other parenting behaviours (e.g., reasoning, ignoring, diverting, time-out). Each item is rated along a 5-point Likert scale ranging from 1 (*never*) to 5 (*more than once a day*). Items are organized into three subscales. The first subscale includes four items that measure Use of Physical Punishment (e.g., Hit, spanked, or slapped this child; Shook this child). Total scores can range from four to 20, with higher scores reflecting more frequent use of physical punishment. In the current sample of 83 parents, this scale was found to have poor internal consistency at pre-program ($\alpha = .44$), post-program ($\alpha = .49$), and 1-month follow-up ($\alpha = .43$). The second subscale includes four items that measure Use of Non-Physical Punishment (e.g., Made this child sit alone, such as in a corner or in another room; Took away a

favourite activity, like watching TV or seeing friends). Total scores can range from four to 20, with higher scores reflecting more frequent use of non-physical (e.g., emotional) punishment. In the current sample of 83 parents, this scale was found to have adequate to good internal consistency at pre-program ($\alpha = .70$), post-program ($\alpha = .73$), and 1-month follow-up ($\alpha = .85$); 3) The third subscale includes nine items that measure Use of Positive Discipline (e.g., Explain why this child's behaviour was a problem; Talked calmly with this child to find a solution together). Total scores can range from nine to 45, with higher scores reflecting more frequent use of positive disciplinary strategies. In the current sample of 83 parents, this scale was found to have adequate to good internal consistency at pre-program ($\alpha = .81$), post-program ($\alpha = .75$), and 1-month follow-up ($\alpha = .76$).

Positive parenting strategies. Sixteen items (out of the 21) of the Parenting Young Children questionnaire (McEachern et al., 2012) were used to assess dimensions of positive parenting strategies that related directly to the main goals of the PDEP program (Appendix W). Each item is rated on a 7-point Likert scale from 1 (*not at all*) to 7 (*most of the time*) indicating the frequency of the parenting behaviour over the past month. After several discussions with the research team, a decision was made to not use five items that made specific references to punishment/rewards and/or the term “rules” as these items are not reflective of PDEP program goals. Items deleted included the following: 1) Reward your child when s/he did something well or showed a new skill; 2) Stick to your rules and not change your mind; 3) Set rules on your child's problem behaviour that you were willing/able to enforce; 4) Make sure your child followed the rules you set all or most of the time; and 5) Avoid struggles with your child by giving clear choices.

This measure is composed of three subscales. The first is called Supportive Positive Behaviour (six items; e.g., Stand back and let your child work through problems s/he might be able to solve), and it measures the frequency of positive parenting behaviours, such as praising the child, teaching the child new skills, and playing with the child. Total scores can range from six to 42, with a higher score

indicating higher frequency of use of supportive parenting strategies. In the current sample of 83 parents, this scale was found to have adequate internal consistency at pre-program ($\alpha = .75$), post-program ($\alpha = .75$), and 1-month follow-up ($\alpha = .75$). The second subscale is called Proactive Parenting (six items; e.g., Plan ways to prevent problem behaviour, such as feeding your child before going to the store), and it measures parenting behaviours used to prevent challenging child behaviours. Total scores can range from six to 42, with a higher score indicating higher frequency of use of proactive parenting strategies. In the current sample of 83 parents, this scale had low to adequate internal consistency at pre-program ($\alpha = .76$), post-program ($\alpha = .69$), and 1-month follow-up ($\alpha = .77$). The third subscale is called Setting Limits (four items; e.g., “Tell your child how you expected him or her to behave, such as in the grocery store), and it measures the use of parenting strategies to help a child understand limits and responsibilities. Total scores can range from four to 28, with a higher score indicating higher frequency of use of disciplinary strategies intended to set limits. In the current sample of 83 parents, this scale was found to have low to adequate internal consistency at pre-program ($\alpha = .70$), post-program ($\alpha = .69$), and 1-month follow-up ($\alpha = .61$).

Treatment attendance. For parents in the PDEP experimental 35 parents in the PDEP group, a treatment attendance form was completed. For each parent, the total number of sessions attended (out of eight) was calculated. Attendance across the 35 parents in the PDEP group ranged from four to eight sessions ($M = 6.9$, $SD = 1.3$).

Data Analysis

Standard data screening procedures. Prior to analyses, the questionnaire items were examined for the accuracy of data entry, missing values, and fit between their distributions and the assumptions of multivariate analysis. All screening procedures were conducted in SPSS 26.

Procedures for dealing with missing values. A Missing Value Analysis (MVA) determined the percentage of missing values for each of the questionnaire items and the pattern of missingness. The Little's MCAR test was used to determine if data were missing completely at random (MCAR; Little & Rubin, 1987). Of the initial sample of 83 parents, between 1% and 3.4% presented sparse missing data on several items at pre-program, post-program, and 1-month follow up. Results from the Little MCAR test were not significant ($\chi^2 = 41.40, p > .05$), indicating that the data were missing completely at random. As such, missing data were imputed using the Expectation-Maximization (EM) method. This method is known to produce unbiased estimates when compared to other imputation methods such as mean substitution (Rubin, Witkiewitz, Andre, & Reilly, 2007). Moreover, multiple imputation was not used because it is not available in HLM.

Normality. The normality assumption was assessed through skewness and kurtosis analyses. Generally, an absolute skew value larger than two or an absolute kurtosis value larger than seven are used as reference values for determining substantial non-normality (West, Finch, & Currant, 1996). Graphical methods, such as histograms and normal probability plots, were also used (Tabachnik & Fidell, 2007). Univariate normality results supported the normality assumption for the majority of the variables as the skewness and kurtosis values were below the cut-off points. However, Use of Physical Punishment was highly skewed and kurtotic (skew = 3.3 and kurtosis = 12.3 at post-PDEP; skew = 3.1 and kurtosis = 9.2 at 1-month follow-up;). This variable was normalized using square root, which improved the skew of the distribution (skew = 2.4 and kurtosis = 6.2 at post-PDEP; skew = 2.1 and kurtosis = 6.56 at 1-month follow-up).

Screening procedures for hierarchical linear modeling. In addition to the standard screening procedures described above, slopes and intercepts were examined to detect outliers in multi-level modelling. In the context of multilevel modelling, screening for outlying slopes and intercepts is necessary to detect influential cases and ensure unbiased estimates of multilevel regression coefficients

(Huta, 2014). The first step is to obtain the file with individual slopes and intercepts and convert individual scores to z-scores. A cut off of 3.3 is used as it corresponds to $p = .001$ (two-tailed; Tabachnick & Fidell, 2007). Findings indicated no outlying intercepts (i.e., outliers with respect to differences in baseline scores), but there were two cases that reported several outlying slopes (i.e., outliers with respect to differences in either baseline to post-program scores or post-program to one-month follow-up). Namely, one participant had an extremely rapid rate of improvement from pre- to post-program, compared to other participants (i.e., highly negative slope), on Approval of Physical Punishment ($z = -4.2$), Subjective Norms ($z = -4.4$), Parenting Stress ($z = -4.1$), and Challenging Behaviour ($z = -4.3$). Another participant showed an extreme increase in Use of Physical Punishment from pre- to post-program ($z = 4.5$). These cases were therefore deleted from the analyses specific to these outcomes.

Interclass correlation coefficient and power analysis: Determining the number of clusters in hierarchical linear modelling. The Interclass Correlation Coefficient (ICC) is a statistic used to indicate the strength of the relationship between units (i.e., observations) within a selected group (Tabachnick & Fidell, 2007). The ICCs are obtained from the unconditional or null models that are run for each outcome variable with no independent variables, and they are helpful in quantifying the proportion of variance explained by a grouping (random) factor in multilevel data. This statistic is also used to assess the viability of mixed (i.e., multilevel, nested) models with a cut off of 0.05 (or 5%) which is usually indicative of the necessity of a level in multilevel modeling (Huta, 2014).

For most of the outcome variables (eight out of 15), the ICCs at Level 3 (facilitator) were found to be below 5%. The ICCs at Level 2 (parent) were all found to be above 5%, lending statistical justification for using multi-level modelling with two levels (Table 29). These ICCs suggest that most of the variance related to parenting outcomes is due to differences at the parent level (range 36-68%),

whereas a much smaller percentage is related to differences at the level of the PDEP facilitator (0-15%).

Power was calculated for the HLM model with nine clusters (PDEP facilitators) at Level 3, a sample of 83 parents at Level 2, and three assessment points (i.e., repeated measures) at Level 1. For this model, an ICC of .50 and a two-sided significance of 5% was employed. The model gives a power of 54% (5.4 out of 10) to detect an intervention effect of 0.50 (medium effect size). This can also be interpreted as low power in HLM models, as it indicates slightly more than a 50% chance to find a significant effect. In addition, statistical evidence indicated that there was not enough variability at Level 3 to treat facilitator as a random effect as demonstrated by the majority of non-significant u_0 coefficients (see Table 29). Moreover, according to previous work, 10 participants per cluster (with at least 10 clusters) is necessary to achieve reliable results in a multi-level design focused on fixed effects (Hoyle & Gottfredson, 2015).

Considering both the low power, inconsistent ICCs across outcomes, and low variability at the facilitator level as well as the minimum sample required as indicated by Hoyle and Gottfredson (2015), Level 3 (PDEP facilitator) was not included in the analyses. Thus, HLM was applied to two levels of data, namely time (Level 1; 249 time points) and participant (Level 2; 83 parents). Furthermore, it is important to mention that this study was part of a larger study that aims to recruit 400 parents across all the participating agencies over a 5-year period. According to Stallard (2012), sample sizes for small-scale studies should be approximately 0.03 times that of the planned sample size for the definitive study. In this case, 83 parents for the pilot trial is well above the recommended sample size.

Growth curve modelling: Measuring the rate of change in the outcome variables. Growth curve analyses using hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) examined the effects of the PDEP program over three time points, namely pre- and post-PDEP and 1-month follow-up. The number of observations varied across participants over the three time periods because of

attrition. However, HLM is based on the full maximum likelihood estimation and is capable of using data from participants at all available time points (Guo, 2005; Raudenbush & Bryk, 2002). To be included in the analyses, parents were required to have pre-PDEP data and then data for at least 1 out of the 2 data collection points thereafter (i.e., pre- and post-PDEP or pre-PDEP and 1-month follow-up). This criterion was based on the principle that two data points are sufficient to estimate a linear trajectory in growth curve models with partially missing data (Curran, Obeidat, & Losardo, 2010). Initially, two separate datasets (one for each level) were created using the aggregation method in SPSS. The two aggregated datasets were then entered in the HLM software where model coefficients at each level were specified and simultaneously estimated. Separate analyses were run for each outcome. Random assignment and average number of services received by the parent were included as covariates. All analyses were conducted using HLM 7 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011).

The outcomes across different time points were examined using two contrast variables: 1) mean pre- to post-program change (PREPOST) and 2) mean post-program to 1-month follow-up change (POSTFLW). At Level 2, a dichotomous indicator variable for group membership (i.e., experimental vs. wait-list control; EXPCON) was created to allow statistical comparison between the experimental and control groups. Given that the majority of parents did not consent to randomization, a series of t-tests and chi-square analyses were conducted to examine whether there were any socio-demographic differences between those parents who agreed to randomization ($n = 22$) and those who did not ($n = 61$). Results indicated that non-randomized parents had fewer children ($M = 1.8$, $SD = 0.7$) than randomized parents ($M = 2.1$, $SD = 1.0$; $t = -2.90$, $p < .01$). Non-randomized parents also had a higher level of education than their non-randomized counterparts ($\chi^2_{(7)} = 11.90$, $p < .05$). To control for the effects of random assignment (i.e., randomized vs. non-randomized) a dichotomous indicator variable (RND) was created and entered as a covariate at Level 2.

Average of number of programs attended during the study (NPROGRAM) was also entered as a covariate at Level 2 when analyzing the data for the full sample. The average number of other parenting services attended by the parents over the course of the study is 0.3 ($SD = 1.4$). A simplified version of the multilevel model is presented as follows:

Level-1 Model

$$Y_{ij} = \beta_{0j} + \beta_{1j}*(PREPOST_{ij}) + \beta_{2j}*(POSTFOLL_{ij}) + r_{ij}$$

Level-2 Model

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}*(NPROGRAM_j) + \gamma_{02}*(RND_j) + \gamma_{03}*(EXPCON_j) + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}*(NPROGRAM_j) + \gamma_{12}*(RND_j) + \gamma_{13}*(EXPCON_j) + u_{1j} \\ \beta_{2j} &= \gamma_{20} + \gamma_{21}*(NPROGRAM_j) + \gamma_{22}*(RND_j) + \gamma_{23}*(EXPCON_j) + u_{2j}\end{aligned}$$

Mixed Model

where Y_{ti} is the outcome variable Y at time t for parent i . Coefficients of interest for the Level 2 model include β_{0j} (i.e., the intercept) which represents differences in baseline scores on the outcome variable between parents in the experimental and wait-list control group after controlling for the average number of other parenting programs attended and random assignment. A statistically significant β_{0j} coefficient indicates baseline differences between parents in the experimental group compared to wait-list control group; depending on the magnitude of the coefficient (-/+), it can be determined if one group scored lower/higher on a specific outcome at baseline.

β_{1j} (i.e., the coefficient of pre- to post-program change) represents the differences in the mean change on the outcome variables from pre- to post-program between parents in the experimental and the wait-list control group after controlling for the average number of other parenting programs attended and random assignment. A statistically significant β_{1j} coefficient indicates the presence of a significant change in the outcome variable from pre- to post-program for the experimental group (compared to the wait-list control group); depending on the magnitude of the coefficient (-/+) it can be

determined if the score of one group on a specific outcome decreased or increased from pre- to post-program.

β_2 (i.e., the coefficient of post-program to 1-month follow-up change) represents the differences in the mean change on the outcome variables from post-program to 1-month follow-up between parents in the experimental and the wait-list control group after controlling for the average number of other parenting programs attended and random assignment. A statistically significant β_{2j} coefficient indicates the presence of a significant change in the outcome variable from post-program to the 1-month follow-up for the experimental group (compared to the wait-list control group); depending on the magnitude of the coefficient (-/+), it can be determined if the score of one group on a specific outcome decreased or increased from post-program to the 1-month follow-up.

Coefficients u are the random effects of the model and represent the heterogeneity associated with the model intercept (u_0), the slope of pre- to post-program (u_1), and the slope of post-program to 1-month follow-up (u_2). Random effects were included to the wait-list control group for the clustering effects at Level 2. The coefficient r_{ij} is a random effect representing the deviation of the outcome of parent j across time points i , from the predicted score based on the Level 2 model.

In terms of effect size, the f^2 coefficient was used to assess the effect sizes for the fixed effects (Aiken & West, 1991; Lorah, 2018). Effect sizes of 0.02 were considered small, 0.15 medium, and 0.35 large (Cohen, 1992; Lorah, 2018).

In order to mitigate the likelihood of Type I error, the Bonferroni correction was applied across all the 17 comparisons. This resulted in a revised family-wise alpha level of 0.003 for determining significance. However, it is important to note that this study was intended to be hypothesis generating and using a statistical correction such as the Bonferroni adjustment could be unduly conservative.

Additional analyses were conducted with treatment attendance as moderator of the changes in outcomes for the PDEP experimental group (treatment attendance was measures only for the experimental group). In multi-level modelling, moderating effects are called cross-level interactions because the relation crosses at different levels of the model (i.e., variable at Level 2 such as treatment attendance and time effects at Level 1; Raudenbush, & Bryk, 2002). In this study, cross-level interactions (i.e., testing varying slope models) were analyzed to examine the effect of treatment attendance on the rate of change in outcomes from pre-to post-PDEP (effect of treatment on β_1 slope) and post-PDEP to 1-month follow-up (effect of treatment attendance on β_2) for the experimental group.

Results

Differences between the Experimental and Wait-list Control Group on Pre-PDEP Outcome Variables

Table 30 shows the mean scores for all outcome variables at pre-program for the two PDEP groups. Findings showed that, overall, there were not many statistically significant differences between the two groups. Of the significant differences, parents in the experimental group scored higher than the wait-list parents on the following: approval of non-physical punishment ($t = -2.78, p = .025$); subjective norms ($t = -3.32, p = .001$); and use of non-physical punishment ($t = -2.10, p = .038$). However, it is important to note that these baseline differences between the PDEP experimental and control group were no longer significant after accounting for the effects of random assignment and number of programs attended by the parents (see section below). In addition, a series of t-tests and crosstabs were conducted to test for the presence of any differences between randomized and non-randomized parents on any the outcomes of interest. No statistically significant differences in any of the outcome variables at pre-program were detected.

Correlations Among Variables at Pre-PDEP

Prior to conducting the multi-level analyses, simple bivariate correlations among all study variables at pre-PDEP assessment were conducted (see Table 31). With respect to cognitive/affective processes, greater approval of physical punishment was significantly correlated with higher score on approval of non-physical punishment and subjective norms. Greater approval of non-physical punishment was also significantly correlated with higher score on subjective norms score. In contrast, higher scores on approval of physical punishment, approval of non-physical punishment and subjective norms were significantly related to lower parenting self-efficacy.

With respect to action goals, higher frequency of parenting stress was significantly related to higher scores on both the Parenting Task (i.e., intensity of stress associated to parenting tasks and Challenging Behaviour (i.e., intensity of stress associated to child behaviour) subscales. Higher score on the frequency of parenting stress, Parenting Tasks, and Challenging Behaviour were significantly related to higher approval of physical punishment, higher approval of non-physical punishment, higher subjective norms, and lower parenting self-efficacy. A higher score on Child Challenging Behaviour was also significantly related to lower DERS Awareness. Moreover, a lower score on DERS Impulse (i.e., lack of impulse control) was significantly correlated with a higher score on DERS Strategies (i.e., limited access to emotion regulation strategies) and a lower score on DERS Awareness. (i.e., lack of emotional awareness). A higher score on DERS Strategies was also significantly related to a lower score on DERS Awareness. Higher scores on Parenting Task and Challenging Behaviour were also significantly related to lower use of proactive parenting strategies. Greater frequency of parenting stress, and higher scores on the Parenting Task and Challenging Behaviour subscales, were related to higher frequency of use of physical and non-physical punishment. Finally, a higher frequency of parenting stress was related to a higher frequency of use of positive discipline.

With respect to parenting practices, a higher frequency of use of non-physical punishment was significantly related to a higher frequency of use of positive discipline. Moreover, a higher frequency of use of physical punishment was significantly related to a higher frequency of use of non-physical punishment. A higher frequency of use of physical punishment was associated with higher approval of physical and non-physical punishment. A higher frequency of use of non-physical punishment was associated with higher approval of physical and non-physical punishment as well as higher subjective norms score. Greater use of supporting parenting strategies was significantly related to both greater use of proactive parenting and strategies aiming at setting limits. Greater use of proactive parenting was significantly related to greater use of strategies aiming at setting limits. Moreover, greater use of supporting parenting and proactive parenting were related to lower scores on approval of non-physical punishment and subjective norms. Conversely, greater use of both supporting parenting and proactive parenting were related to higher levels of parenting self-efficacy. Finally, higher levels of parenting self-efficacy were related to higher scores on setting limits.

Changes in Outcome Variables Over Time Between the PDEP Experimental and Wait-list Control Group

Table 32 presents the fixed effects of the conditional growth curve models for all the parenting outcomes with the time variable (effect coded) entered as the predictor, as well as the two covariates (i.e., random assignment and average number of other programs attended). Starting with the measures that tap into cognitive/affective processes, approval of physical and non-physical punishment, subjective norms, and parenting self-efficacy were examined. For Approval of Physical Punishment, there were no statistically significant differences in the pre-PDEP score (i.e., baseline) between the wait-list control and experimental group ($\beta_0 = -0.01$, $t = -0.01$, $p = .992$), after accounting for the effects of randomization and average number of parenting programs attended. The difference in mean change in Approval of Physical Punishment score from pre- to post-program was significant ($\beta_1 = -2.64$, $t = -$

3.45, $p < .001$), meaning that the experimental group's approval of physical punishment score decreased more than the wait-list control group. The pre- to post-PDEP mean change score for the experimental group was 2.64 units lower than that of the wait-list control group. The effect size for the fixed effect was large ($f^2 = .32$). The difference in mean change score from post-program to 1-month follow-up was also statistically significant ($\beta_2 = -1.68$, $t = -2.36$, $p = .020$), indicating a 1.68 unit decrease in the approval of physical punishment for the experimental group compared to the wait-list control group. The effect size for the fixed effect was small ($f^2 = .01$).

For Approval of Non-Physical Punishment, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = -0.57$, $t = -0.58$, $p = .056$), after accounting for the effects of randomization and average number of parenting programs attended. The difference in mean change in Approval of Non-Physical Punishment score from pre- to post-program was significant ($\beta_1 = -3.09$, $t = -4.45$, $p < .001$), meaning that the experimental group's approval of non-physical punishment score decreased more than the wait-list control group. The pre-to post-PDEP mean change score for the experimental group was 3.09 units lower than that of the wait-list control group. The effect size for the fixed effect was large ($f^2 = .54$). The difference in mean change from post-PDEP to 1-month follow-up was also statistically significant ($\beta_2 = -1.47$, $t = -2.89$, $p = .005$), indicating a 1.47 unit decrease in approval of non-physical punishment from post-program to 1-month follow-up for the experimental group compared with the wait-list control group. The effect size for the fixed effect was small ($f^2 = .03$).

For Subjective Norms, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 0.08$, $t = 0.07$, $p = .944$), after accounting for the effects of randomization and average number of parenting programs attended. The mean difference in change in Subjective Norms score from pre- to post-program score was significant ($\beta_1 = -4.34$, $t = -5.71$, $p < .001$), meaning that the experimental group's subjective norms score decreased more than the

wait-list control group. The pre-to post-PDEP mean change score was 4.34 units lower for the experimental group compared to the wait-list control group. The effect size for the fixed effect was large ($f^2 = .58$). The mean difference in change in Subjective Norms score from post-PDEP to 1-month follow-up was also statistically significant ($\beta_2 = -2.33$, $t = -3.55$, $p < .001$), indicating a 2.33 unit decrease in subjective norms from post-program to 1-month follow-up for the experimental group compared to the wait-list control group. The effect size for the fixed effect was small ($f^2 = .01$).

For Parenting Self-Efficacy, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = -0.66$, $t = 0.71$, $p = .480$), after accounting for the effects of randomization and average number of parenting programs attended. The mean difference in change in Parenting Self-Efficacy score from pre- to post-PDEP score was significant ($\beta_1 = 1.01$, $t = 1.99$, $p = .049$), meaning that the experimental group's self-efficacy score increased more than the wait-list control group. The pre-to post-PDEP mean change score was 1.01 units higher for the experimental group compared to the wait-list control group. The effect size for the fixed effect was large ($f^2 = .37$). The mean difference in change score from post-PDEP to 1-month follow-up, however, was not significant ($\beta_2 = 0.77$, $t = 1.75$, $p = .085$).

For DERS Impulse, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 0.67$, $t = 1.35$, $p = .180$), after accounting for the effects of randomization and average number of parenting programs attended. The mean difference in change in DERS Impulse score from pre- to post-PDEP score was not significant ($\beta_1 = -0.37$, $t = -1.23$, $p = .220$). The mean difference in change in DERS Impulsivity score from post-PDEP to 1-month follow-up was significant ($\beta_2 = -0.61$, $t = -2.07$, $p = .041$), indicating 0.61 unit decrease in impulsivity from post-program to follow-up for the experimental group compared to the wait-list control group. The effect size for the fixed effect was small ($f^2 = .01$).

For DERS Strategies, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = -0.21, t = -0.48, p = .635$), after accounting for the effects of randomization and average number of parenting programs attended. The mean difference in change in DERS Strategy score from pre- to post-PDEP score was not significant ($\beta_1 = -0.06, t = -0.19, p = .847$), nor was the mean difference in change score from post-PDEP to 1-month follow-up ($\beta_2 = -0.08, t = -0.28, p = .782$).

For DERS Awareness, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = -0.48, t = -0.76, p = .451$), after accounting for the effects of randomization and average number of parenting programs attended. The mean difference in change in DERS Awareness score from pre- to post-PDEP score was not significant ($\beta_1 = 0.22, t = 0.49, p = .626$), nor was the mean difference in change in score from post-PDEP to 1-month follow-up ($\beta_2 = -0.02, t = -0.05, p = .957$).

With respect to measures taping into action goals, changes in parenting stress and problem-solving were assessed. For the frequency total score on the Parenting Daily Hassles scale, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 0.78, t = 0.68, p = .906$), after accounting for the effects of randomization and average number of parenting programs attended. Neither the mean difference in change in Parenting Daily Hassles total score from pre- to post-PDEP ($\beta_1 = 0.06, t = 0.05, p = .959$) nor the mean difference in score from post-PDEP to 1-month follow-up ($\beta_2 = -0.91, t = -0.79, p = .431$) was statistically significant.

For the Parenting Task subscale of the Parenting Daily Hassles, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 0.37, t = 0.35, p = .728$), after accounting for the effects of randomization and average number of

parenting programs attended. Neither the mean difference in change in Parenting Task score from pre- to post-PDEP ($\beta_1 = -0.87, t = -1.40, p = .162$) nor the mean difference in from post-PDEP to 1-month follow-up was statistically significant ($\beta_2 = -0.91, t = -1.42, p = .157$).

For the Challenging Behaviour subscale of the Parenting Daily Hassles, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 0.63, t = 0.64, p = .519$), after accounting for the effects of randomization and average number of parenting programs attended. The mean difference in change in Challenging Behaviour score from pre- to post-PDEP was significant ($\beta_1 = -1.34, t = -2.48, p = .014$), indicating a 1.34 unit decrease in parenting stress associated to children's challenging behaviour for the experimental group compared to the wait-list control group from pre-to post-program. The effect size for the fixed effect was large ($f^2 = .17$). The mean difference in Challenging behaviour score from post-PDEP to 1-month follow-up was also statistically significant ($\beta_{21} = -1.32, t = -2.35, p = .020$), indicating a 1.32 unit decrease in parenting stress associated to children's challenging behaviour for the experimental group compared to the wait-list control group. The effect size for the fixed effect was small ($f^2 = .01$).

For the Approach/Avoidance subscale of Problem Solving, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 1.38, t = 1.12, p = .267$), after accounting for the effects of randomization and average number of parenting programs attended. Neither the mean difference in change in Approach/Avoidance score from pre- to post-PDEP ($\beta_1 = 0.02, t = 0.02, p = .983$) nor the post-PDEP to 1-month follow-up mean difference was statistically significant ($\beta_2 = 0.73, t = 0.84, p = .399$).

With respect to measures that tap into parenting practices, the use of physical and non-physical punishment, positive discipline and parenting strategies were assessed. For Use of Physical Punishment, there were no statistically significant differences in the pre-PDEP score between the wait-

list control and experimental group ($\beta_0 = -0.01, t = -0.26, p = .875$), after accounting for the effects of randomization and average number of parenting programs attended. Neither the mean difference in change in Use of Physical Punishment score from pre- to post-PDEP ($\beta_1 = -0.02, t = -0.82, p = .074$), nor the mean difference in change score from post-PDEP to 1-month follow-up ($\beta_2 = -0.18, t = -1.55, p = .124$) were statistically significant.

For Use of Non-Physical Punishment, there were no statistically significant differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = -0.27, t = -0.51, p = .612$), after accounting for the effects of randomization and average number of parenting programs attended. The mean difference in change in Use of Non-Physical Punishment score from pre- to post-PDEP was significant ($\beta_1 = -1.71, t = -5.89, p < .001$), indicating a 1.71 unit decrease in the use of non-physical punishment from pre- to post-program for the experimental group compared to the wait-list control group. The effect size for the fixed effect was large ($f^2 = .62$). The mean difference in change score from post-PDEP to 1-month follow-up was also statistically significant ($\beta_2 = -0.72, t = -2.41, p = .017$), indicating a decrease of 0.72 units in the use of non-physical punishment from post-PDEP to 1-month follow-up for the experimental group compared to the wait-list control group. The effect size for the fixed effect was small ($f^2 = .08$).

For Use of Positive Discipline, there were no statistically differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 0.04, t = 0.03, p = .976$), after accounting for the effects of randomization and average number of parenting programs attended. Neither the mean difference in change in Use of Positive Discipline score from pre- to post-PDEP ($\beta_1 = 1.42, t = 1.86, p = .066$) nor the mean difference in change score post-PDEP to 1-month follow-up score ($\beta_2 = 1.08, t = 1.57, p = .121$) were statistically significant.

For Supporting Parenting, there were no statistically differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 0.28, t = 0.50, p = .618$), after accounting for the effects of randomization and average number of parenting programs attended. The mean difference in change in Supporting Parenting score from pre- to post-program score was significant ($\beta_1 = 1.14, t = 3.11, p = .003$). The pre-to post-PDEP mean change score was 1.14 units higher for the experimental group compared to the wait-list control group. The effect size for the fixed effect was small ($f^2 = .12$). The mean difference in change in Supporting Parenting score from post-PDEP to 1-month follow-up was also statistically significant ($\beta_2 = 0.93, t = 2.25, p = .027$), indicating an increase of 0.93 units in supporting parenting from post-program to 1-month follow-up for the experimental group compared to the wait-list control group. The effect size for the fixed effect was small ($f^2 = .01$).

For Proactive Parenting, there were no statistically differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 0.28, t = 0.47, p = .653$), after accounting for the effects of randomization and average number of parenting programs attended. The mean difference in change in Proactive Parenting score from pre- to post-PDEP score was significant ($\beta_1 = 1.06, t = 2.56, p = .011$). The pre-to post-PDEP mean change score was 1.06 units higher for the experimental group compared to the wait-list control group. The effect size for the fixed effect was large ($f^2 = .24$). The mean difference in change score from post-PDEP to 1-month follow-up was not statistically significant ($\beta_2 = 0.40, t = 0.94, p = .346$).

For Setting Limits, there were no statistically differences in the pre-PDEP score between the wait-list control and experimental group ($\beta_0 = 0.28, t = 0.77, p = .445$), after accounting for the effects of randomization and average number of parenting programs attended. Neither the mean difference in change in Setting Limits scores from pre- to post-PDEP ($\beta_1 = 0.21, t = 0.70, p = .484$) nor the mean

difference in change score from post-PDEP to 1-month follow-up ($\beta_2 = 0.13$, $t = 0.45$, $p = .651$) were statistically significant.

Bonferroni Adjustment for Multiple Comparisons: Follow-up to the Analyses on the Changes in the Outcome Between the Experimental and Control Groups.

Following Bonferroni corrections, some differences were noted for several of the fixed effects of the conditional growth model (see Table 32). Namely, the mean change from post-PDEP to 1-month follow-up for the experimental group (relative to the control group) for Approval of Physical Punishment, Approval of Non-Physical Punishment, Parenting Self-Efficacy, DERS Impulse, Use of Non-Physical Punishment, and Proactive Parenting were no longer significant. Moreover, for Challenging Behaviour and Supporting Parenting, both the mean change from pre- to post-PDEP and from post-PDEP to 1-month follow-up of the experimental group (relative to the control group) were no longer significant.

Treatment Attendance as Moderator of the Change in Outcomes for the PDEP-Experimental Group.

Additional analyses were conducted to examine the moderating effects of treatment attendance on the pre-to post-PDEP and post-PDEP to 1-month follow-up slopes of the 35 parents in the experimental group. No significant moderating effects associated with treatment attendance were found for any of the outcomes of interest, thereby indicating that none of the slopes (i.e., pre- to post-program and post-program to 1-month follow-up) were significantly related to treatment attendance (all p -values $> .05$; see Table 33).

Discussion

Results from the growth curve analysis indicated significant changes for some of the outcomes of interest. With respect to the cognitive/affective processes, and in line with the initial hypotheses,

parents who completed PDEP reported decreases in approval of physical punishment relative to non-PDEP parents both shortly and after completion of the PDEP program and at 1-month follow-up. The effect sizes were predominantly in the medium to large range for the pre-to-post change and small for the post- to follow-up change. These findings corroborate previous results from a small single-group (i.e., no control) PDEP outcomes study conducted with a sample of Canadian parents (Durrant et al., 2014). Findings are also in line with previous program implementation data across different countries (Ademi-Shala et al., 2016; Mori et al., 2016). Indeed, it appears that information regarding the harmful effects of physical punishment, along with shifting the attention toward long-term goals (as opposed to short-term goals based on the child immediate compliance), may be a key determinant in reducing positive attitudes toward physical punishment (Stenason et al., 2020).

The current study indicates that PDEP also is effective in changing attitudes toward non-physical (emotional) punishment. In this sample, PDEP parents reported lower scores on approval of non-physical punishment (at both program completion and 1-month follow-up), compared to non-PDEP parents. This is an important finding in that it provides evidence of PDEP's ability to change parents' attitudes about non-physical punishment (e.g., time-outs, shaming), whose effectiveness is currently debated among researchers in the field of child development and parenting (Larzelere et al., 2017). Studies reporting beneficial effects of punitive non-physical methods have not been consistently replicated across different samples (Durrant et al., 2017; Holden et al., 2017). Moreover, non-physical punishment is considered a power assertive technique that does not coincide with the goals of positive parenting approaches, which instead focus on conflict resolution strategies based on collaboration, problem-solving, and relationship building (Holden et al., 2017).

Turning to subjective norms, the initial hypothesis was also supported. That is, compared with non-PDEP parents, those who completed the program reported a shift in attributions about children's behaviours both at completion of the program and at 1-month follow-up. They were less likely to

consider the child normative behaviour as intentional and undesirable. These changes were manifested at both the end of the program and at follow-up, and they corroborate previous findings from a non-controlled study (Durrant et al., 2014). The current findings support the idea that PDEP is effective in changing the perception of parents with respect to children's behaviours. It is possible that after taking the program, parents were able to contextualize challenging child behaviours within the child's stage of development and other factors (e.g., fatigue, temperament). Such shifts in perspective could have, in turn, favoured a positive change in the parental belief system (which includes both attitudinal predispositions toward the use of physical and emotional punishment and negative child behavioural attributions).

With respect to parenting self-efficacy (i.e., parents' perception of their own efficacy in using positive disciplinary strategies), the initial hypothesis was partially supported. Compared to non-PDEP parents, PDEP parents reported greater parenting self-efficacy upon program completion, and there were no additional increases from post-program to 1-month follow-up. The effect size for the pre- to post-program change was large. The current findings of an effect for parenting self-efficacy from pre- to post-program contrast with those from Durrant and colleagues (2014) who did not report any significant change pre- to post-program using the total score of a reduced version of the self-efficacy scale (i.e., original version of the self-efficacy scale composed of only three items). The current study utilized a modified scale for parenting self-efficacy which included relevant items specifically targeting parents' sense of competence with respect to the use of disciplinary strategies. These results not only provide support for PDEP's impact on enhancing parenting self-efficacy but also provide further predictive validity of the parenting self-efficacy scale that was modified as part of this dissertation. There are several mechanisms that might explain the effects of PDEP on parenting self-efficacy. In one of the initial sessions of the PDEP program, parents learned about the stress response and how stress affects the brain and can trigger unwanted behaviours on the part of the parent. As a result, it may be

that parents were better able to regulate difficult emotions triggered by stressful parenting situations and, hence, implement positive parenting strategies while avoiding punishments. Such improvements in emotion regulation and parenting practices might have, in turn, contributed to an increased sense of parenting competence.

In regard to parental emotion regulation, the initial hypothesis was partially supported in that PDEP parents (relative to non-PDEP parents) reported a significant decline in impulsivity from post-program to 1-month follow-up; however, the size of this effect was small. Moreover, there was no significant change in impulsivity immediately after completion of the program. It is possible that parents require some time to elaborate their feelings and gain control over their emotions. It is important to note that this the first study to examine the effects of the PDEP program on parental emotional regulation. One of the PDEP modules focuses on teaching parents about the neural mechanisms underlying the stress response and emotion regulation. Parents who participated in the PDEP program seem to have learned some practical strategies and real-life examples (e.g., blowing bubbles) that have helped them manage their distress and challenging emotions (Stenanson et al., 2020).

Turning to action goals, the PDEP parents (relative to the parents who did not complete the program) experienced an improvement in parenting stress in terms of the perceived intensity of challenging child behaviours, both from pre- to post-program and from post-program to 1-month follow-up. The effect sizes were predominantly in the medium to large range for the pre- to post-change and small for the post- to follow-up change. Conversely, in contrast with the initial hypothesis, no significant effects were reported for the frequency of perceived parenting stress or for stress related to parenting tasks. However, in interpreting these effects, it is important to note that past research has examined the global score of parenting stress rather than analyzing different subcomponents of this construct, so it is difficult to compare the current study results to previous findings. As proposed in the

PDEP logic model, it is possible that by end of the program, parents were able to gain better control their emotions in relation to children's challenging behaviour and feel more competent with respect to the use of positive disciplinary strategies. This might have, in turn, contributed to lowering their perceived stress related to challenging child behaviour.

Contrary to hypotheses, no differences in changes between PDEP and non-PDEP parents were found for problem-solving abilities. The explanation may lie in the measure that was used, because the problem-solving approach/avoidance subscale appears to target individuals' approach and avoidance tendencies toward seeking solutions to a problem without any specific reference to parenting situations. As such, it is possible that PDEP effects would not be seen through this scale because it lacks specificity. Moreover, problem-solving is a topic that is covered only in the very last sessions of the PDEP program. In a recent qualitative study conducted with a small sample of parents who completed PDEP, some parents indicated that the problem-solving module consisted of a large amount of material that was covered in a short period of time. Parents felt they would have benefitted from spending more time on the material presented in this module (Stenanson et al., 2020).

In terms of parenting practices, findings indicated that PDEP parents (relative to non-PDEP parents) reported less frequent use of non-physical punishment immediately after the program and at 1-month follow-up. This is a novel finding that deserves particular attention since the PDEP program is one of few parenting programs that explicitly discourages the use of non-physical punishment, such as time-outs, parental psychological control (e.g., shaming, guilt-induction), and harsh verbal discipline. In line with the PDEP logic model, the decline in the use of non-physical punishment could be the result of an increase in parents' understanding of children's feelings and behaviours, as well as parents' greater self-reflection and insight following program completion.

Contrary to hypotheses, the PDEP and non-PDEP groups did not differ in terms of changes in the use of physical punishment. This is surprising given PDEP's focus on physical punishment and its

main intent to reduce the use of harsh parenting practices. Given the low mean scores prior to the start of the program for both PDEP and non-PDEP parents, it is possible that there was a floor effect so that it was difficult for scores to decrease much further in any meaningful way. Moreover, the scale had poor internal consistency ($\alpha = .43-.49$); therefore, this subscale might not have constituted a reliable instrument to assess the frequency of parental use of physical punishment.

Similarly, the lack of significant changes with respect to the use of positive disciplinary strategies for parents who completed PDEP (relative to non-PDEP parents) was not in line with the initial hypotheses. However, some of the items of the scale do not refer specifically to positive parenting strategies but rather to the ability of parents to reflect on their own behaviour in an attempt to address a challenging situation with their child (e.g., breathing deeply to calm oneself, talking calming to the child, apologizing to the child). Moreover, this scale has not undergone rigorous psychometric testing, which might raise questions regarding the appropriateness of this instrument.

In line with expectations, the PDEP parents (relative to non-PDEP parents) experienced a greater increase in supportive (from both pre- to post-program and from post-program to 1-month follow-up) and proactive parenting (from pre- to post-program only). This supports the main aim of PDEP in helping parents improve warmth in their relationship with their children and providing support to face daily challenges. However, no significant effects were found for setting limits. It is possible that this scale did not adequately capture the concept of “structure” as intended by the PDEP program. For example, the scale refers to setting limits to help reduce challenging behaviour (e.g., Explain what you wanted your child to do in clear and simple ways?; Tell your child what you wanted him or her to do rather than tell him/her to stop doing something?), but it misses several key aspects of structure that are addressed in the PDEP program, such as establishing routines for children and breaking down tasks to allow for scaffolding and successful task completion.

In sum, the abovementioned results from this small-scale study provide some preliminary evidence for the effectiveness of the PDEP program as demonstrated by the significant findings for the majority of the outcomes of interest. The ongoing research on the evaluation of the PDEP will employ a much larger sample to improve statistical power and potentially contribute to a better understanding of the impact of the PDEP program in Ontario.

Limitations and Future Research Directions

One of the limitations of this study was the inability to use randomization across all participants, which led to the design being quasi-experimental in nature. Quasi-experimental designs have several advantages over purely correlational studies in that they allow relevant comparisons (i.e., experimental vs. control group) designed to reduce confounding effects (Kazdin, 2016). However, the lack of randomization raises concerns regarding participant selection biases. In the current study, parents' preference for which group they attended was prioritized in an effort to balance both the needs of parents with adherence to the requirements of a randomized controlled trial. Despite this limitation, the randomized and non-randomized participants were comparable on the majority of demographic factors, and the variable of randomization was entered as a covariate in all analyses. Moreover, this study is part of a larger project that will continue recruitment of families for the next 2-3 years. As recruitment continues, it is expected that differences between randomized and non-randomized parents will reduce even further.

Another limitation was the exclusive reliance on parental self-report data to assess program outcomes. Given the sensitivity of the topics that were investigated (e.g., physical punishment), it is possible that these measures were subject to social desirability biases. However, it would be ethically difficult to measure parents' use of physical and emotional punishment in terms of actual behaviours given the obligations of mandated reporting. Moreover, the measurement instrument used to assess disciplinary strategies, namely the use of physical punishment, non-physical punishment, and positive

discipline, had not undergone formal assessment of its psychometric properties. In a future stage of this larger research project, once we obtain data from a larger number of parents, these measures will undergo more rigorous testing (e.g., IRT analysis). Furthermore, the research team is currently examining the possibility of integrating the use of observational measures of other proxy variables. An example of these measures is the Parent-Child Aggression Acceptability Movie Task (Parent-CAAM Task; Rodriguez, Russa, & Harmon, 2011). The Parent-CAAM Task is an analog task to implicitly assess beliefs regarding the use of punishment, consisting of eight 90-sec movie clips depicting varying levels of parent-child aggression (five physical abuse, three physical discipline). Participants are asked, based only on that scene, to stop the video if and when they believe the scene has become physically abusive. Scores are derived from the number of seconds until the parent makes a decision to halt the video. Slower response time in judging a scene as abusive is conceptualized as indicating greater acceptability of parent-child aggression. This measure, however, has several issues with respect to its level of acceptability of severity among Canadian parents and as such it needs to be validated before being able to be used with this population.

Finally, the sample size was sufficient to capture potential program effects at the parent level (Level 2 of the multilevel model), but it did not allow for the modelling of the effects at either the facilitator (Level 3) or agency levels (Level 4). However, it should be noted that none of effects at the facilitator level were significant, as demonstrated by both the magnitude of the ICCs and the significance of the random Level 3 effects. Nonetheless, higher-order effects at both the facilitator and agency level will be adequately investigated after the data collection of the 5-year project is completed and the sample size will meet the requirements to run a more complex model.

General Discussion

The first part of this dissertation (Study 1 and 2) focused on the evaluation and modification of a measurement instrument designed to assess constructs from the Positive Discipline in Everyday

Parenting (PDEP) program. The constructs from this PDEP questionnaire consist of several scales, namely parental approval of punishment (e.g., physical and emotional), subjective norms, and parenting self-efficacy. In Study 1, the psychometric characteristics of the original version of the PDEP questionnaire were examined to obtain general information about the fit of each of the scales. Following these analyses, the scales were modified in that problematic items were either re-worked or eliminated, and additional items were added. Following these modifications, the content of the revised PDEP questionnaire was examined by two expert samples (e.g., professionals in the area of child development and parenting research; parents who had not completed the program). The experts provided useful feedback that was used to make additional changes and arrive at a second revised version of the PDEP questionnaire. In Study 2, the psychometric characteristics of the revised version of the questionnaire were examined using a new sample of 400 parents. A subsample of the parents also was re-assessed 4-weeks later for a test-retest reliability check. The validity of the constructs was also examined by correlating the PDEP scales with other measures assessing similar constructs. At the end of the process, a valid and reliable measure of the constructs of interest (e.g., approval of physical and non-physical punishment, subjective norms, parenting self-efficacy) was obtained and it was used to assess the impact of PDEP on various aspects of parenting (Study 3).

Generally, an effective evaluation of a program relies on the use of valid and reliable measurement instruments that are grounded in theoretical and empirical evidence (Lichiello & Turnock, 1999). Results from the first study indicated that the revised PDEP questionnaire had high content relevance, and formal examination of the questionnaire's factor structure indicated good psychometric properties of items within each of the four scales. The scales also demonstrated good convergent validity, adequate to excellent internal consistency and test-retest reliability. These findings are important since the original PDEP questionnaire had never undergone such rigorous examination to ensure its validity and reliability.

The second part of the dissertation (Study 3) evaluated PDEP outcomes by way of a quasi-experimental design that made use of the revised PDEP questionnaire, as well as other instruments. Results indicated that parents who completed the PDEP program reported some changes in line with program goals. PDEP targets two main components of parenting, namely cognitive/emotional processes (i.e., approval of punishment, subjective norms, and parenting self-efficacy) and action goals (i.e., reduction of parenting stress, enhanced problem-solving and emotion regulation). With the exception of a few outcomes (e.g., problem-solving, awareness of emotion regulation abilities and strategies), the findings provided evidence for the effectiveness of PDEP in addressing both cognitive/affective processes and action goals. The results also indicate that PDEP was effective in reducing parental use of non-physical punishment and increasing supportive and proactive parenting.

Implications

Theoretical/research implications. This study provides evidence of the applicability of the PDEP conceptual model, hence contributing to the advancement of theoretical knowledge in PDEP outcome research. First, the validation of the PDEP measures has allowed for a stronger connection between PDEP goals and outcomes. These results advanced the validity of the PDEP conceptual model, in that they provided an alignment between the measures validated in this study and the PDEP program outcomes. This was also confirmed in the third study (evaluation of the PDEP's effectiveness) as most of hypotheses regarding the outcome variables representing the key aspects of the program were supported.

In terms of outcome assessment, the current findings provide evidence of the validity of the PDEP questionnaire primarily in a community-based sample of Canadian parents. Given that physical and emotional punishment is a problem often encountered with maltreating parents, future studies should be conducted to validate the PDEP questionnaire with this population in order to establish normative scores. The availability of such scores will provide additional information on the distinction

between “normative” levels of physical and non-physical punishment use and other forms of maltreatment such as physical and verbal abuse.

Moreover, a validity assessment of other measures, such as the use of physical and non-physical punishment and the use of positive discipline, was not carried out because this was beyond the thesis requirements. Future studies should conduct a formal validity and reliability assessment of these measures to gather more information on the psychometric quality of these instruments. In such endeavour, it would also be useful to examine the content of the items of each measure to ensure their alignment with the PDEP theoretical framework.

The third study also served to highlight the interplay of both cognitive/affective processes and action goals in changing parenting behaviours. Further research is needed to examine such dynamics, including both mediating and moderating mechanisms through which these variables operate in the context of the PDEP program. This can be accomplished through the use of sophisticated techniques such as structural equation modelling (SEM) with the intent to uncover cause-effect relationships in experimental studies.

Clinical/applied implications. Findings from this dissertation, in particular Study 3, point to the importance of early intervention programs in helping parents shift their attitudes and behaviours toward more positive forms of child discipline. Findings highlighted the benefits of participating in the PDEP program, particularly with respect to cognitive/affective processes and parenting behaviours. PDEP reduced positive attitudes toward physical and non-physical (e.g., emotional) punishment, reduced negative attributions about children’s challenging behaviour, and enhanced parenting self-efficacy. Part of the PDEP program is dedicated to teaching parents about different stages of child development and also demystifying common misinterpretations of challenging behaviour. PDEP helped parents shift their focus from perceiving common challenging behaviour (e.g., “tantrums”) as intentional to contextualizing the behaviours as part of the developmental growth process. As a result

of the PDEP program, parents gained a better understanding of children's thinking and feelings that, in turn, helped them reflect on the best positive approach to raise their children without using punishment (physical and/or emotional). Moreover, in the PDEP program, parents learned about warmth and structure and how these two concepts go hand in hand. In fact, parents who complete PDEP appeared more likely to rely on supportive parenting behaviour, which is centred on encouragement and warmth, and to use proactive strategies to reduce the likelihood of challenging behaviour and parent-child conflict.

Moreover, as part of the PDEP program, parents learned about the stress response and how to identify sources of stress. While there were no effects of PDEP with respect to changes in the frequency of parenting stress and with parenting stress relative to parenting tasks, the program appeared effective in helping parents lower the stress associated with handling challenging child behaviours. Similarly, PDEP appeared to have an impact on parents' self-reported impulsivity. Part of the PDEP program is dedicated to helping parents understand the stress response and the limbic system's response to stress (e.g., emotion dysregulation). This content, coupled with a review of the various developmental stages and associated typical behaviours, might have contributed to lowering the stress of parents in response to challenging child behaviours.

Finally, PDEP has demonstrated its effectiveness in reducing the frequency of non-physical punishment. As indicated by the ongoing debate on the effects of non-physical punishment on children's behaviour (e.g., Holden et al., 2017; Larzelere et al., 2017), it remains uncertain whether or not non-physical punishment can be regarded as a negative parenting technique. In the PDEP program, it is argued that any form of punishment deteriorates the parent-child bond and, as such, should be avoided. This remains a sensitive topic that requires further investigation.

References

- Ademi-Shala, R., Hoxha, L., & Ateah, C. (2016, August). *Delivering Kosovo's first parenting program: Challenges, strategies, and outcomes*. Paper presented at the conference of the International Society for Prevention of Child Abuse and Neglect, Calgary, Alberta, Canada.
- Afifi, T. O., Brownridge, D. A., Cox, B. J., & Sareen, J. (2006). Corporal punishment, childhood abuse and psychiatric disorders. *Child Abuse & Neglect*, 30, 1093-1103. DOI: 10.1016/j.chiabu.2006.04.006
- Afifi, T. O., Mota, N. P., Dasiewicz, P., MacMillan, H. L., & Sareen, J. (2012). Physical punishment and mental disorders: results from a nationally representative US sample. *Pediatrics*, 130, 184-192. DOI: 10.1542/peds.2011-2947. DOI: DOI: <https://doi.org/10.1542/peds.2011-2947>
- Afifi, T. O., Ford, D., Gershoff, E. T., Merrick, M., Grogan-Kaylor, A., Ports, K. A., ... & Bennett, R. P. (2017). Spanking and adult mental health impairment: The case for the designation of spanking as an adverse childhood experience. *Child Abuse & Neglect*, 71, 24-31. DOI: <https://doi.org/10.1016/j.chiabu.2017.01.014>
- Afifi, T. O., Fortier, J., MacMillan, H. L., Gonzalez, A., Kimber, M., Georgiades, K., ... & Struck, S. (2019). Examining the relationships between parent experiences and youth self-reports of slapping/spanking: a population-based cross-sectional study. *BMC Public Health*, 19(1), 1-8. DOI: 10.1186/s12889-019-7729-6
- Aiken, L., & West, S. (1991). *Multiple regression: Testing and interpreting interactions*. Sage Publications.
- Ajzen, I. (2012). The theory of planned behavior. In P. A. M. Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. 1, pp. 438-459). London: Sage.

Akaike, H. (1974). A new look at the statistical model identification. *Transactions on Automatic Control*, 19, 716–723. DOI: 10.1109/TAC.1974.1100705.

American Academy of Pediatrics. (2018). *What's the best way to discipline my child?*. Retrived from: <https://www.healthychildren.org/English/family-life/family-dynamics/communication-discipline/Pages/Disciplining-Your-Child.aspx>.

American Educational Research Association. (2014). *Standards for educational and psychological testing*. Washington, DC. AERA.

Ardelt, M., & Eccles, J. S. (2001). Effects of mothers' parental efficacy beliefs and promotive parenting strategies on inner-city youth. *Journal of Family Issues*, 22, 944–972. DOI: <https://doi-org.proxy.bib.uottawa.ca/10.1177%2F019251301022008001>

Ateah, C. A., & Durrant, J. E. (2005). Maternal use of corporal punishment in response to child misbehavior: implications for child abuse prevention. *Child Abuse & Neglect*, 29, 169-185. DOI: <https://doi.org/10.1016/j.chiabu.2004.10.010>

Azar, S. T., Miller, E. A., McGuier, D. J., Stevenson, M. T., O'Donnell, E., Olsen, N., & Spence, N. (2016). Maternal social information processing and the frequency and severity of mother-perpetrated physical abuse. *Child Maltreatment*, 21(4), 308-316. DOI: <https://doi.org/10.1177/1077559516668047>

Azar, S. T., Reitz, E. B., & Goslin, M. C. (2008). Mothering: Thinking is part of the job description: Application of cognitive views to understanding maladaptive parenting and doing intervention and prevention work. *Journal of Applied Developmental Psychology*, 29(4), 295-304. DOI: <https://doi.org/10.1016/j.appdev.2008.04.009>

- Azar, S. T., Robinson, D. R., Hekimian, E., & Twentyman, C. T. (1984). Unrealistic expectations and problems solving ability in maltreating and comparison mothers. *Journal of Consulting and Clinical Psychology*, 52, 687–691. DOI: <https://doi.org/10.1037/0022-006X.52.4.687>
- Azar, S. T., & Rohrbeck, C. A. (1986). Child abuse and unrealistic expectations: Further validation of the Parent Opinion Questionnaire. *Journal of Consulting and Clinical Psychology*, 54(6), 867–868. DOI: <https://doi.org/10.1037/0022-006X.54.6.867>
- Azar, S. T., Stevenson, M. T., & Johnson, D. R. (2012). Intellectual disabilities and neglectful parenting: Preliminary findings on the role of cognition in parenting risk. *Journal of Mental Health Research in Intellectual Disabilities*, 5(2), 94-129. DOI: <https://doi.org/10.1080/19315864.2011.615460>
- Baker, F. (2001). *The basics of item response theory*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland College Park, MD.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215. DOI: <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura A. (1982) Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122– 147. DOI: <https://doi.org/10.1037/0003-066X.37.2.122>
- Barber, B. K., & Harmon, E. L. (2002). Violating the self: Parental psychological control of children and adolescents. In B. K. Barber (Ed.), *Intrusive parenting: How psychological control affects children and adolescents* (pp. 15–52). Washington, DC: APA.
- Bavolek SJ, & Keene RG (2001). *Adult-Adolescent Parenting Inventory (AAPI-2): Administration and Development Handbook*. Park City, UT: Family Development Resources, Inc.

- Berlin, L. J., Ispa, J. M., Fine, M. A., Malone, P. S., Brooks-Gunn, J., Brady-Smith, C., Ayoub, C., & Bai, Y. (2009). Correlates and consequences of spanking and verbal punishment for low-income White, African American, and Mexican American toddlers. *Child Development, 80*, 1403-1420. DOI: <https://doi.org/10.1111/j.1467-8624.2009.01341.x>
- Breitenstein, S. M., Fogg, L., Ocampo, E. V., Acosta, D. I., & Gross, D. (2016). Parent use and efficacy of a self-administered, tablet-based parent training intervention: A randomized controlled trial. *JMIR mHealth and uHealth, 4*(2), e36. DOI: 10.2196/mhealth.5202
- Breitenstein, S. M., Gross, D., & Bettencourt, A. F. (2020). The Chicago Parent Program. In E. T. Gershoff & S. J. Lee (Eds.), *Ending the physical punishment of children: A guide for clinicians and practitioners* (p. 109–119). American Psychological Association. Doi: <https://doi.org/10.1037/0000162-012>
- Breitenstein, S. M., Gross, D., Fogg, L., Ridge, A., Garvey, C., Julion, W., & Tucker, S. (2012). The Chicago Parent Program: Comparing 1-year outcomes for African American and Latino parents of young children. *Research in Nursing & Health, 35*(5), 475-489. DOI: 10.1002/nur.21489
- Breitenstein, S. M., Shane, J., Julion, W., & Gross, D. (2015). Developing the eCPP: Adapting an evidence-based parent training program for digital delivery in primary care settings. *Worldviews on Evidence-Based Nursing, 12*(1), 31-40. DOI: <https://doi.org/10.1111/wvn.12074>
- Breitenstein, S. M., Schoeny, M., Risser, H., & Johnson, T. (2016). A study protocol testing the implementation, efficacy, and cost effectiveness of the ezParent program in pediatric primary care. *Contemporary Clinical Trials, 50*, 229-237. DOI: <https://doi.org/10.1016/j.cct.2016.08.017>

Bugental, D. B., & Johnston, C. (2000). Parental and child cognitions in the context of the family.

Annual Review of Psychology, 51, 315–344. DOI:

<https://doi.org/10.1146/annurev.psych.51.1.315>

Bugental, D. B., Johnston, C., New, M., & Silvester, J. (1998). Measuring parental attributions:

Conceptual and methodological issues. *Journal of Family Psychology*, 12, 459. DOI:

<https://psycnet.apa.org/doi/10.1037/0893-3200.12.4.459>

Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge.

Cai, L., Thissen, D., & du Toit, S. H. C. (2011). *IRTPRO for Windows [Computer software]*.

Lincolnwood, IL: Scientific Software International.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and Validity Assessment* (Vol. 17). Sage publications.

Chen, W. H., & Thissen, D. (1997). Local dependence indexes for item pairs using item response

theory. *Journal of Educational and Behavioral Statistics*, 22, 265-289. DOI: 10.2307/1165285

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159. DOI: 10.1037//0033-

2909.112.1.155

Coley, R., Kull, M., & Carrano, J. (2014). Parental endorsement of spanking and children's

internalizing and externalizing problems in African American and Hispanic Families. *Journal of Family Psychology*, 28, 22-31. DOI: 10.1037/a0035272

Coleman, P. K., & Karraker, K. H. (2000). Parenting self-efficacy among mothers of school-age

children: Conceptualization, measurement, and correlates. *Family Relations*, 49, 13-24. DOI:

<https://psycnet.apa.org/doi/10.1111/j.1741-3729.2000.00013.x>

- Conners, N. A., Whiteside-Mansell, L., Deere, D., Ledet, T., & Edwards, M.C. (2006). Measuring the potential for child maltreatment: The reliability and validity of the Adult Adolescent Parenting Inventory-2. *Child Abuse & Neglect*, 30, 39–53. DOI: <https://doi.org/10.1016/j.chiabu.2005.08.011>
- Corralejo, S. M., Jensen, S. A., Greathouse, A. D., & Ward, L. E. (2018). Parameters of time-out: Research update and comparison to parenting programs, books, and online recommendations. *Behavior Therapy*, 49, 99-112. DOI: <https://doi.org/10.1016/j.beth.2017.09.005>
- Coyl, D. D., Roggman, L. A., & Newland, L. A. (2002). Stress, maternal depression, and negative mother–infant interactions in relation to infant attachment. *Infant Mental Health Journal*, 23, 145-163. DOI: <https://doi.org/10.1002/imhj.10009>
- Coyne, J. C., & Kwakkenbos, L. (2013). Triple P-Positive Parenting programs: the folly of basing social policy on underpowered flawed studies. *BMC Medicine*, 11(1), 11. DOI: <https://doi.org/10.1186/1741-7015-11-11>
- Criminal Code, R.S.C. 1985, c.46, s.231(6).
- Crnic, K. A., & Greenberg, M. T. (1990). Minor parenting stresses with young children. *Child Development*, 61, 1628-1637. DOI: [10.1111/j.1467-8624.1990.tb02889.x](https://doi.org/10.1111/j.1467-8624.1990.tb02889.x)
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334. DOI: <https://doi.org/10.1007/BF02310555>
- Cunningham, C. E., Bremner, R., & Boyle, M. (1995). Large group community-based parenting programs for families of preschoolers at risk for disruptive behaviour disorders: Utilization, cost effectiveness, and outcome. *Journal of Child Psychology and Psychiatry*, 36(7), 1141-1159. DOI: [10.1111/j.1469-7610.1995.tb01362.x](https://doi.org/10.1111/j.1469-7610.1995.tb01362.x)

- Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, 11(2), 121-136. DOI: 10.1080/15248371003699969
- Dadds, M. R., & Tully, L. A. (2019). What is it to discipline a child: What should it be? A reanalysis of time-out from the perspective of child mental health, attachment, and trauma. *American Psychologist*. DOI: 10.1037/amp0000449
- De Ayala, R. J. (2013). *The theory and practice of item response theory*. Guilford Publications.
- Deater-Deckard, K. (2004). *Parenting Stress*. New Haven, CT: Yale University Press.
- Deater-Deckard, K., & Dodge, K. A. (1997). Externalizing behavior problems and discipline revisited: Nonlinear effects and variation by culture, context, and gender. *Psychological Inquiry*, 8, 161–175. DOI: https://doi.org/10.1207/s15327965pli0803_1
- Derella, O. J., Burke, J. D., Stepp, S. D., & Hipwell, A. E. (2019). Reciprocity in undesirable parent–child behavior? Verbal aggression, corporal punishment, and girls’ oppositional defiant symptoms. *Journal of Clinical Child & Adolescent Psychology*, 1-14. DOI: <https://doi.org/10.1080/15374416.2019.1603109>
- DeVellis, R. F. (2003). *Scale development: Theory and applications (2nd ed.)*. Thousand Oaks, California, US: Sage.
- Dix, T. H., & Grusec, J. E. (1985). Parent attribution processes in the socialization of children. In I. E. Sigel (Ed.), *Parental belief systems: The psychological consequences for children* (pp. 319–346). Hillsdale, NJ: Erlbaum.

- Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and perspectives*, 38(1), 105. Retrieved from:
<https://search.informit.com.au/documentSummary;dn=491551710186460;res=IELIAC>
- Dubowitz, H. (2020). Safe Environment for Every Kid (SEEK). Holden, G. W., & Holland, G. W. (2020). Motivational interviewing. In E. T. Gershoff, & S. J. Lee. *Ending the physical punishment of children: A guide for clinicians and practitioners* (pp. 63-72). American Psychological Association.
- Durrant, J. E. (2007, 2013). *Positive discipline: What it is and how to do it*. Save the Children Sweden Southeast Asia and the Pacific.
- Durrant, J. E., & Ensom, R. (2020). An Eriksonian perspective on physical punishment and its impacts on mental health. *The Psychoanalytic Study of the Child*, 73(1), 30-45. DOI:
<https://doi.org/10.1080/00797308.2020.1690866>
- Durrant, J. E., & Ensom, R. (2012). Corporal punishment of children: Lessons from 20 years of research. *Canadian Medical Association Journal*, 184, 1373-1377. DOI:
<https://doi.org/10.1503/cmaj.101314>
- Durrant J. E., & Ensom R. (2004) *Joint statement on physical punishment of children and youth*. Ottawa, Canada: Coalition on Physical Punishment of Children and Youth.
- Durrant, J., Baker, L., Jones, A. D., Cavers, V., Shanks, T., & Roloff, L. (2018, June). *Preventing punitive violence against children: A three-province evaluation of a primary prevention program*. Final report to the Prairie Action Foundation. Winnipeg Manitoba: Mosaic Newcomer Family Resource Network, Saskatchewan Prevention Institute, & Alberta Home Visitation Network Association.

- Durrant, J., Plateau, D. P., Ateah, C. A., Holden, G. W., Barker, L. A., Stewart-Tufescu, A., ... & Ahmed, R. (2017). Parents' views of the relevance of a violence prevention program in high, medium, and low human development contexts. *International Journal of Behavioral Development, 41*(4), 523-531.
- Durrant, J. E., Plateau, D. P., Ateah, C., Stewart-Tufescu, A., Jones, A., Ly, G., Barker, L., Holden, G., Kearley, C., MacAulay, J., Peters, R. D., & Tapanya, S. (2014). Preventing punitive violence: preliminary data on the Positive Discipline in Everyday Parenting (PDEP) Program. *Canadian Journal of Community Mental Health, 33*, 109-125. DOI: <https://doi.org/10.7870/cjcmh-2014-018>
- Durrant, J. E., Rose-Krasnor, L., & Broberg, A. G. (2003). Physical punishment and maternal beliefs in Sweden and Canada. *Journal of Comparative Family Studies, 34*, 585-604. DOI: <https://www.jstor.org/stable/41603902>
- Durrant, J. E., & Stewart-Tufescu, A. (2017). What is "Discipline" in the age of children's rights? *The International Journal of Children's Rights, 25*(2), 359-379. DOI: <https://doi.org/10.1163/15718182-02502007>
- Durrant, J. E., Trocmé, N., Fallon B., Milne, C., Black, T., & Knoke, D. (2006). *Punitive violence against children in Canada [CECW information sheet #41E]*. Toronto, ON: University of Toronto, Faculty of Social Work.
- Duncombe, M., Havighurst, S., Kehoe, C., Holland, K., Frankling, E., & Stargatt, R. (2016). Comparing an emotion- and a behavior-focused parenting program as part of a multisystemic intervention for child conduct problems. *Journal of Clinical Child & Adolescent Psychology, 45*(3), 320–334. DOI: 10.1080/15374416.2014.963855

Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: L. Erlbaum Associates.

Fabrizio, C. S., Lam, T. H., Hirschmann, M. R., Pang, I., Yu, N. X., Wang, X., & Stewart, S. M. (2015). Parental emotional management benefits family relationships: a randomized controlled trial in Hong Kong, China. *Behaviour Research and Therapy*, 71, 115–124. DOI: 10.1016/j.brat.2015.05.011.

Fréchette, S., & Romano, E. (2015). Change in corporal punishment over time in a representative sample of Canadian parents. *Journal of Family Psychology*, 29, 507-519. DOI: 10.1037/fam0000104

Fréchette, S., Zoratti, M., & Romano, E. (2015). What is the link between corporal punishment and child physical abuse? *Journal of Family Violence*, 30, 135-148. DOI: <https://doi.org/10.1007/s10896-014-9663-9>

Gagné, M-H., Tourigny, M., Joly, J & Pouliot-Lapointe, J. (2007). Predictors of adult attitudes toward corporal punishment of children. *Journal of Interpersonal Violence*, 22 (10), 1285-1304. DOI: 10.1177/0886260507304550

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.)*. Boston: Allyn & Bacon.

Gershoff, E.T. (2002). Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. *Psychological Bulletin*, 128, 539-579. DOI: 10.1037//0033-2909.128.4.539

- Gershoff, E. T. (2010). More harm than good: A summary of scientific research on the intended and unintended effects of corporal punishment on children. *Law and Contemporary Problems*, 73, 31-56. DOI: <https://scholarship.law.duke.edu/lcp/vol73/iss2/3>
- Gershoff, E. T., & Grogan-Kaylor, A. (2016). Spanking and child outcomes: Old controversies and new meta-analyses. *Journal of Family Psychology*. DOI: <http://dx.doi.org/10.1037/fam0000191>.
- Gershoff, E. T., Grogan-Kaylor, A., Lansford, J. E., Chang, L., Zelli, A., Deater-Deckard, K., & Dodge, K. A. (2010). Parent discipline practices in an international sample: Associations with child behaviors and moderation by perceived normativeness. *Child Development*, 81(2), 487-502. DOI: 10.1111/j.1467-8624.2009.01409.x
- Gershoff, E. T., & Lee, S. J. (2020). *Ending the physical punishment of children: A guide for clinicians and practitioners* (pp. viii-189). American Psychological Association.
- Gidron Y. (2013) Group Therapy/Intervention. In: Gellman M.D., Turner J.R. (eds) *Encyclopedia of Behavioral Medicine*. Springer, New York, NY. DOI: https://doi.org/10.1007/978-1-4419-1005-9_1409
- Global Initiative to End All Corporal Punishment of Children. (2017). *Prohibiting corporal punishment: A human rights imperative*. Retrieved from: <http://www.endcorporalpunishment.org/prohibiting-corporal-punishment/introduction.html>.
- Goger, P., Rozenman, M., & Gonzalez, A. (2020). The association between current maternal psychological control, anxiety symptoms, and emotional regulatory processes in emerging adults. *Journal of Behavior Therapy and Experimental Psychiatry*, 68, 101563. DOI: <https://doi.org/10.1016/j.jbtep.2020.101563>

- Gonzalez, M., Ateah, C. A., Durrant, J. E., & Feldgaier, S. (2019). The Impact of the Triple P seminar series on Canadian parents' use of physical punishment, non-physical punishment and non-punitive responses. *Behaviour Change*, 36(2), 102-120. DOI: 10.1017/bec.2019.7
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual review of psychology*, 60, 549-576. DOI: 10.1146/annurev.psych.58.110405.085530
- Graham-Bermann, S. A. (2012). *The Moms' Empowerment Program: A training manual*. Ann Arbor: Department of Psychology, University of Michigan.
- Graham-Bermann, S., & Miller-Graff, L. (2015). Community-based intervention for women exposed to intimate partner violence: A randomized control trial. *Journal of Family Psychology*, 29, 537-547. DOI:10.1037/fam0000091.
- Grogan-Kaylor, A., Castillo, B., Ma, J., Ward, K., Lee, S., Pace, G., & Park, J. (2020). A Bayesian analysis of associations between neighborhoods, spanking and child externalizing behavior. *Children and Youth Services Review*, 112. DOI: <https://doi.org/10.1016/j.childyouth.2020.104930>.
- Grogan-Kaylor, A., Galano, M. M., Howell, K. H., Miller-Graff, L., & Graham-Bermann, S. A. (2019). Reductions in parental use of corporal punishment on pre-school children following participation in the moms' empowerment program. *Journal of Interpersonal Violence*, 34(8), 1563-1582. DOI: <https://doi.org/10.1177%2F0886260516651627>
- Gromoske, A. N., & Maguire-Jack, K. (2012). Transactional and cascading relations between early spanking and children's social-emotional development. *Journal of Marriage and Family*, 74, 1054-1068. DOI: 10.1111/j.1741-3737.2012.01013.x

- Gross, D., Fogg, L., Webster-Stratton, C., Garvey, D., Julion, W., & Grady, J. (2003). Parent training with multi-ethnic families of toddlers in day care in low- income urban communities. *Journal of Consulting and Clinical Psychology*, 71, 261–278. DOI: 10.1037/0022-006X.71.2.261
- Gülseven, Z., Carlo, G., Streit, C., Kumru, A., Selcuk, B., & Sayıl, M. (2018). Longitudinal relations among parenting daily hassles, child rearing, and prosocial and aggressive behaviors in Turkish children. *Social Development*, 27(1), 45-57. DOI: <https://doi.org/10.1111/sode.12247>
- Guo, S. (2005). Analyzing grouped data with hierarchical linear modeling. *Children and Youth Services Review*, 27(6), 637-652. DOI: 10.1016/j.childyouth.2004.11.017
- Hambleton, R.K., & Jones, R.W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, 12, 38-47. DOI: 10.1111/j.1745-3992.1993.tb00543.x
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hamovitch, E. K., Acri, M. C., & Bornheimer, L. A. (2019). An analysis of the relationship between parenting self-efficacy, the quality of parenting, and parental and child emotional health. *Journal of Family Social Work*, 22(4-5), 337-351. DOI: 10.1080/10522158.2019.1635939
- Haskett, ME., Scott, SS., Willoughby, M., Ahern, L., & Nears, K. (2006). The parent opinion questionnaire and child vignettes for use with abusive parents: Assessment of psychometric properties. *Journal of Family Violence*, 21(2), 137-151. DOI: <https://doi.org/10.1007/s10896-005-9010-2>
- Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7, 238. DOI: 10.1037/1040-3590.7.3.238

- Heppner, P. P., & Petersen, C. H. (1982). The development and implications of a personal problem-solving inventory. *Journal of Counseling Psychology*, 29(1), 66. DOI: <https://doi.org/10.1037/0022-0167.29.1.66>
- Hill, N. E., & Bush, K. R. (2001). Relationships between parenting environment and children's mental health among African American and European American mothers and children. *Journal of Marriage and the Family*, 63, 954–966. DOI: <https://psycnet.apa.org/doi/10.1111/j.1741-3737.2001.00954.x>
- Holden, G. W., & Buck, M. J. (2002). Parental attitudes toward child rearing. In M. H. Bornstein (Ed.), *Handbook of parenting* (2nd ed., pp. 537–562). Mahwah, NJ: Erlbaum
- Holden, G. W., Coleman, S. M., & Schmidt, K. L. (1995). Why 3-year-old children get spanked: Parent and child determinants as reported by college-educated mothers. *Merrill-Palmer Quarterly*, 41, 431-452. DOI: <https://www.jstor.org/stable/23087935>
- Holden, G. W., & Edwards, L. A. (1989). Parental attitudes toward child rearing: Instruments, issues, and implications. *Psychological Bulletin*, 106(1), 29. DOI: <https://psycnet.apa.org/doi/10.1037/0033-2909.106.1.29>
- Holden, G. W., Grogan-Kaylor, A., Durrant, J. E., & Gershoff, E. T. (2017). Researchers deserve a better critique: Response to Larzelere, Gunnoe, Roberts, and Ferguson. *Marriage & Family Review*, 53(5), 465-490. DOI: 10.1080/01494929.2017.1308899
- Holden, G. W., & Holland, G. W. (2020). Motivational interviewing. In E. T. Gershoff, & S. J. Lee. *Ending the physical punishment of children: A guide for clinicians and practitioners* (pp. 41-46). American Psychological Association.

- Hoyle, R. H., & Gottfredson, N. C. (2015). Sample size considerations in prevention research applications of multilevel modeling and structural equation modeling. *Prevention Science*, 16(7), 987-996. DOI: 10.1007/s11121-014-0489-8
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. DOI: <https://doi.org/10.1080/10705519909540118>
- Hulin, C. L., Drasgow, F., & Parsons, C.K. (1983). *Item response theory*. Homewood, IL: Dow Jones-Irwin.
- Huta, V. (2014). When to use hierarchical linear modeling. *The Quantitative Methods for Psychology*, 10, 13–28. DOI: 10.20982/tqmp.10.1.p013
- Jones, T. L., & Prinz, R. J. (2005). Potential roles of parental self-efficacy in parent and child adjustment: A review. *Clinical Psychology Review*, 25(3), 341–363. DOI: 10.1016/j.cpr.2004.12.004
- John, O. P., Soto, C. J., Robins, R. W., Fraley, R. C., & Krueger, R. F. (2007). The importance of being valid. In R.W. Robins, R.C., Fraley, & R. F., Krueger (Eds.) *Handbook of research methods in personality psychology*, (pp. 461-494). Guilford Press.
- Jud, A., & Trocmé, N. (2012). *Physical abuse and physical punishment in Canada*. *Child Canadian Welfare Research Portal Information Sheet # 122E*. Toronto, ON: University of Toronto, Faculty of Social Work.
- Kaminski, J. W., & Claussen, A. H. (2017). Evidence base update for psychosocial treatments for disruptive behaviors in children. *Journal of Clinical Child & Adolescent Psychology*, 46, 477–499. DOI:10.1080/15374416.2017.1310044.

- Kane, G. A., Wood, V. A., & Barlow, J. (2007). Parenting programmes: a systematic review and synthesis of qualitative research. *Child: Care, Health and Development*, 33(6), 784-793. DOI: 10.1111/j.1365-2214.2007.00750.x
- Kazdin, A. (2016). *Research design in clinical psychology* (5th ed.). Allyn and Bacon.
- Knight, R. M., Albright, J., Deling, L., Dore-Stites, D., & Drayton, A. K. (2020). Longitudinal relationship between time-out and child emotional and behavioral functioning. *Journal of Developmental & Behavioral Pediatrics*, 41(1), 31-37. DOI: 10.1097/DBP.0000000000000725
- Kolko, D. J., Dorn, L. D., Bukstein, O., & Burke, J. D. (2008). Clinically referred ODD children with or without CD and healthy controls: comparisons across contextual domains. *Journal of Child and Family Studies*, 17(5), 714–734. DOI: 10.1007/s10826-007-9186-6
- Krug, E. G., Mercy, J. A., Dahlberg, L. L., & Zwi, A. B. (2002). The world report on violence and health. *The lancet*, 360, 1083-1088. Retrieved from:
https://www.who.int/violence_injury_prevention/violence/world_report/en/.
- Larzelere, R. E., Gunnoe, M. L., Roberts, M. W., & Ferguson, C. J. (2017). Children and parents deserve better parental discipline research: Critiquing the evidence for exclusively “positive” parenting. *Marriage & Family Review*, 53(1), 24-35. DOI:
<https://doi.org/10.1080/01494929.2016.1145613>
- Lee, S. J., Grogan-Kaylor, A., & Berger, L. M. (2014). Parental spanking of 1-year-old children and subsequent child protective services involvement. *Child Abuse & Neglect*, 38, 875-883. DOI: 10.1016/j.chiabu.2014.01.018
- Lichiello, P., & Turnock, B. J. (1999). *Guidebook for Performance measurement*. Turning Point.
- Little, R. J. A., & Rubin, D. B. (1987). *Analysis with missing data*.

- Liu, L., & Wang, M. (2015). Parenting stress and harsh discipline in China: The moderating roles of marital satisfaction and parent gender. *Child Abuse & Neglect*, 43, 73-82. DOI: <https://doi.org/10.1016/j.chiabu.2015.01.014>
- Loeb, E. L., Kansky, J., Tan, J. S., Costello, M. A., & Allen, J. P. (2020). Perceived psychological control in early adolescence predicts lower levels of adaptation into mid-adulthood. *Child Development*. DOI: <https://doi.org/10.1111/cdev.13377>
- Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, 6(1), 8. DOI: 10.1186/s40536-018-0061-2
- Lorber, M. F., O'Leary, S. G., & Kendziora, K. T. (2003). Mothers' overreactive discipline and their encoding and appraisals of toddler behavior. *Journal of Abnormal Child Psychology*, 31(5), 485–494. DOI: <https://doi.org/10.1023/A:1025496914522>
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Routledge.
- Lord, F.M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Lundahl, B., Risser, H.J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*, 26, 86–104. DOI: 10.1016/j.cpr.2005.07.004
- MacKenzie, M. J., Nicklas, E., Waldfogel, J., & Brooks-Gunn, J. (2013). Spanking and child development across the first decade of life. *Pediatrics*, 132, 1118-1125. DOI: <https://doi.org/10.1542/peds.2013-1227>
- Maydeu-Olivares, A., & Joe, H. (2005). Limited-and full-information estimation and goodness-of-fit testing in 2 n contingency tables: A unified framework. *Journal of the American Statistical Association*, 100(471), 1009-1020. DOI: <https://doi.org/10.1198/016214504000002069>

- Maydeu-Olivares, A., & Joe, H. (2006). Limited information goodness-of-fit testing in multidimensional contingency tables. *Psychometrika*, 71(4), 713. DOI: 10.1007/s11336-005-1295-9
- Mas, S. (2015, December 14). Truth and reconciliation offers 94 ‘calls to action’. *CBC News*. Retrieved from: <http://www.cbc.ca/news/politics/truth-and-reconciliation-94-calls-to-action-1.3362258>.
- Milner, J. S. (2003). Social information processing in high-risk and physically abusive parents. *Child Abuse & Neglect*, 27, 7-20. DOI: [https://psycnet.apa.org/doi/10.1016/S0145-2134\(02\)00506-9](https://psycnet.apa.org/doi/10.1016/S0145-2134(02)00506-9)
- Mori, I., Mochizuki, R., & Stewart-Tufescu, A. (2016, August). *Transforming parents’ beliefs about physical punishment through “Positive Discipline in Everyday Parenting” in Japan*. Paper presented at the Conference of the International Society for Prevention of Child Abuse and Neglect, Calgary, Alberta, Canada.
- Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3, 1-18. DOI: [https://doi.org/10.1016/0022-2496\(66\)90002-2](https://doi.org/10.1016/0022-2496(66)90002-2)
- Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.
- Olofsson, V., Skoog, T., & Tillfors, M. (2016). Implementing group based parenting programs: A narrative review. *Children and Youth Services Review*, 69, 67-81. DOI: <https://doi.org/10.1016/j.childyouth.2016.07.004>
- Osterlind, S. J. (1983). *Test item bias*. Beverly Hills, CA: Sage.
- Palmer, E. J., & Hollin, C. R. (2001). Sociomoral reasoning, perceptions of parenting and self-reported delinquency in adolescents. *Applied Cognitive Psychology*, 15, 85-100. DOI: [https://doi.org/10.1002/1099-0720\(200101/02\)15:1%3C85::AID-ACP691%3E3.0.CO;2-6](https://doi.org/10.1002/1099-0720(200101/02)15:1%3C85::AID-ACP691%3E3.0.CO;2-6)

- Parliament of Canada. (2019). *Senate public bill. S-206: An act to amend the Criminal Code (protection of children against standard child-rearing violence)*. Retrieved on April 2020 from: <https://www.parl.ca/LegisInfo/BillDetails.aspx?Language=E&billId=8063354>.
- Perron, J. L., Lee, C. M., LaRoche, K. J., Ateah, C., Clément, M. È., & Chan, K. (2014). Child and parent characteristics associated with Canadian parents' reports of spanking. *Canadian Journal of Community Mental Health*, 33, 31-45. DOI: <https://doi.org/10.7870/cjcmh-2014-014>
- Plotkin, R. (1983). *Cognitive mediation in disciplinary actions among mothers who have abused or neglected their children: Dispositional and environmental factors*. Doctoral dissertation, University of Rochester. Dissertation Abstracts International, 45-02B, 684.
- Piché, G., Huynh, C., Clément, M. È., & Durrant, J. E. (2017). Predicting externalizing and prosocial behaviors in children from parental use of corporal punishment. *Infant and Child Development*, 26(4), e2006. DOI: <https://doi.org/10.1002/icd.2006>
- Polcari, A., Rabi, K., Bolger, E., & Teicher, M. H. (2014). Parental verbal affection and verbal aggression in childhood differentially influence psychiatric symptoms and wellbeing in young adulthood. *Child Abuse & Neglect*, 38(1), 91-102. DOI: 10.1016/j.chiabu.2013.10.003
- Portwood, S. G., Lambert, R. G., Abrams, L. P., & Nelson, E. B. (2011). An evaluation of the adults and children together (ACT) against violence parents raising safe kids' program. *The Journal of Primary Prevention*, 32(3-4), 147-160. DOI: 10.1007/s10935-011-0249-5
- Posada, G., & Pratt, D. (2008). Physical aggression in the family and preschoolers' use of the mother as a secure base. *Journal of Marital and Family Therapy*, 34, 14-27. DOI: <https://doi.org/10.1111/j.1752-0606.2008.00050.x>

- Prendergast, S., & MacPhee, D. (2020). Trajectories of maternal aggression in early childhood: associations with parenting stress, family resources, and neighborhood cohesion. *Child Abuse & Neglect*, 99, 104315. DOI: <https://doi.org/10.1016/j.chiabu.2019.104315>
- Prinz, R. J. Triple P- Positive Parenting Program. In E. T. Gershoff, & S. J. Lee. *Ending the physical punishment of children: A guide for clinicians and practitioners* (pp. 133-144). American Psychological Association.
- 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, DOI: <http://dx.doi.org/10.1007/s11121-009-0123-3>
- Querido, J. G., Bearss, K., & Eyberg, S. M. (2002). Parent/child interaction therapy. In F. W. Kaslow & T. Patterson (Eds.), *Comprehensive handbook of psychotherapy: Cognitive-behavioral approaches* (Vol. 2, pp. 91–113). Hoboken, NJ: Wiley.
- Rajyaguru, P., Moran, P., Cordero, M., & Pearson, R. (2019). Disciplinary parenting practice and child mental health: Evidence from the UK Millennium Cohort study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 58(1), 108-116. DOI: 10.1016/j.jaac.2018.06.033
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X. F., Martinez, A., Bloom, H., & Hill, C. (2011b). *Optimal design software for multi-level and longitudinal research* (Version 3.01) [Software]. Retrieved from: <https://www.wtgrantfoundation.org>.
- Reckase, M. D. (1979) Unifactor latent trait models applied to multifactor test: Results and implications. *Journal of Educational Statistics*, 4, 207–230. DOI: 10.2307/1164671

- Reise, S. P., & Henson, J. M. (2003). A discussion of modern versus traditional psychometrics as applied to personality assessment scales. *Journal of Personality Assessment*, 81, 93-103. DOI: 10.1207/S15327752JPA8102_01.
- Roberts, M. W., & Powers, S. W. (1990). Adjusting chair timeout enforcement procedures for oppositional children. *Behavior Therapy*, 21(3), 257-271. DOI: 10.1016/S0005-7894(05)80329-6
- Rodriguez, C. M., Russa, M. B., & Harmon, N. (2011). Assessing abuse risk beyond self-report: Analog task of acceptability of parent-child aggression. *Child Abuse & Neglect*, 35(3), 199-209. DOI: 10.1016/j.chiabu.2010.12.004
- Rodriguez, C. M., Smith, T. L., & Silvia, P. J. (2016). Multimethod prediction of physical parent-child aggression risk in expectant mothers and fathers with Social Information Processing theory. *Child Abuse & Neglect*, 51, 106-119. DOI: 10.1016/j.chiabu.2015.10.028
- Rodriguez, C. M., & Wittig, S. M. (2019). Predicting child problem behavior and maternal/paternal parent-child aggression: identifying early prevention targets. *Journal of Applied Developmental Psychology*, 60, 76-86. DOI: 10.1016/j.appdev.2018.11.001
- Romano, E., Bell, T., & Norian, R. (2013). Corporal punishment: Examining attitudes toward the law and factors influencing attitude change. *Journal of Family Violence*, 28, 265-275. DOI: <https://psycnet.apa.org/doi/10.1007/s10896-013-9494-0>
- Rominov, H., Giallo, R., & Whelan, T. A. (2016). Fathers' postnatal distress, parenting self-efficacy, later parenting behavior, and children's emotional-behavioral functioning: A longitudinal study. *Journal of Family Psychology*, 30(8), 907. DOI: 10.1037/fam0000216
- Rubin, L. H., Witkiewitz, K., Andre, J. S., & Reilly, S. (2007). Methods for handling missing data in the behavioral neurosciences: Don't throw the baby rat out with the bath water. *Journal of*

Undergraduate Neuroscience Education, 5(2), A71. DOI:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3592650/>

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores.

Psychometrika Monograph Supplement, 34, 100–114. DOI: <https://doi.org/10.1002/j.2333-8504.1968.tb00153.x>

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

Sanders, M. R., Turner, K. M., & Metzler, C. W. (2019). Applying self-regulation principles in the delivery of parenting interventions. *Clinical Child and Family Psychology Review*, 22(1), 24-42. DOI: <https://link.springer.com/article/10.1007/s10567-019-00287-z>

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

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461-464. DOI: <https://www.jstor.org/stable/2958889>

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological methods*, 7, 147-177. DOI: <https://psycnet.apa.org/doi/10.1037/1082-989X.7.2.147>

Silva, J. (2007). *Parents Raising Safe Kids: ACT 8-week program for parents*. Washington, DC: American Psychological Association.

Simms, L. J., & Watson, D. (2007). The construct validation approach to personality scale construction. In R.W. Robins, R.C., Fraley, & R. F., Krueger (Eds.) *Handbook of research methods in personality psychology* (pp. 240-258). Guilford Press.

- Slep, A. M. S., & O'Leary, S. G. (1998). The effects of maternal attributions on parenting: An experimental analysis. *Journal of Family Psychology*, 12, 234. DOI: <https://psycnet.apa.org/doi/10.1037/0893-3200.12.2.234>
- Smith, A. M., & O'Leary, S. G. (1995). Attributions and arousal as predictors of maternal discipline. *Cognitive Therapy and Research*, 19, 459-471. DOI: <https://doi.org/10.1007/BF02230412>
- Smith, J. D., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (2015). Negative relational schemas predict the trajectory of coercive dynamics during early childhood. *Journal of Abnormal Child Psychology*, 43(4), 693-703. DOI: <https://doi.org/10.1007/s10802-014-9936-z>
- Snarr, J. D., Slep, A. M. S., & Grande, V. P. (2009). Validation of a new self-report measure of parental attributions. *Psychological Assessment*, 21, 390-401. DOI: 10.1037/a0016331
- Soenens, B., & Vansteenkiste, M. (2010). A theoretical upgrade of the concept of parental psychological control: Proposing new insights on the basis of self-determination theory. *Developmental Review*, 30(1), 74-99. DOI: <https://doi.org/10.1016/j.dr.2009.11.001>
- Stallard, N. (2012). Optimal sample sizes for phase II clinical trials and pilot studies. *Statistics in Medicine*, 31, 1031-1042. DOI: <https://doi.org/10.1002/sim.4357>
- Steiger, J. H., & Lind, J. C. (1980). *Statistically based tests for numbers of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.
- Stenason, L., Moorman, J., & Romano, E. (2020). The experiences of parents and facilitators in a positive parenting program. *The Qualitative Report*, 25(1), 1-13. DOI: <https://nsuworks.nova.edu/tqr/vol25/iss1/>

- Straus, M. A., & Field, C. J. (2003). Psychological aggression by American parents: National data on prevalence, chronicity, and severity. *Journal of Marriage and Family*, 65(4), 795-808. DOI: <https://psycnet.apa.org/doi/10.1111/j.1741-3737.2003.00795.x>
- Straus, M. A., & Paschall, M. J. (2009). Corporal punishment by mothers and development of children's cognitive ability: A longitudinal study of two nationally representative age cohorts. *Journal of Aggression, Maltreatment & Trauma*, 18, 459-483. DOI: <https://doi.org/10.1080/10926770903035168>
- Straus, M. A., & Stewart, J. H. (1999). Corporal punishment by American parents: National data on prevalence, chronicity, severity, and duration in relation to child and family characteristics. *Clinical Child and Family Psychology Review*, 2, 55–70. DOI: <https://doi.org/10.1023/A:1021891529770>
- Sturge-Apple, M. L., Li, Z., Martin, M. J., Jones-Gordils, H. R., & Davies, P. T. (2020). Mothers' and fathers' self-regulation capacity, dysfunctional attributions and hostile parenting during early adolescence: A process-oriented approach. *Development and Psychopathology*, 32(1), 229-241. DOI: 10.1017/S0954579418001694
- Sturge-Apple, M. L., Skibo, M. A., Rogosch, F. A., Ignjatovic, Z., & Heinzelman, W. (2011). The impact of allostatic load on maternal sympathovagal functioning in stressful child contexts: Implications for problematic parenting. *Development and Psychopathology*, 23(3), 831. DOI: 10.1017/s0954579411000332
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics*. Boston: Pearson/Allyn & Bacon.

- Taylor, C. A., Guterman, N. B., Lee, S. J., & Rathouz, P. J. (2009). Intimate partner violence, maternal stress, nativity, and risk for maternal maltreatment of young children. *American Journal of Public Health, 99*(1), 175-183. DOI: <https://dx.doi.org/10.2105%2FAJPH.2007.126722>
- UN General Assembly. (1989). Convention on the Rights of the Child. U.N. Doc. A/RES/44/25 (1989). Retrieved on February 8, 2014, from <http://www.unhchr.ch/html/menu3/b/k2crc.htm>
- UNICEF. (2014). *Hidden in plain sight: A statistical analysis of violence against children*. New York, NY: UNICEF.
- United Nations Committee on the Rights of the Child. (2006). *General Comment No. 8, The right of the child to protection from corporal punishment and other cruel or degrading forms of punishment* (arts. 19; 28, para. 2; and 37, inter alia) (Forty-second session, 2006), U.N. Doc. CRC/C/GC/8.
- Vittrup, B., Holden, G. W., & Buck, J. (2006). Attitudes predict the use of physical punishment: A prospective study of the emergence of disciplinary practices. *Pediatrics, 117*, 2055-2064. DOI: 10.1542/peds.2005-2204
- Wang, M. T., & Kenny, S. (2014). Longitudinal links between fathers' and mothers' harsh verbal discipline and adolescents' conduct problems and depressive symptoms. *Child Development, 85*(3), 908-923. DOI: 10.1111/cdev.12143
- Ward, K. P., Lee, S. J., Limb, G. E., & Grogan-Kaylor, A. (2019). Physical punishment and child externalizing behavior: Comparing American Indian, White, and African American children. *Journal of Interpersonal Violence* (advanced online publication). DOI: 10.1177/0886260519861678.
- Webster-Stratton, C. (2001). The incredible years: Parents, teachers, and children training series. *Residential Treatment for Children & Youth, 18*, 31-45. DOI: http://dx.doi.org/10.1300/J007v18n03_04

- Wei, C., & Kendall, P. C. (2014). Child perceived parenting behavior: Childhood anxiety and related symptoms. *Child & Family Behavior Therapy*, 36(1), 1-18. DOI: 10.1080/07317107.2014.878175
- West, S. G. Finch, J. F., & Carrant, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In Hoyle R. H. (Ed), *Structural equation modeling: Concepts, issues and applications* (pp. 56-75). Newbery Park, CA: Sage.
- Wilson, C., Gardner, F., Burton, J., & Leung, S. (2007). Maternal attributions and observed maternal behaviour: Are they linked? *Behavioural and Cognitive Psychotherapy*, 35, 165–178. DOI: <https://psycnet.apa.org/doi/10.1017/S1352465806003195>
- Wilson, P., Rush, R., Hussey, S., Puckering, C., Sim, F., Allely, C. S., ... & Gillberg, C. (2012). How evidence-based is an evidence-based parenting program? A PRISMA systematic review and meta-analysis of Triple P. *BMC Medicine*, 10(1), 130. DOI: 10.1186/1741-7015-10-130
- Wittkowski, A., Garrett, C., Calam, R. et al. Self-Report Measures of Parental Self-Efficacy: A Systematic Review of the Current Literature. *Journal of Child Family Studies*, 26, 2960–2978 (2017). DOI: <https://doi-org.proxy.bib.uottawa.ca/10.1007/s10826-017-0830-5>
- Wright, B. D. (1967). *Sample free test calibration and person measurement*. Paper presented at the ETS Invitational Conference on Testing Problems. Retrieved from <http://www.rasch.org/memol.htm>.
- Xu, J., Paek, I., & Xia, Y. (2017). Investigating the Behaviors of M² and RMSEA² in Fitting a Unidimensional Model to Multidimensional Data. *Applied Psychological Measurement*, 41(8), 632-644. DOI: 10.1177/0146621617710464
- Zolotor, A. J., Theodore, A. D., Chang, J. J., Berkoff, M. C., & Runyan, D. K. (2008). Speak softly – and forget the stick: Corporal punishment and child corporal abuse. *American Journal of Preventive Medicine*, 35, 364–369. DOI: <https://doi.org/10.1016/j.amepre.2008.06.031>

Table 1

Mean, Standard Deviation, Skewedness and Kurtosis of the PDEP items (current version).

Item code^a	Mean	SD	Skewedness	Kurtosis
APP1	2.09	1.27	0.88	-0.35
APP2	2.34	1.45	0.79	-0.42
APP3	2.93	1.61	0.27	-1.06
APP4	2.45	1.51	0.72	-0.59
APP5	2.51	1.36	0.67	-1.46
ANP1	1.82	1.11	1.46	1.89
ANP2	1.90	1.09	1.15	0.87
ANP3	2.14	1.31	1.14	0.68
ANP4	4.25	1.32	-0.62	-0.77
SN1	3.04	1.36	0.17	-0.67
SN2	2.33	1.29	0.77	-0.18
SN3	2.31	1.22	0.81	0.19
SN4	2.38	1.38	0.86	0.02
SN5	3.50	1.62	-0.01	-1.13
SN6	2.21	1.01	4.00	17.31
SE1	2.50	1.29	0.55	-0.50
SE2	3.59	1.39	0.07	0.15
SE3	4.78	1.14	-1.03	1.02
SE4	3.64	1.44	-0.16	-0.73

APP = Approval of Physical Punishment; ANP = Approval of Non-Physical Punishment; SN = Subjective Norms; SE = Parenting Self-Efficacy.

^aItems in bold were removed from further analyses.

Table 2

Model Fit Indices for the PDEP Scale before the Revisions to the Questionnaire.

	AIC	BIC	RMSEA	M²
APP (original version)	15,172.79	15,322.30	.030	283.56
ANP (original version)	11,717.15	11,846.76	.050	749.16
SN (original version)	16,578.94	16,728.45	.040	595.68
SE (original version)	13,343.21	13,462.82	.050	495.15

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; RMSEA = Root Mean Square Error of Approximation; M² = Maydeu-Olivares and Joe's statistic; APP = Approval of Physical Punishment; ANP = Approval of Non-Physical Punishment; SN = Subjective Norms; SE = Parenting Self-Efficacy.

Table 3*Discrimination and Threshold Parameters for Approval of Physical Punishment Items.*

Item Code	Wording	a	b1	b2	b3	b4	b5
APP1	Sometimes a spank or swat is the best way to get a child to listen.	2.71	-0.04	0.50	0.82	1.72	2.28
APP2	Spanking is fine as long as the parent is not angry.	3.30	-0.17	0.32	0.73	1.33	1.73
APP3	Parents should have the right to decide whether to spank their children.	2.27	-0.54	-0.12	0.27	0.97	1.37
APP4	If parents don't use punishment, their children will be spoiled.	0.87	-0.23	0.23	0.60	1.22	1.64
APP5	It's ok to spank a 5-year-old's bottom if she does something dangerous.	1.89	-0.72	-0.22	0.22	0.87	1.38

a = item slope or discrimination parameter; b = between category threshold (difficulty parameter)

APP = Approval of Physical Punishment

Table 4*SD- χ^2 Values of Approval of Physical Punishment Items.*

	1.	2.	3.	4.	5.
1. APP1	-				
2. APP2	1.3	-			
3. APP3	7.2	5.8	-		
4. APP4	2.8	3.3	3.4	-	
5. APP5	4.0	4.3	3.6	1.9	-

Table 5*Discrimination and Threshold Parameters for Approval of Non-physical Punishment Items.*

Item Code	Wording	a	b1	b2	b3	b4	b5
ANP1	If a 14-year-old is failing in school, his parent should make him do extra chores.	2.00	0.22	1.99	3.68	5.08	6.25
ANP2	If a 12-year-old gets into trouble at school, her parent should punish her before she has a chance to give excuses.	1.31	-0.04	1.35	2.99	4.42	5.88
ANP3	If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay inside until he changes it.	0.89	-0.33	-0.86	2.02	3.06	3.82
ANP4	If children break the rules, their parents should take away privileges.	0.43	-3.27	-2.07	-1.30	2.31	1.57

a = item slope or discrimination parameter; b = between category threshold (difficulty parameter)

ANP = Approval of Non-physical Punishment

Table 6

SD- χ^2 Values of Approval of Non-Physical Punishment Items.

	1.	2.	3.	4.
1. ANP1	-			
2. ANP2	7.1	-		
3. ANP3	5.2	9.4	-	
4. ANP4	5.9	1.5	4.0	-

Table 7*Discrimination and Threshold Parameters for Subjective Norms Items.*

Item Code	Wording	a	b1	b2	b3	b4	b5
SN1	Young children who say "No!" to their parents are being defiant.	1.55	-0.94	-0.32	0.27	1.09	1.68
SN2	If children have tantrums, they are probably spoiled.	1.93	-0.37	0.25	0.82	1.45	2.13
SN3	Four-year-olds who interrupt adults are rude.	1.38	-0.46	0.26	0.93	1.57	2.15
SN4	A teenager who does not want to be seen with his mother should be ashamed of himself.	0.99	-0.36	0.19	0.81	1.27	1.77
SN5	If an 8-year-old uses bad words in front of his parents, this is a sign of disrespect.	0.78	-1.01	-0.43	-0.04	0.50	0.98

a = item slope or discrimination parameter; b = between category threshold (difficulty parameter)

SN = Subjective Norms

Table 8*SD- χ^2 Values of Subjective Norms Items.*

	1.	2.	3.	4.	5.
1. SN1	-				
2. SN2	7.1	-			
3. SN3	5.2	9.4	-		
4. SN4	5.9	1.5	4.0	-	
5. SN5	4.0	4.3	3.6	1.9	-

Table 9*Discrimination and Threshold Parameters for Self-Efficacy Items.*

Item Code	Wording	a	b1	b2	b3	b4	b5
SE1	Most people are better parents than I am.	1.20	-1.23	-0.21	1.57	3.01	4.87
SE2	I can solve most of my parenting challenges.	1.08	-2.60	1.60	-0.39	1.16	3.18
SE3	I have the skills to be a good parent.	1.17	-4.44	-3.62	1.84	0.74	1.02
SE4	As a parent, I often feel like I just don't know what to do.	1.36	-2.90	-1.51	0.47	1.38	2.64

a = item slope or discrimination parameter; b = between category threshold (difficulty parameter)

SE = Parenting Self-Efficacy

Table 10*SD- χ^2 Values of Self-Efficacy Items.*

	1.	2.	3.	4.
1. SE1	-			
2. SE2	13.1	-		
3. SE3	12.8	9.4	-	
4. SE4	5.9	1.5	4.0	-

Table 11

PDEP Questionnaire – First revisions following Preliminary Analyses.

Approval of Physical Punishment	Initial Item	Modifications
APP1	Sometimes a spank or swat is the best way to get a child to listen.	Sometimes a spank or <i>slap</i> is the best way to get <i>a young</i> child to listen.
APP2	Spanking is fine as long as the parent is not angry.	Spanking <i>or slapping</i> is fine as long as the parent is not angry.
APP3	Parents should have the rights to decide whether to spank their children.	Parents should have the rights to decide whether to spank <i>or slap</i> their children.
APP4	If parents don't use punishment, their children will be spoiled.	If parents don't <i>spank or slap</i> their children, they will be spoiled.
APP5	It's ok to spank a 5-year old's bottom if she does something dangerous.	No change
	New Items	
APP6	It's ok to slap a teenager's face if she uses bad words in front of her parents.	
APP7	Young children who aren't spanked will get into trouble when they are teenagers.	
APP8	Spanking teaches children right from wrong.	
APP9	If a 4-year-old won't leave a store, his parent should pull his arm and drag him out.	
APP10	It's ok to shake a 3-year-old to make her stop screaming.	

APP11	If a 2-year-old bites another child, it's ok to bite him to teach him how it feels.	
APP12	Children who are spanked learn to respect their parents.	
APP13	It's ok to slap a 2-year-old's hand for touching dangerous objects.	
Approval of Non-Physical Punishment	Initial Items	Modifications
ANP1	If a 14-year old is failing school, his parent should make him do extra chores.	No change
ANP2	If a 12-year-old gets into trouble at school, her parent should punish her before she has a chance to give excuses.	No change
ANP3	If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay inside until he changes it.	No change
ANP4	If children break the rules, their parents should take away privileges.	No change
	New Items	
ANP5	If a 3-year-old won't eat, his parent should make him stay at the table until he finishes his food.	
ANP6	If two sisters won't stop fighting, the parent should threaten to take away their phones.	
ANP7	A good way to teach a 13-year-old to dress properly is to embarrass her in front of her friends.	

ANP8	If a 6-year-old won't get dressed, his parent should threaten to take him out in his underwear.	
ANP9	A 10-year-old who hurts another child should be forced to apologize.	
ANP10	If a 3-year-old won't share his toys, he should be made to sit on a chair for 3 minutes.	
Subjective Norms	Initial Items	Modifications
SN1	Young children who say "No!" to their parents are being defiant.	<i>Two-year-olds</i> who say "no!" are being defiant.
SN2	If children have tantrums, they are probably spoiled.	<i>Usually, 3-year-olds</i> have tantrums because they are spoiled.
SN3	Four-year-olds who interrupt adults are rude.	No change
SN4	A teenager who does not want to be seen with his mother should be ashamed of himself.	<i>A 13-year-old</i> who does not want to be seen with his mother should be ashamed of himself.
SN5	If an 8-year-old uses bad words in front of his parents, this is a sign of disrespect.	No change
	New Items	
SN6	A 16-year-old who questions her parents' religious beliefs is being disrespectful.	
SN7	A 5-year-old child who won't go to sleep is being stubborn.	
SN8	A 6-year-old child who won't show affection to his grandmother is being rude.	
SN9	Boys should learn not to cry by the time they are 12 years old.	

SN10	Teenagers who live at home should always follow their parents' rules, even if they don't agree.	
Parenting Self-Efficacy	Initial Item	Modification
	Most people are better parents than I am.	Item deleted
SE1	I can solve most of my parenting challenges.	I can solve challenging situations without punishing this child.
SE2	I have the skills to be a good parent.	I believe I have the skills to be a good parent to this child.
SE3	As a parent, I often feel like I just don't know what to do.	I feel like I just don't know what to do with this child.
	New Items	
SE4	I can control my anger in challenging situations with this child.	
SE5	I feel badly about how I handle challenging situations with this child.	
SE6	I feel hopeful that I can improve my relationship with this child.	
SE7	In challenging situations, I end up shouting at this child.	
SE8	I believe that I will be able to handle future challenges as this child gets older.	
SE9	I feel hopeless about my relationship with this child.	

Table 12*Content Validity Assessment Content Experts (n = 8).*

Approval of Physical Punishment	Mean (SD) Representativeness	Percentage Representativeness	Mean (SD) Clarity	Percentage Clarity	Suggestions
1. Sometimes a spank or slap is the best way to get a young child to listen.	3.00 (.00)	(Slightly) 0% (Moderately) 0% (Highly) 100%	2.88 (.35)	(Low) 0% (Moderate) 12.5% (High) 87.5%	
2. Slapping or spanking is fine as long as the parent is not angry.	2.75 (.46)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.75 (.46)	(Low) 0% (Moderate) 25% (High) 75%	
3. Parents should have the right to decide whether to spank or slap their children.	2.63 (.52)	(Slightly) 0% (Moderately) 37.5% (Highly) 62.5%	2.75 (.46)	(Low) 0% (Moderate) 25% (High) 75%	• Parents can think they have a right to decide without actually endorsing corporal punishment. • Right v. approval might not be interchangeable.
4. If parents don't spank or slap,	2.88 (.35)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.75 (.46)	(Low) 0% (Moderate) 25% (High) 75%	•

their children will be spoiled.					
5. It's ok to spank a 5-year-old's bottom if she does something dangerous.	2.88 (.35)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.75 (.46)	(Low) 0% (Moderate) 25% (High) 75%	Danger is a qualifier--probably not ideal
6. It's ok to slap a teenager's face if she uses bad words in front of her parents.	2.75 (.46)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.63 (.46)	(Low) 12.5% (Moderate) 12.5% (High) 75%	Seems too specific
7. Young children who aren't spanked will get into trouble when they are teenagers.	2.88 (.35)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.75 (.46)	(Low) 0% (Moderate) 25% (High) 75%	
8. Spanking teaches children right from wrong.	3.00 (.00)	(Slightly) 0% (Moderately) 0% (Highly) 100%	3.00 (.00)	(Low) 0% (Moderate) 0% (High) 100%	
9. If a 4-year-old won't leave a store, his parent should pull his arm and drag him out.	2.75 (.46)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.63 (.74)	(Low) 12.5% (Moderate) 12.5% (High) 75%	"Pull" and "drag" less clear
10. It's ok to shake a 3-year-old to	2.75 (.46)	(Slightly) 0% (Moderately) 25%	2.75 (.46)	(Low) 0% (Moderate) 25%	

make her stop screaming.		(Highly) 75%		(High) 75%	
11. If a 2-year-old bites another child, it's ok to bite him to teach him how it feels.	2.63 (.46)	(Slightly) 0% (Moderately) 37.5% (Highly) 62.5%	2.63 (.52)	(Low) 0% (Moderate) 37.5% (High) 62.5%	
12. Children who are spanked learn to respect their parents.	3.00 (.00)	(Slightly) 0% (Moderately) 0% (Highly) 100%	3.00 (.00)	(Low) 0% (Moderate) 0% (High) 100%	
13. It's ok to slap a 2-year-old's hand for touching dangerous objects.	2.50 (.76)	(Slightly) 12.5% (Moderately) 25% (Highly) 62.5%	2.25 (.89)	(Low) 25% (Moderate) 25% (High) 50%	• Similar to earlier dangerous qualifier. • Dangerous too vague.

Approval of Non-Physical Punishment

1. If a 14-year-old is failing in school, his parent should make him do extra chores.	2.63 (.74)	(Slightly) 12.5% (Moderately) 12.5% (Highly) 75%	2.71(.49)	(Low) 0% (Moderate) 28.6% (High) 71.4%	
2. If a 12-year-old gets into trouble at school, her parent should punish her before she has a	2.25(.89)	(Slightly) 25% (Moderately) 0% (Highly) 100%	2.13(.83)	(Low) 25% (Moderate) 37.5% (High) 37.5%	• "Punish" is too open to interpretation-- could easily assume physical.

chance to give excuses.					- Punish could be interpreted as physical. - Specify what the punishment is.
3. If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay inside until he changes it.	2.63(.74)	(Slightly) 12.5% (Moderately) 12.5% (Highly) 75%	2.63(.74)	(Low) 12.5% (Moderate) 12.5% (High) 75%	These three teen ones are just odd responses.
4. If children break the rules, their parent should take away privileges.	2.75(.46)	(Slightly) 0% (Moderately) 25% (Highly) 75%	3.00(.00)	(Low) 0% (Moderate) 0% (High) 100%	Maybe too general
5. If a 3-year-old won't eat, his parent should make him stay at the table until he finishes his food.	2.88(.35)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.75(.46)	(Low) 0% (Moderate) 25% (High) 75%	
6. If two sisters won't stop fighting, the parent should threaten to take away their phones.	2.75(.46)	(Slightly representative) 0% (Moderately representative) 25% (Highly representative) 75%	2.75(.46)	(Low) 0% (Moderate) 25% (High) 75%	
7. A good way to teach a 13-year-old to dress properly is to embarrass her in front of her friends.	2.38(.91)	(Slightly representative) 25% (Moderately representative) 12.5%	2.25 (.89)	(Low) 25% (Moderate) 25% (High) 50%	Would omit

		(Highly representative) 62.5%			
8. If a 6-year-old won't get dressed, his parent should threaten to take him out in his underwear.	2.25(1.04)	(Slightly representative) 37.5% (Moderately representative) 0% (Highly representative) 62.5%	2.25 (1.04)	(Low) 37.5% (Moderate) 0% (High) 62.5%	Replace the word threaten to something else (e.g., say that they will take him out in his underwear)
9. A 10-year-old who hurts another child should be forced to apologize.	2.88(.35)	(Slightly representative) 0% (Moderately representative) 12.5% (Highly representative) 87.5%	2.63(.52)	(Low) 0% (Moderate) 37.5% (High) 62.5%	"Forced" seems off
10. If a 3-year-old won't share his toys, he should be made to sit on a chair for 3 minutes.	3.00(.00)	(Slightly representative) 0% (Moderately representative) 0% (Highly representative) 100%	3.00(.00)	(Low) 0% (Moderate) 0% (High) 100%	Most of these items are still harsh. Is that by intent rather than including items that capture more inductive forms of discipline?
Subjective Norms					
1. Two-year-olds who say "no!" are being defiant.	2.88 (.35)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.88 (.35)	(Low) 0% (Moderate) 12.5% (High) 87.5%	

2.	Usually, 3-year-olds have tantrums because they are spoiled.	2.88 (.35)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.75 (.46)	(Low) 0% (Moderate) 25% (High) 75%	
3.	Four-year-olds who interrupt adults are rude.	2.88 (.35)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.88 (.35)	(Low) 0% (Moderate) 12.5% (High) 87.5%	
4.	A 13-year-old who does not want to be seen with his mother should be ashamed of himself.	2.50 (.76)	(Slightly) 0% (Moderately) 50% (Highly) 50%	2.50 (.53)	(Low) 0% (Moderate) 50% (High) 50%	• Again, these teen ones feel off
5.	If an 8-year-old uses bad words in front of his parents, this is a sign of disrespect.	2.63 (.74)	(Slightly) 12.5% (Moderately) 12.5% (Highly) 75%	2.63 (.74)	(Low) 12.5% (Moderate) 12.5% (High) 75%	
6.	A 16-year-old who questions her parents' religious beliefs is being disrespectful.	2.50 (.76)	(Slightly) 12.5% (Moderately) 25% (Highly) 62.5%	2.50 (.76)	(Low) 12.5% (Moderate) 25% (High) 62.5%	
7.	A 5-year-old child who won't go to sleep is being stubborn.	2.88 (.35)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.75 (.46)	(Low) 0% (Moderate) 25% (High) 75%	
8.	A 6-year-old child who won't show affection to his	2.75 (.46)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.76 (.46)	(Low) 0% (Moderate) 25% (High) 75%	

grandmother is being
rude.

9. Boys should learn not to cry by the time they are 12 years old.	2.38 (.91)	(Slightly) 25% (Moderately) 12.5% (Highly) 62.5%	2.50 (.93)	(Low) 25% (Moderate) 0% (High) 75%
10. Teenagers who live at home should always follow their parents' rules, even if they don't agree.	2.63 (.52)	(Slightly) 0% (Moderately) 37.5% (Highly) 62.5%	2.63 (.52)	(Low) 0% (Moderate) 37.5% (High) 62.5%

Parenting Self-efficacy

1. I can solve challenging situations without punishing this child.	2.88 (.35)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.75 (.46)	(Low) 0% (Moderate) 25% (High) 75%
2. I believe I have the skills to be a good parent to this child.	2.50 (.76)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.88 (.35)	(Low) 0% (Moderate) 12.5% (High) 87.5%
3. I feel like I just don't know what to do with this child.	2.88 (.35)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.63 (.74)	(Low) 12.5% (Moderate) 12.5% (High) 75%
4. I can control my anger in	2.88 (.35)	(Slightly) 0% (Moderately) 12.5%	2.75 (.46)	(Low) 0% (Moderate) 25%

challenging situations with this child.			(Highly) 87.5%		(High) 75%	
5.	I feel badly about how I handle challenging situations with this child.	2.88 (.35)	(Slightly) 12.5% (Moderately) 12.5% (Highly) 75%	2.50 (.76)	(Low) 12.5% (Moderate) 25% (High) 62.5%	It doesn't seem to get at efficacy
6.	I feel hopeful that I can improve my relationship with this child.	2.63 (.74)	(Slightly) 12.5% (Moderately) 25% (Highly) 75%	2.38 (.92)	(Low) 25% (Moderate) 12.5% (High) 62.5%	- Not as clear for efficacy on conflict - There are things I can do to improve my relationship with this child
7.	In challenging situations, I end up shouting at this child.	2.75 (.46)	(Slightly) 12.5% (Moderately) 12.5% (Highly) 75%	2.63 (.74)	(Low) 12.5% (Moderate) 12.5% (High) 75%	
8.	I believe that I will be able to handle future challenges as this child gets older.	2.63 (.74)	(Slightly) 0% (Moderately) 12.5% (Highly) 87.5%	2.63 (.74)	(Low) 12.5% (Moderate) 12.5% (High) 75%	
9.	I feel hopeless about my relationship with this child.	2.75 (.71)	(Slightly) 25% (Moderately) 25% (Highly) 50%	2.13 (.99)	(Low) 37.5% (Moderate) 12.5% (High) 50%	Not convinced that hopelessness = low self-efficacy about handling conflict

Table 13*Content Validity Assessment Parents (n = 12).*

Approval of Physical Punishment	Mean (SD) Representativeness	Percentage Representativeness	Mean (SD) Clarity	Percentage Clarity	Suggestions
1. Sometimes a spank or slap is the best way to get a young child to listen.	2.33 (.65)	(Slightly) 8.3% (Moderately) 33.3% (Highly) 58.4%	2.50 (.67)	(Low) 8.3% (Moderate) 33.3% (High) 58.4%	• Change “the best way” • One word is enough- spank or slap • The best way should be changed to “will get a young child to listen”
2. Slapping or spanking is fine as long as the parent is not angry.	2.58 (.51)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.75 (.45)	(Low) 0% (Moderate) 25% (High) 75%	
3. Parents should have the right to decide whether to spank or slap their children.	2.42 (.67)	(Slightly) 8.3% (Moderately) 16.7% (Highly) 75%	2.67 (.65)	(Low) 8.3% (Moderate) 16.7% (High) 75%	• No need for both spank and slap if both means physical punishment

4. If parents don't spank or slap, their children will be spoiled.	2.25 (.75)	(Slightly) 8.3% (Moderately) 16.7% (Highly) 75%	2.67 (.65)	(Low) 8.3% (Moderate) 16.7% (High) 75%	
5. It's ok to spank a 5-year-old's bottom if she does something dangerous.	2.50 (.35)	(Slightly) 0% (Moderately) 16.7% (Highly) 83.3%	2.83 (.39)	(Low) 0% (Moderate) 16.7% (High) 83.3%	
6. It's ok to slap a teenager's face if she uses bad words in front of her parents.	2.42 (.79)	(Slightly) 8.3% (Moderately) 8.3% (Highly) 83.4%	2.75 (.62)	(Low) 8.3% (Moderate) 8.3% (High) 83.4%	• Replace he/she with they (also for other questions) • Use they instead of she • "When" they use bad words
7. Young children who aren't spanked will get into trouble when they are teenagers.	2.75 (.62)	(Slightly) 8.3% (Moderately) 8.3% (Highly) 83.4%	2.75 (.62)	(Low) 8.3% (Moderate) 8.3% (High) 83.4%	
8. Spanking teaches children right from wrong.	2.33 (.78)	(Slightly) 16.7% (Moderately) 33.3% (Highly) 50%	2.67 (.65)	(Low) 8.3% (Moderate) 16.7% (High) 75%	
9. If a 4-year-old won't leave a store, his parent should	2.33 (.78)	(Slightly) 16.7% (Moderately) 33.3% (Highly) 50%	2.67 (.65)	(Low) 8.3% (Moderate) 16.7% (High) 75%	• Drag him out?

pull his arm and drag him out.					
10. It's ok to shake a 3-year-old to make her stop screaming.	2.33 (.89)	(Slightly) 25% (Moderately) 16.7% (Highly) 58.3%	2.75 (.46)	(Low) 8.3% (Moderate) 8.3% (High) 83.4%	
11. If a 2-year-old bites another child, it's ok to bite him to teach him how it feels.	2.42 (.79)	(Slightly) 16.7% (Moderately) 25% (Highly) 58.3%	2.75 (.62)	(Low) 8.3% (Moderate) 8.3% (High) 83.4%	
12. Children who are spanked learn to respect their parents.	2.17 (.84)	(Slightly) 25% (Moderately) 8.3% (Highly) 66.7%	2.75 (.62)	(Low) 8.3% (Moderate) 8.3% (High) 83.4%	
13. It's ok to slap a 2-year-old's hand for touching dangerous objects.	2.50 (.76)	(Slightly) 25% (Moderately) 33.3% (Highly) 41.7%	2.75 (.89)	(Low) 0% (Moderate) 25% (High) 75%	Replace dangerous with another word "maybe" something that is not safe?
Approval of Non-Physical Punishment					
1. If a 14-year-old is failing in school, his parent should	2.42 (.67)	(Slightly) 8.3% (Moderately) 41.7% (Highly) 50%	2.83(.39)	(Low) 0% (Moderate) 16.7% (High) 83.3%	Extra not always clear

make him do extra chores.					
2. If a 12-year-old gets into trouble at school, her parent should punish her before she has a chance to give excuses.	2.58(.67)	(Slightly) 8.3% (Moderately) 25% (Highly) 66.7%	2.25(.75)	(Low) 16.6% (Moderate) 41.7% (High) 41.7%	- I would say before she could explain herself. - Punish should be changed because it is too general
3. If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay inside until he changes it.	2.67(.65)	(Slightly) 8.3% (Moderately) 16.7% (Highly) 75%	2.75(.45)	(Low) 0% (Moderate) 25% (High) 75%	- Change “stay inside” with other punishment - Hairstyle?
4. If children break the rules, their parents should take away privileges.	2.75(.45)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.58(.67)	(Low) 8.3% (Moderate) 25% (High) 66.7%	I would take away sleepover but if to go outside I would not take away physical activities take away mostly electronics

5. If a 3-year-old won't eat, his parent should make him stay at the table until he finishes his food.	2.75(.45)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.75(.45)	(Low) 0% (Moderate) 25% (High) 75%	
6. If two sisters won't stop fighting, the parent should threaten to take away their phones.	2.67(.49)	(Slightly) 0% (Moderately) 33.3% (Highly) 66.7%	2.75(.46)	(Low) 0% (Moderate) 25% (High) 75%	• Change sisters to siblings • Threaten? Obligated
7. A good way to teach a 13-year-old to dress properly is to embarrass her in front of her friends.	2.42(.79)	(Slightly) 16.7% (Moderately) 25% (Highly) 58.3%	2.75 (.45)	(Low) 0% (Moderate) 25% (High) 75%	• Not representative since is not non-physical punishment in my opinion is humiliation • To embarrass? • To make fun in front of friends

8. If a 6-year-old won't get dressed, his parent should threaten to take him out in his underwear.	2.50(.90)	(Slightly) 16.7% (Moderately) 16.7% (Highly) 66.7%	2.83 (.39)	(Low) 0% (Moderate) 16.7% (High) 83.3%	Again, this is humiliation not punishment
9. A 10-year-old who hurts another child should be forced to apologize.	2.75(.62)	(Slightly) 8.3% (Moderately) 8.3% (Highly) 83.3%	2.83(.39)	(Low) 0% (Moderate) 16.7% (High) 83.3%	
10. If a 3-year-old won't share his toys, he should be made to sit on a chair for 3 minutes.	2.75(.45)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.92(.29)	(Low) 0% (Moderate) 8.3% (High) 91.7%	"Chair for 3 minutes" can put other alternatives along with this option
Subjective Norms					
1. Two-year-olds who say "no!" are being defiant.	2.50 (.67)	(Slightly) 8.3% (Moderately) 33.3% (Highly) 58.4%	2.75 (.49)	(Low) 0% (Moderate) 25% (High) 75%	- What is defiant - Stubborn instead of defiant

2.	Usually, 3-year-olds have tantrums because they are spoiled.	2.50 (.67)	(Slightly) 8.3% (Moderately) 33.3% (Highly) 58.4%	2.67 (.49)	(Low) 0% (Moderate) 33.3% (High) 66.7%	• Tantrums?
3.	Four-year-olds who interrupt adults are rude.	2.42 (.79)	(Slightly) 16.7% (Moderately) 25% (Highly) 58.3%	2.92 (.28)	(Low) 0% (Moderate) 16.7% (High) 83.3%	
4.	A 13-year-old who does not want to be seen with his mother should be ashamed of himself.	2.50 (.90)	(Slightly) 16.7% (Moderately) 16.7% (Highly) 66.6%	2.92 (.28)	(Low) 0% (Moderate) 16.7% (High) 83.3%	
5.	If an 8-year-old uses bad words in front of his parents, this is a sign of disrespect.	2.50(.67)	(Slightly) 8.3% (Moderately) 33.3% (Highly) 58.4%	2.92 (.29)	(Low) 0% (Moderate) 8.3% (High) 91.7%	
6.	A 16-year-old who questions her parents' religious beliefs is being disrespectful.	2.42 (.79)	(Slightly) 16.7% (Moderately) 25% (Highly) 58.3%	2.75 (.62)	(Low) 8.3% (Moderate) 8.3% (High) 83.4%	• All those questions should use the plural "they"
7.	A 5-year-old child who won't go to sleep is being stubborn.	2.64 (.50)	(Slightly) 0% (Moderately) 36.4% (Highly) 63.6%	2.82(.41)	(Low) 0% (Moderate) 18.2% (High) 81.8%	
8.	A 6-year-old child who won't show	2.67 (.65)	(Slightly) 8.3% (Moderately) 16.7% (Highly) 75%	2.83 (.39)	(Low) 0% (Moderate) 16.7% (High) 83.3%	• "Love" instead of affection

affection to his grandmother is being rude.						Rude” probably use another word (could be that is mad)
9. Boys should learn not to cry by the time they are 12 years old.	2.58 (.79)	(Slightly) 16.7% (Moderately) 8.3% (Highly) 75%	2.75(.62)	(Low) 8.3% (Moderate) 8.3% (High) 83.4%		- Boys and girls are the same - Don’t agree with this item
10. Teenagers who live at home should always follow their parents’ rules, even if they don’t agree.	2.75 (.62)	(Slightly) 8.3% (Moderately) 8.3% (Highly) 83.4%	2.83 (.39)	(Low) 0% (Moderate) 16.7% (High) 83.3%		
Parenting Self-Efficacy						
1. I can solve challenging situations without punishing this child.	3.00 (0)	(Slightly) 0% (Moderately) 0% (Highly) 100%	3.00 (0)	(Low) 0% (Moderate) 0% (High) 100%		
2. I believe I have the skills to be a good parent to this child.	2.92 (.29)	(Slightly) 0% (Moderately) 8.3% (Highly) 91.7%	3.00 (0)	(Low) 0% (Moderate) 0% (High) 100%		

3.	I feel like I just don't know what to do with this child.	2.58 (.67)	(Slightly) 8.3% (Moderately) 25% (Highly) 66.7%	2.75 (.62)	(Low) 8.3% (Moderate) 8.3% (High) 83.4%	
4.	I can control my anger in challenging situations with this child.	2.92 (.29)	(Slightly) 0% (Moderately) 8.3% (Highly) 91.7%	3.00 (0)	(Low) 0% (Moderate) 0% (High) 100%	
5.	I feel badly about how I handle challenging situations with this child.	2.75 (.45)	(Slightly) 0% (Moderately) 25% (Highly) 75%	2.75 (.45)	(Low) 0% (Moderate) 25% (High) 75%	- I feel sad - I feel guilty instead of badly
6.	I feel hopeful that I can improve my relationship with this child.	2.92(.29)	(Slightly) 0% (Moderately) 8.3% (Highly) 91.7%	3.00 (0)	(Low) 0% (Moderate) 0% (High) 100%	
7.	In challenging situations, I end up	2.83 (.39)	(Slightly) 10% (Moderately) 6.7% (Highly) 83.3%	2.83 (.39)	(Low) 0% (Moderate) 16.7% (High) 83.3%	I end up maybe another verb?

shouting at this child.					
8. I believe that I will be able to handle future challenges as this child gets older.	2.92 (.29)	(Slightly) 0% (Moderately) 8.3% (Highly) 91.7%	3.00 (0)	(Low) 0% (Moderate) 0% (High) 100%	When the child grows up instead of gets older
9. I feel hopeless about my relationship with this child.	2.67 (.78)	(Slightly) 0% (Moderately) 16.7% (Highly) 83.3%	2.50 (.79)	(Low) 16.7% (Moderate) 16.7% (High) 66.6%	Hopeless or helpless?

Table 14

Changes to PDEP Items Based on the Content Experts' and Parents' Feedback.

Item	Initial	Modified
APP6	It's ok to slap a teenager's face if she uses bad words in front of her parents.	It's ok to slap a teenager's face if she insults her parent .
APP9	If a 4-year-old won't leave a store, his parent should pull his arm and drag him out.	If a 4-year-old won't leave a store, his parent should pull him out by his arm.
APP13	It's ok to slap a 2-year-old's hand for touching dangerous objects.	It's ok to slap a 2-year-old's hand for touching an object he was told not to touch .
ANP1	If a 14-year-old is failing in school, his parent should make him do extra chores.	If a 14-year-old is failing in school, his parent should make him do extra chores to teach him to work harder .
ANP2	If a 12-year-old gets into trouble at school, her parent should punish her before she has a chance to give excuses.	If a 12-year-old gets into trouble at school, her parent should scold her before she has a chance to give excuses.
ANP3	If a 16-year-old wears a hairstyle that his parent disapproves of, he should	If a 16-year-old wears clothing that his parent disapproves of, he should

	have to stay inside until he changes it.	have to stay inside until he changes it.
ANP4	If children break the rules, their parents should take away privileges.	If a 10-year-old breaks the rules, her parents should take away favourite activities, like watching TV or seeing friends.
ANP6	If two sisters won't stop fighting, the parent should threaten to take away their phones	If two teenage siblings won't stop fighting, their parent should threaten to take something away, such as their phones.
ANP8	If a 6-year-old won't get dressed, his parent should threaten to take him out in his underwear.	If a 6-year-old won't get dressed for school, his parent should make him stay in his room until he changes his mind.
ANP9	A 10-year-old who hurts another child should be forced to apologize.	A 7-year -old who hurts another child should be forced to apologize.
SN3	Four-year-olds who interrupt adults are rude.	Four-year-olds who interrupt adults should know better.
SN9	Boys should learn not to cry by the time they are 12 years old.	<u>Deleted from the final version</u>
SN10	Teenagers who live at home should always follow their	Teenagers should follow their parents' rules without

	parents' rules, even if they don't agree.	question, even if they don't agree.
SE6	I feel hopeful that I can improve my relationship with this child.	I feel confident that I can improve my relationship with this child.
SE7	In challenging situations, I end up shouting at this child.	In challenging situations, I am able to stop myself from shouting at this child.

APP = Approval of Physical Punishment; ANP = Approval of Non-Physical Punishment; SN = Subjective Norms;
SE = Self-Efficacy

Table 15

Demographic characteristics of parents participating in the online survey (n = 400).

	N	Valid %	Mean	SD
Age (Respondent)			35.18	4.64
Number of children			2.04	0.95
Age (Child)			7.12	4.96
<u>Gender</u>				
Female	375	93.8		
Male	21	5.3		
Trans	3	0.8		
Non-Binary	1	0.3		
<u>Marital Status</u>				
Married	274	68.5		
Living with a partner	68	17.0		
Single (never married)	27	6.8		
Separated	13	3.3		
Divorced	16	4.0		
Widowed	2	0.5		
<u>Race/Ethnicity</u>				
White	371	92.8		
Middle Eastern	1	0.3		
African Canadian	4	1.0		

Asian (e.g., Chinese, Korean)	6	1.5
South Asian (e.g., Indian)	3	0.8
First Nations	7	1.8
More than one ethnicity/race	8	2.0

Education

Less than High school	6	1.5
High school diploma	62	15.5
College diploma	110	27.5
Bachelor's degree	154	38.5
Master's degree	53	13.3
Doctoral degree or equivalent	15	3.8

Employment Status

Not employed	111	27.8
Employed part-time	57	14.0
Employed full-time	195	48.8
Self-employed	11	2.8
Disability	7	1.8
Parental Leave	19	4.8

Income

Less than \$29,999	41	10.3
\$30,000-\$49,999	59	14.5
\$50,000-\$89,999	134	33.6
\$90,000-\$119,999	85	21.3
\$120,000-\$150,000	41	10.3

Over \$150,000	40	10.0
----------------	----	------

Recruitment site (How did you hear about the study?)

Community organizations	1	0.3
-------------------------	---	-----

Friends	4	0.8
---------	---	-----

Social media	395	98.9
--------------	-----	------

Table 16

Demographic Characteristics of Parents Participating in the Test-Retest Study (n = 100).

	N	Valid %	Mean	SD
Age (Respondent)			36.31	7.32
Number of children			1.94	0.85
Age (Child)			8.12	6.69
<u>Gender</u>				
Female	94	94.0		
Male	5	5.0		
Non-Binary	1	1.0		
<u>Marital Status</u>				
Married	67	67.0		
Living with a partner	16	16.0		
Single (never married)	6	6.0		
Separated	6	6.0		
Divorced	5	5.0		
<u>Race/Ethnicity</u>				
White	93	93.0		
African Canadian	1	1.0		
Asian (i.e., Chinese, Korean)	2	2.0		
First Nations	4	4.0		

Education

Less than High school	3	3.0
High school diploma	11	11.0
College diploma	29	29.0
Bachelor's degree	38	38.0
Master's degree	15	15.0
Doctoral degree or equivalent	4	4.0

Employment Status

Not employed	22	22.0
Employed part-time	20	20.0
Employed full-time	43	43.0
Self-employed	2	2.0
Disability	3	3.0
Parental Leave	10	10.0

Income

Less than \$29,999	13	13.0
\$30,000-\$49,999	17	17.0
\$50,000-\$89,999	34	34.0
\$90,000-\$119,999	14	14.0
\$120,000-\$150,000	6	6.0
Over \$150,000	16	16.0

Table 17

Mean, Standard Deviation, Skewedness and Kurtosis of the Revised PDEP Questionnaire Items.

	Mean	SD	Skew	Kurtosis
APP1	2.23	1.64	1.01	-.54
APP2	2.25	1.68	1.05	-0.37
APP3	3.23	1.86	0.10	-1.52
APP4	1.47	1.00	2.06	6.99
APP5	2.64	1.86	0.62	1.23
APP6	1.29	0.83	3.58	14.08
APP7	1.54	1.14	2.04	4.96
APP8	2.22	1.64	1.05	-0.43
APP9	2.94	1.70	0.38	-1.29
APP10	1.07	0.35	6.34	50.72
APP11	2.64	1.86	0.62	1.23
APP12	2.17	1.60	1.14	-0.11
APP13	2.81	1.79	0.47	-1.31
ANP1	2.41	1.49	0.94	-0.23
ANP2	1.93	1.14	1.48	2.21
ANP3	3.36	1.74	0.09	1.14
ANP4	4.73	1.30	-0.99	0.41
ANP5	2.21	1.43	1.18	0.42
ANP6	4.43	1.42	0.12	0.24
ANP7	1.22	0.59	3.39	13.97
ANP8	3.42	1.70	0.02	-1.44
ANP9	4.49	1.40	-0.84	0.35
ANP10	2.41	1.49	0.12	0.24
SN1	2.29	1.45	1.19	0.43
SN2	1.96	1.22	1.51	1.71
SN3	1.83	1.19	1.62	1.91
SN4	2.61	1.47	0.85	-0.35

SN5	1.74	1.18	1.86	2.92
SN6	2.39	1.41	0.92	-0.77
SN7	2.83	1.53	0.58	-0.88
SN8	2.98	1.59	0.49	-0.92
SN9	3.09	1.65	0.34	-1.18
SE1	4.04	0.71	-0.74	1.60
SE2	3.81	0.73	0.12	0.24
SE3	2.60	0.79	0.45	0.64
SE4	4.48	0.71	-1.23	1.29
SE5	4.47	0.65	-0.95	0.32
SE6	2.24	0.94	0.32	-0.28
SE7	3.73	0.81	-0.65	0.54
SE8	4.01	0.79	-0.74	1.05
SE9	1.43	0.73	1.82	3.29

APP = Approval of Physical Punishment; ANP = Approval of Non-Physical Punishment;
 SN = Subjective Norms; SE = Parenting Self-Efficacy.

*Items in bold were removed from further analyses.

Table 18*Model Fit Indices for the PDEP Scales after Revisions.*

	AIC	BIC	RMSEA	M²
APP (11 items)	9,254.22	9,513.50	.060	2911.48
APP (10 items)	8,063.09	8,298.44	.060	2544.03
APP (9 items)	6,974.06	7,185.43	.060	1882.05
ANP (9 items)	10,547.11	10,762.52	.050	1813.16
ANP (8 items)	9,263.71	9,693.89	.040	1003.35
ANP (7 items)	8,136.88	8,304.42	.040	811.46
SN (9 items)	9,513.25	9,728.65	.050	1244.53
SE (9 items)	6,575.93	6,751.44	.050	534.73
SE (8 items)	5,902.10	6,058.01	.040	298.07

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion;
 RMSEA = Root Mean Square Error of Approximation; M² = Maydeu-Olivares and Joe's statistic

Table 19*Discrimination and Threshold Parameters for Approval of Physical Punishment Items.*

Item Code^a	Wording	a	b1	b2	b3	b4	b5
APP1	Sometimes a spank or a slap is the best way to get a young child to listen.	4.40	0.88	2.88	4.28	4.34	7.82
APP2	Spanking or slapping is fine as long as the parent is not angry.	5.33	1.05	3.35	5.18	5.55	7.91
APP3^b	Parents should have the right to decide whether to spank or slap their children.	2.29	-0.69	-0.14	0.30	0.59	1.36
APP4	If parents don't spank or slap their children, they will be spoiled.	3.66	2.54	3.47	4.63	4.84	5.67
APP5	It's ok to spank a 5-year old's bottom if she does something dangerous.	2.53	-0.14	1.05	1.75	2.24	3.91
APP7	Young children who aren't spanked will get into trouble when they are teenagers.	3.26	2.63	4.35	5.10	5.15	7.63
APP8	Spanking teaches children right from wrong.	4.97	1.06	3.36	4.89	5.07	8.33
APP9^b	If a 4-year old won't leave a store, his parent should pull him out by his arm	1.04	-1.25	2.45	2.18	3.75	4.91

APP11	If a 2-year old bites another child, it's ok to bite him to teach him how it feels.	2.53	-0.14	1.05	1.75	2.24	3.91
APP12	Children who are spanked learn to respect their parents.	4.83	1.03	3.08	4.88	5.15	7.63
APP13	It's ok to slap a 2-year old's hand for touching an object he was told not to touch.	2.28	-0.94	0.54	1.57	1.78	3.93

a = item slope or discrimination parameter; b = between category threshold (difficulty parameter)

APP = Approval of Physical Punishment

^aNote that APP6 "*It's ok to slap a teenager's face if she insults her parent*" and APP10 "*It's ok to shake a 3-year old to make her stop screaming*" were deleted prior to IRT analyses because were highly skewed.

^bItem in bold was deemed problematic and therefore deleted.

Table 20*SD- χ^2 Values of Approval of Physical Punishment Items.*

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. APP1	-										
2. APP2	8.7	-									
3. APP3	13.7	19.1	-								
4. APP4	3.4	8.8	15.4	-							
5. APP5	4.0	4.3	14.1	1.9	-						
6. APP6	5.0	4.5	14.2	0.3	1.2	-					
7. APP7	2.2	3.2	22.4	0.5	2.9	0.1	-				
8. APP8	8.0	4.3	1.9	3.8	5.0	2.4	4.7	-			
9. APP9	7.6	7.3	7.6	6.1	10.9	2.9	2.6	6.9	-		
10. APP10	2.3	2.0	29.3	2.9	3.9	7.6	1.9	1.3	2.1	-	
11. APP11	9.0	3.5	12.0	5.1	9.5	4.7	2.7	2.1	8.3	4.0	-

Table 21

Discrimination and Threshold Parameters for Approval of Non-Physical Punishment Items.

Item Code^a	Wording	a	b1	b2	b3	b4	b5
ANP1	If a 14-year-old is failing in school, his parent should make him do extra chores to teach him to work harder.	1.00	-0.67	0.71	1.75	1.88	3.61
ANP2	If a 12-year-old gets into trouble at school, her parent should scold her before she has a chance to give excuses.	1.19	-0.17	1.38	3.18	3.24	4.72
ANP3	If a 16-year-old wears clothing that his parent disapproves of, he should have to stay inside until he changes it.	1.11	-1.79	-0.54	0.46	0.70	2.32
ANP4^b	If a 10-year-old breaks the rules, her parent should take away favourite activities, like watching TV or seeing friends.	0.89	-4.01	-2.79	-1.80	-0.54	0.71
ANP5	If a 3-year old won't eat, his parent should make him stay at the table until he finishes his food.	1.46	-0.33	1.10	2.41	2.43	4.15
ANP6	If two teenage siblings won't stop fighting, their parent should threaten to	0.82	-3.38	-2.14	-1.27	-0.18	1.14

	take something away, such as their phones.						
ANP8	If a 6-year-old won't get dressed for school, his parent should make him stay in his room until he changes his mind.	1.15	-2.03	-0.57	0.40	0.51	2.55
ANP9^b	A 7-year old who hurts another child should be forced to apologize.	0.66	-3.21	-2.16	-1.50	-0.16	0.98
ANP10	If a 3-year-old won't share his toys, he should be made to sit on a chair for 3 minutes.	1.65	-0.54	0.97	2.13	2.23	3.79

a = item slope or discrimination parameter; b = between category threshold (difficulty parameter)

ANP = Approval of Non-Physical Punishment

^aNote that item ANP7 "A good way to teach a 13-year-old to dress properly is to embarrass her in front of her friend" was deleted prior to IRT analyses because because was highly skewed.

^bItems in bold were deemed problematic and therefore deleted.

Table 22*SD- χ^2 Values of Approval of Non-Physical Punishment Items.*

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. ANP1	-								
2. ANP2	6.8	-							
3. ANP3	2.5	3.2	-						
4. ANP4	1.3	6.3	11.0	-					
5. ANP5	5.6	1.1	7.3	1.0	-				
6. ANP6	6.4	1.5	6.3	24.9	2.9	-			
7. ANP8	0.4	4.2	0.1	1.8	4.8	4.2	-		
8. ANP9	3.4	2.5	5.6	13.1	3.3	6.5	0.3	-	
9. ANP10	5.8	3.5	3.6	1.5	9.4	7.6	3.6	1.3	-

Table 23*Discrimination and Threshold Parameters for Subjective Norms Items.*

Item Code^a	Wording	a	b1	b2	b3	b4	b5
SN1	Two-year-olds who say “No!” are being defiant.	1.78	-0.83	1.33	2.59	2.64	4.32
SN2	Usually, 3-year-olds have tantrums because they are spoiled.	1.77	-0.22	1.95	3.21	3.35	5.61
SN3	Four-year-olds who interrupt adults should know better.	1.38	-1.54	0.52	1.82	1.87	3.87
SN4	A 13-year-old who does not want to be seen with his mother should be ashamed of himself.	1.30	-0.87	0.65	2.09	2.27	4.36
SN5	If an 8-year-old uses bad words in front of his parent, this is a sign of disrespect.	1.11	-1.69	-0.25	0.91	1.15	2.78
SN6	A 16-year-old who challenges her parents’ religious beliefs is being disrespectful.	1.47	0.59	2.04	3.28	3.32	5.30
SN7	A 5-year-old who won’t go to sleep is being stubborn.	1.37	-1.76	0.13	1.46	1.52	3.73
SN8	A 6-year-old who won’t show affection to his grandmother is being rude.	2.05	0.38	2.28	3.63	3.76	7.03

SN10	Teenagers should follow their parents' rules without question, even if they don't agree.	1.00	-1.58	-0.20	1.19	1.33	2.81
------	--	------	-------	-------	------	------	------

a = item slope or discrimination parameter; b = between category threshold (difficulty parameter)

SN = Subjective Norms

^aNote that SN9 "*Boys should learn not to cry by the time are 12 years old*" was deleted after content validity analysis

Table 24*SD- χ^2 Values of Subjective Norms Items*

		1.	2.	3.	4.	5.	6.	7.	8.	9.
1.	SN1	-								
2.	SN2	3.0	-							
3.	SN3	2.1	1.5	-						
4.	SN4	1.6	0.7	0.3	-					
5.	SN5	1.8	1.8	1.0	1.0	-				
6.	SN6	1.0	0.1	6.3	1.3	0.5	-			
7.	SN7	3.1	1.7	0.1	2.0	0.4	1.2	-		
8.	SN8	0.7	1.5	5.6	2.4	0.4	2.4	2.3	-	
9.	SN10	4.3	2.5	3.6	3.5	1.4	1.2	0.8	3.2	-

Table 25*Discrimination and Threshold Parameters for Self-Efficacy Items*

Item Code	Wording	a	b1	b2	b3	b4
SE1	I can solve challenging situations without punishing this child.	1.88	-8.02	-4.47	-1.35	2.60
SE2	I believe I have the skills to be a good parent to this child.	1.62	-6.81	-5.68	-2.65	-0.46
SE3	I feel like I just don't know what to do with this child.	2.31	-6.65	-3.42	-0.24	2.56
SE4	I can control my anger in challenging situations with this child.	1.88	-7.42	-5.28	-2.47	1.82
SE5	I feel badly about how I handle challenging situations with this child.	2.21	-4.46	-0.29	3.63	5.94
SE6^a	I feel confident that I can improve my relationship with this child.	1.17	-2.90	-1.51	.47	1.38
SE7	In challenging situations, I am able to stop myself from shouting at this child.	1.73	-2.12	0.77	4.27	7.28
SE8	I believe that I will be able to handle future challenges as this child gets older.	2.12	-6.92	-5.23	-2.03	1.64
SE9	I feel hopeless about my relationship with this child.	2.35	1.28	3.94	6.16	9.22

a = item slope or discrimination parameter; b = between category threshold (difficulty parameter)

SE = Self-Efficacy

^aItems in bold was deemed problematic and therefore deleted from revised version.

Table 26*SD- χ^2 Values of Self-Efficacy Items*

		1.	2.	3.	4.	5.	6.	7.	8.	9.
1.	SE1	-								
2.	SE2	0.2	-							
3.	SE3	1.2	2.7	-						
4.	SE4	0.2	0.2	2.2	-					
5.	SE5	3.6	1.6	0.6	4.7	-				
6.	SE6	0.7	0.3	0.3	4.2	1.3	-			
7.	SE7	5.1	0.2	0.9	2.7	1.2	0.2	-		
8.	SE8	3.3	9.1	2.4	2.4	3.0	0.3	2.1	-	
9.	SE9	1.4	0.8	2.2	3.4	1.5	2.5	4.7	0.7	-

Table 27

Means and Standard Deviations of All Study Variables (Convergent Validity Assessment).

Variable	Mean	Standard Deviation	Skew	Kurtosis	Theoretical Range
Approval of Physical Punishment (PDEP)	20.15	6.08	0.47	-1.24	9-54
Approval of Non-Physical Punishment (PDEP)	19.42	11.06	-0.63	-0.05	7-42
Subjective Norms (PDEP)	33.44	12.23	-0.01	-0.65	9-54
Parenting Self-Efficacy (PDEP)	27.31	4.05	-0.90	1.92	8-40
ATS	37.03	10.36	0.78	-0.60	10-70
Child-Responsible Attributions	30.03	8.39	0.18	-0.43	9-54
Discipline Self-Efficacy	33.29	5.73	-0.38	-0.15	8-48

ATS = Attitudes Toward Spanking Scale; Child-Responsible= Parent Cognition Scale -Child-Responsible Attributions subscale ; PCS-Parent = Parent Cognition Scale –Parent-Causal Attributions subscale; Discipline Self-Efficacy- Discipline subscale of the Parenting Self-Efficacy Index

Table 28

Demographic Characteristics of Parents Participating in the Quasi-Experimental Evaluation Study (PDEP Experimental vs. Wait-List Control Group).

	All Participants (N = 83)		Experimental (n = 47)		Wai-list Control (n = 36)		t-test	Chi-squared
Age (Parent) <i>M, SD</i>	36.31	7.32	34.53	6.04	34.41	6.48	- 0.08	
Number of children <i>M, SD</i>	1.87	0.77	1.94	0.79	1.85	0.73	1.36	
Age (Children) <i>M, SD</i>	1.98	0.86	1.87	0.77	2.15	0.98	1.30	
<u>Gender, <i>n Valid %</i>^a</u>								2.43
Female	69	84.10	37	78.70	32	91.4		
Male	13	15.90	10	21.30	3	8.3		
<u>Marital Status, <i>n Valid %</i>^a</u>								2.10
Married	58	70.70	35	74.50	23	65.7		
Living with a partner	12	14.60	6	12.80	6	17.1		
Single (never married)	8	9.80	4	8.50	4	11.4		
Separated/Divorced	4	4.90	2	4.20	2	5.8		
<u>Race/Ethnicity, <i>n Valid %</i>^b</u>								3.85
White	61	75.30	33	71.70	28	80.00		
African Canadian	6	7.40	4	8.70	2	5.70		
Asian (i.e., Chinese, Korean)	2	2.50	2	4.30	2	5.70		
South Asian (e.g. India)	5	6.20	3	6.50	0	0		
Middle Eastern	4	4.90	3	6.50	1	2.90		
Latino	2	2.50	1	2.20	1	2.90		
More than one ethnicity	1	1.20	0	0	1	2.90		

<u>Education, n Valid ^a</u>							4.16
Less than High school	4	4.90	3	6.40	1	2.90	
High school diploma	6	7.30	4	8.50	2	5.70	
College diploma	23	28.00	15	31.90	8	22.90	
Bachelor's degree	31	37.80	15	31.90	16	45.70	
Postgraduate degree	18	22.00	10	21.30	8	22.90	
<u>Partner Education, n Valid %^c</u>							4.24
Less than High school	4	5.60	2	5.00	2	6.50	
High school diploma	14	19.70	10	25.00	4	12.90	
College diploma	22	31.00	10	25.00	12	38.70	
Bachelor's degree	18	25.40	10	25.00	8	25.80	
Postgraduate degree	13	18.30	8	20.00	5	16.10	
<u>Employment Status, n Valid ^a</u>							6.22
Not employed	30	36.50	16	34.10	14	40.10	
Employed outside the home	52	63.50	31	65.90	21	60.00	
<u>Partner Employment Status n Valid^c</u>							4.55
Not employed	6	8.45	4	10.0	2	6.40	
Employed	65	91.55	36	90.0	29	93.60	
<u>Income</u>							
Less than \$29,999	13	16.00	9	19.10	4	11.70	
\$30,000-\$49,999	8	10.20	4	8.50	4	11.70	
\$50,000-\$89,999	24	29.50	16	34.00	8	23.50	
\$90,000-\$119,999	17	20.90	10	21.40	7	20.70	
\$120,000-\$150,000	10	12.30	5	12.80	4	11.70	
Over \$150,000	9	11.10	2	4.20	7	20.70	

a: Full sample n = 82, PDEP Experimental group n = 47, Wait-list control group n = 35

b: Full sample n = 81 PDEP Experimental group n = 46, Wait-list control n = 35

c: Full sample n = 71, PDEP Experimental n = 40, Wait-list control group n = 31

Table 29*Unconditional/Null Growth Curve Models for Parenting Outcomes.*

Outcome	Random Effect	Variance Component	SD	Wald χ^2 (df)	Probability Wald χ^2 (df)	Level 2 ICC	Level 3 ICC
Approval of Physical Punishment	Intercept r_0	21.38	4.62	340.38 (75)	$p < .001$	.54	.05
	Residual e	17.92	4.23				
	Intercept μ_0	2.19	1.48	13.91 (7)	$p = .052$		
Approval of Non-Physical Punishment	Intercept r_0	11.99	3.46	250.42 (75)	$p < .001$	.44	.01
	Residual e	14.99	3.87				
	Intercept μ_0	0.24	0.60	8.71 (7)	$p = .273$		
Subjective Norms	Intercept r_0	19.31	4.39	271.23 (75)	$p < .001$	.47	.00
	Residual e	21.23	4.67				
	Intercept μ_0	0.02 15.36	0.11	5.58 (7)	$p = .500$		

Parenting Self-Efficacy	Intercept r_0		3.91	487.89 (75)	$p < .001$	.66	.08
	Residual e	7.92	2.81				
	Intercept μ_0	2.23	1.49	18.23 (7)	$p = .011$		
DERS Impulse							
	Intercept r_0	3.69	1.92	429.94 (75)	$p < .001$	.62	.00
	Residual e	1.50	2.25				
	Intercept μ_0	0.01	0.27	2.57 (7)	$p = .500$		
DERS Strategies							
	Intercept r_0	3.37	1.83	431.17 (75)	$p < .001$	.61	.00
	Residual e	2.11	1.45				
	Intercept μ_0	0.01	0.03	4.36 (7)			
DERS Awareness							
	Intercept r_0	1.27	2.23	307.22 (75)	$p < .001$	.54	.15
	Residual e	4.11	2.03				
	Intercept μ_0	1.61	1.27	24.32 (7)	$p = .001$		

Parenting Daily Hassles (Frequency)	Intercept r_0	11.75	3.42	518.86 (75)	$p < .001$	.68	.00
	Residual e	5.52	2.35				
	Intercept μ_0	0.11	0.33	10.50 (7)	$p = .162$		
Parenting Task	Intercept r_0	20.19	4.48	480.72 (75)	$p < .001$	.65	.00
	Residual e	10.61	3.27				
	Intercept μ_0	0.003	0.06	8.52 (7)	$p = .288$		
Challenging Behaviour	Intercept r_0	14.85	3.85	421.82 (75)	$p < .001$	.61	.02
	Residual e	9.31	3.05				
	Intercept μ_0	0.60	0.77	9.86 (7)	$p = .196$		
Problem Solving	Intercept r_0	19.95	4.47	269.74 (75)	$p < .001$	.48	.02
	Residual e	22.01	4.69				
	Intercept μ_0	1.12	1.06	10.63 (7)	$p = .154$		

Use of Physical Punishment	Intercept r_0	0.24	0.48	219.13 (75)	$p < .001$	.39	.09
	Residual e	0.36	0.60				
	Intercept μ_0	0.01	0.00 1	5.60 (7)	$p = .500$		
Use of Non-Physical Punishment	Intercept r_0	3.59	1.89	299.93 (75)	$p < .001$	.52	.05
	Residual e	3.26	1.80				
	Intercept μ_0	0.72	0.85	20.55 (7)	$p = .005$		
Positive Discipline	Intercept r_0	28.69	5.35	467.02 (75)	$p < .001$	.64	.05
	Residual e	15.66	3.95				
	Intercept μ_0	2.17	1.47	12.20 (7)	$p = .094$		
Supportive Parenting	Intercept r_0	4.57	2.13	318.53 (75)	$p < .001$	.53	.00
	Residual e	3.91	1.97				
	Intercept μ_0	0.01 5.17	0.04	7.88 (7)	$p = .342$		

Proactive Parenting	Intercept r_0		2.27	263.30 (75)	$p < .001$	.46	.00
	Residual e	6.06	2.46				
	Intercept μ_0	0.01	0.02	5.27 (7)	$p = .500$		
Setting Limits	Intercept r_0	1.54	1.24	202.12 (75)	$p < .001$	.36	.00
	Residual e	2.70	1.64				
	Intercept μ_0	0.01	0.01	5.65 (7)	$p = .500$		

Intercept r_0 = random effect of parent on the intercept (Level 2); Residual e = random error; Intercept μ_0 = random effect of facilitator on the intercept (Level 3); SD = standard deviation; df = degrees of freedom; ICC = intraclass correlation coefficient.

Table 30*Means (Standard Deviations) of All Variables at Pre-Program for the Two PDEP Groups*

Variable	Potential Range	<i>PDEP Experimental M (SD)</i>	<i>Wait-list Control M (SD)</i>	T-test	p-values¹
		Pre (n = 47)	Pre (n = 36)		
Approval of Physical Punishment	9-54	19.23 (7.75)	16.75 (6.06)	-1.58	.117
Approval of Non-Physical Punishment	7-42	18.65 (6.77)	15.88 (4.61)	-2.28	.025
Subjective Norms	9-54	22.96 (7.02)	18.27 (5.37)	-3.32	.001
Parenting Self-Efficacy	8-40	28.89 (5.65)	30.17 (4.95)	1.07	.287
DERS Impulse	3-15	5.74 (3.01)	4.92 (2.44)	-1.34	.346
DERS Strategies	3-15	5.40 (2.57)	5.36 (2.69)	1.35	.359
DERS Awareness	3-15	10.08 (3.53)	11.05 (2.83)	-0.07	.406
Parenting Daily Hassles Frequency Total score	20-80	45.62 (10.27)	44.40 (8.82)	0.58	.950
Parenting Task	8-40	18.10 (6.59)	17.17 (5.12)	1.06	.486
Challenging Behaviour	7-35	17.98 (5.58)	16.41 (4.64)	-1.36	.179

Problem Solving (Approach/Avoidance)	16-96	53.62 (7.85)	51.86 (6.49)	-1.08	.280
Use of Physical Punishment Use	4-20	4.53 (0.92)	4.36 (0.63)	-0.94	.348
Use of Non-Physical Punishment	4-20	8.70 (2.89)	7.43 (2.46)	-2.10	.038
Positive Discipline	9-45	27.56 (6.97)	29.34 (6.92)	1.16	.250
Supportive Parenting	6-42	35.49 (2.90)	36.22 (3.15)	0.89	.276
Proactive Parenting	6-42	32.69 (3.69)	35.55 (3.29)	-1.09	.272
Setting Limits	4-28	21.72 (1.95)	21.75 (2.28)	-1.11	.955

¹Bold values indicate statistically significant results using alpha less than .05.

Table 31.*Correlation Among All Outcome Variables at the Pre-PDEP Data Collection*

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Approval of Physical Punishment	-															
2. Approval of Non-Physical Punishment	.65***	-														
3. Subjective Norms	.53***	.74***	-													
4. Parenting Self-Efficacy	-.40***	.41***	.41***	-												
5. DERS-Impulse	-.021	.01	.054	-.12	-											
6. DERS-Strategies	-.04	.05	.099	-.17	.56***	-										
7. DERS-Awareness	-.01	-.03	-.42	-.10	-	.39***										
8. Problem Solving (Approach/Avoidance)	-.04	.04	-.05	-.20	-.02	-.05	-.01	-								
9. Parenting Daily	.24*	.22*	.23**	-.45***	.10	.09	-.16	.15	-							

Hassles																	
Frequency																	
Total score																	
10. Parenting Task	.29*	.24*	.34**	-.52***	.21	.32	-.21	.07	.77***	—							
11. Challenging Behaviour	.34*	.32*	.45***	-.60***	.24*	.12	-.26*	.09	.52***	.77***	-						
12. Use of Physical Punishment	.46***	.31*	.18	-.41***	.03	-.06	-.02	-.16	.23***	.25**	.29**	-					
13. Use of Non-Physical Punishment	.33***	.46**	.41***	-.57***	.01	-.07	.05	.03	.27***	.25**	.33**	.40**	-				
14. Positive Discipline	-.08	-.08	-.04	.06	-.04	.02	-.02	.01	.33***	.17	.18	.10	.34**	-			
15. Supporting Parenting	-.17	-.34**	-.31**	.47***	-.13	.04	.04	-.05	-.10	-.16	-.14	-.15	-.34**	.18	-		
16. Proactive Parenting	-.14	-.28**	-.28**	.38***	-.08	.12	.12	-.11	-.14	-.29**	-.31**	-.10	-.33**	.09	.65***	-	
17. Setting Limits	-.11	-.15	-.15	.42***	.11	.07	.07	-.01	-.20	-.26*	-.19	-.17	-.28**	.14	.63***	.73***	

*p < .05, ** p < .01, ***p < .001

Table 32.

Fixed Effects of the Conditional Growth Curve Models for All the Parenting Outcomes¹.

Outcome	Fixed Effects	Coefficient	Standard Error	t-ratio	p-value ²	Family-wise alpha $p < .003^3$
Approval of Physical Punishment	Intercept β_0	-0.01	1.24	-0.01	.992	
	Slope β_1	-2.64	0.76	-3.45	<.001	
	Slope β_2	-1.68	0.71	-2.36	.020	ns.
Approval of Non-Physical Punishment	Intercept β_0	-0.57	0.96	-0.58	.056	
	Slope β_1	-3.09	0.69	-4.45	<.001	
	Slope β_2	-1.47	0.51	-2.89	.005	ns.
Subjective Norms	Intercept β_0	0.08	1.17	0.07	.944	
	Slope β_1	-4.34	0.76	-5.71	<.001	
	Slope β_2	-2.33	0.65	-3.55	<.001	
Parenting Self-Efficacy	Intercept β_0	-0.66	0.92	-0.71	.480	
	Slope β_1	1.01	0.50	1.99	.049	
	Slope β_2	0.77	0.43	1.75	.085	

DERS Impulse						
	Intercept β_0	0.67	0.47	1.35	.180	
	Slope β_1	-0.37	0.31	1.23	.220	
	Slope β_2	-0.61	0.29	-2.07	.041	ns.
DERS Strategies						
	Intercept β_0	-0.21	0.45	-0.48	.635	
	Slope β_1	-0.06	0.30	-0.19	.847	
	Slope β_2	-0.08	0.28	-0.28	.782	
DERS Awareness						
	Intercept β_0	-0.48	0.64	-0.76	.451	
	Slope β_1	0.22	0.45	0.49	.626	
	Slope β_2	-0.02	0.39	-0.05	.957	
Parenting Daily Hassles (Frequency)						
	Slope β_2	0.78	0.44	1.68	.906	
	Slope β_1	0.06	0.91	0.05	.959	
	Slope β_2	-0.91	1.04	-0.79	.431	
Parenting Task						
	Intercept β_0	0.37	1.07	0.35	.728	
	Slope β_1	-0.87	0.62	-1.40	.162	
	Slope β_2	-0.91	0.63	-1.42	.157	
Challenging Behaviour		0.63	0.98	0.64		

	Intercept β_0				.519	
	Slope β_1	-1.34	0.54	-2.48	.014	ns.
	Slope β_2	-1.32	0.56	-2.35	.020	ns.
Problem Solving	Intercept β_0	1.38	1.23	1.12	.267	
	Slope β_1	0.02	0.91	0.02	.983	
	Slope β_2	0.73	0.86	0.84	.339	
Use of Physical Punishment	Intercept β_0	-0.01	0.02	-0.26	.875	
	Slope β_1	-0.02	0.03	-0.82	.074	
	Slope β_2	-0.18	0.12	-1.55	.124	
Use of Non-Physical Punishment	Intercept β_0	-0.27	0.52	-0.51	.612	
	Slope β_1	-1.71	0.29	-5.89	<.001	
	Slope β_2	-0.72	0.30	-2.41	.017	ns.
Positive Discipline	Intercept β_0	0.04	1.32	0.03	.976	
	Slope β_1	1.42	0.76	1.86	.066	
	Slope β_2	1.08	0.68	1.57	.121	
Supporting Parenting	Intercept β_0	0.28	0.56	0.50	.618	
		1.14	0.37	3.11		

	Slope β_1				.003	ns.
	Slope β_2	0.93	0.40	2.25	.027	ns.
Proactive Parenting						
	Intercept β_0	0.28	0.63	0.47	.653	
	Slope β_1	1.06	0.41	2.56	.011	ns.
	Slope β_2	0.40	0.43	0.94	.346	
Setting Limits						
	Intercept β_0	0.28	0.36	0.77	.445	
	Slope β_1	0.21	0.30	0.70	.484	
	Slope β_2	0.13	0.31	0.45	.651	

Intercept β_0 = mean change in parenting outcomes at baseline (pre-PDEP) between the PDEP experimental and wait-list control group; Slope β_1 = mean change in parenting outcomes from pre- to post-PDEP between the PDEP experimental and wait-list control group; Slope β_2 = mean change in parenting outcomes from post-PDEP to 1-month follow-up between the PDEP experimental and wait-list control group.

¹Note that the models were controlled for randomization (randomized vs. non-randomized) and average number of services across the three time points.

²Bold values indicate statistically significant results using alpha less than .05.

³Non-statistically significant results following Bonferroni adjustment.

Table 33.

Fixed Effects of the Conditional Growth Curve Models for the PDEP Experimental group across the Parenting Outcomes with Treatment Attendance as Moderator.

Outcome	Fixed Effects	Coefficient	Standard Error	t-ratio	p-value
Approval of Physical Punishment	Treatment Attendance on Intercept β_0	-0.42	0.58	-0.73	.469
	Treatment Attendance on Slope β_1	-0.44	0.44	0.99	.324
	Treatment Attendance on Slope β_2	0.24	0.37	0.64	.524
Approval of Non-Physical Punishment	Treatment Attendance on Intercept β_0	-0.85	0.43	-2.00	.052
	Treatment Attendance on Slope β_1	-0.27	0.44	-0.61	.547
	Treatment Attendance on Slope β_2	0.06	0.27	0.22	.829
Subjective Norms	Treatment Attendance on Intercept β_0	-0.56	0.60	-0.93	.356
	Treatment Attendance on Slope β_1	0.44	0.45	0.98	.335

	Treatment Attendance on Slope β_2	0.23	0.37	0.64	.528
Parenting Self- Efficacy	Treatment Attendance on Intercept β_0	0.80	0.50	1.60	.116
	Treatment Attendance on Slope β_1	-0.08	0.33	-0.25	.805
	Treatment Attendance on Slope β_2	0.21	0.23	0.91	.367
DERS Impulse	Treatment Attendance on Intercept β_0	-0.14	0.26	-0.55	.582
	Treatment Attendance on Slope β_1	-0.12	0.17	-0.68	.499
	Treatment Attendance on Slope β_2	0.05	0.16	0.31	.756
DERS Strategies	Treatment Attendance on Intercept β_0	-0.05	0.24	-0.12	.831
	Treatment Attendance on Slope β_1	0.11	0.15	0.71	.476
	Treatment Attendance on Slope β_2	0.18	0.15	1.27	.210
DERS Awareness	Treatment Attendance on Intercept β_0	0.18	0.31	0.59	.556
	Treatment Attendance on Slope β_1	-0.19	0.23	-0.82	.413

	Treatment Attendance on Slope β_2	-0.01	0.24	-0.03	.979
Parenting Daily Hassles (Frequency)	Treatment Attendance on Slope β_0				
		-0.21	0.42	-0.50	.622
	Treatment Attendance on Slope β_1	0.14	0.23	0.62	.539
	Treatment Attendance on Slope β_2	0.22	0.19	1.16	.252
Parenting Task	Treatment Attendance on Intercept β_0	0.03	0.56	0.04	.968
	Treatment Attendance on Slope β_1	-0.39	0.38	-1.00	.320
	Treatment Attendance on Slope β_2	-0.30	0.35	-0.86	.397
Challenging Behaviour	Treatment Attendance on Intercept β_0	0.37	0.50	0.73	.466
	Treatment Attendance on Slope β_1	-0.42	0.34	-1.22	.229
	Treatment Attendance on Slope β_2	-0.35	0.25	-1.44	.158
Problem Solving	Treatment Attendance on Intercept β_0	1.38	1.23	1.12	.267
	Treatment Attendance on	0.02	0.91	0.02	.983

	Slope β_1				
	Treatment Attendance on Slope β_2	0.73	0.86	0.84	.339
Use of Physical Punishment	Treatment Attendance on Intercept β_0	-0.08	0.07	-1.12	.267
	Treatment Attendance on Slope β_1	0.03	0.06	0.65	.518
	Treatment Attendance on Slope β_2	0.01	0.05	0.17	.859
Use of Non-Physical Punishment	Treatment Attendance on Intercept β_0	-0.29	0.25	-1.18	.245
	Treatment Attendance on Slope β_1	0.13	0.15	0.86	.396
	Treatment Attendance on Slope β_2	-0.15	0.15	-0.98	.334
Positive Discipline	Treatment Attendance on Intercept β_0	0.75	0.70	1.06	.293
	Treatment Attendance on Slope β_1	0.80	0.39	2.00	.052
	Treatment Attendance on Slope β_2	0.32	0.33	0.95	.347
Supporting Parenting	Treatment Attendance on Intercept β_0	0.30	0.26	1.16	.253
	Treatment Attendance on	0.28	0.22	1.27	.209

	Slope β_1				
	Treatment Attendance on Slope β_2	0.11	0.24	0.47	.647
Proactive Parenting	Treatment Attendance on Intercept β_0	0.30	0.28	1.05	.300
	Treatment Attendance on Slope β_1	0.27	0.25	1.09	.283
	Treatment Attendance on Slope β_2	0.11	0.27	0.41	.684
Setting Limits	Treatment Attendance on Intercept β_0	0.21	0.17	1.29	.203
	Treatment Attendance on Slope β_1	0.25	0.16	1.57	.123
	Treatment Attendance on Slope β_2	0.19	0.17	1.09	.278

Intercept β_0 = cross-level interaction between baseline score in parenting outcomes for PDEP experimental and treatment attendance; Slope β_1 = cross-level interaction of between the mean change in parenting outcomes from pre- to post-PDEP between the PDEP experimental and wait-list control group; Slope β_2 = mean change in parenting outcomes from post-PDEP to 1-month follow-up between the PDEP experimental and wait-list control group

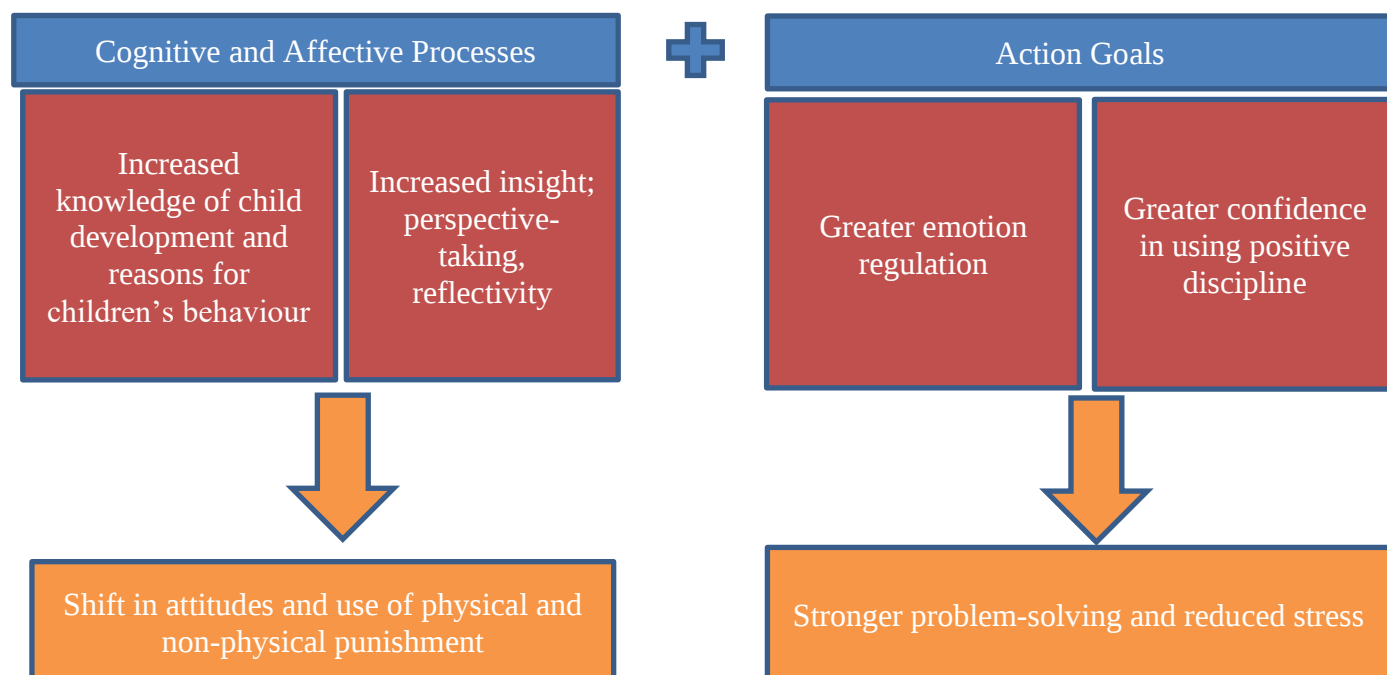


Figure 1. PDEP Conceptual Model

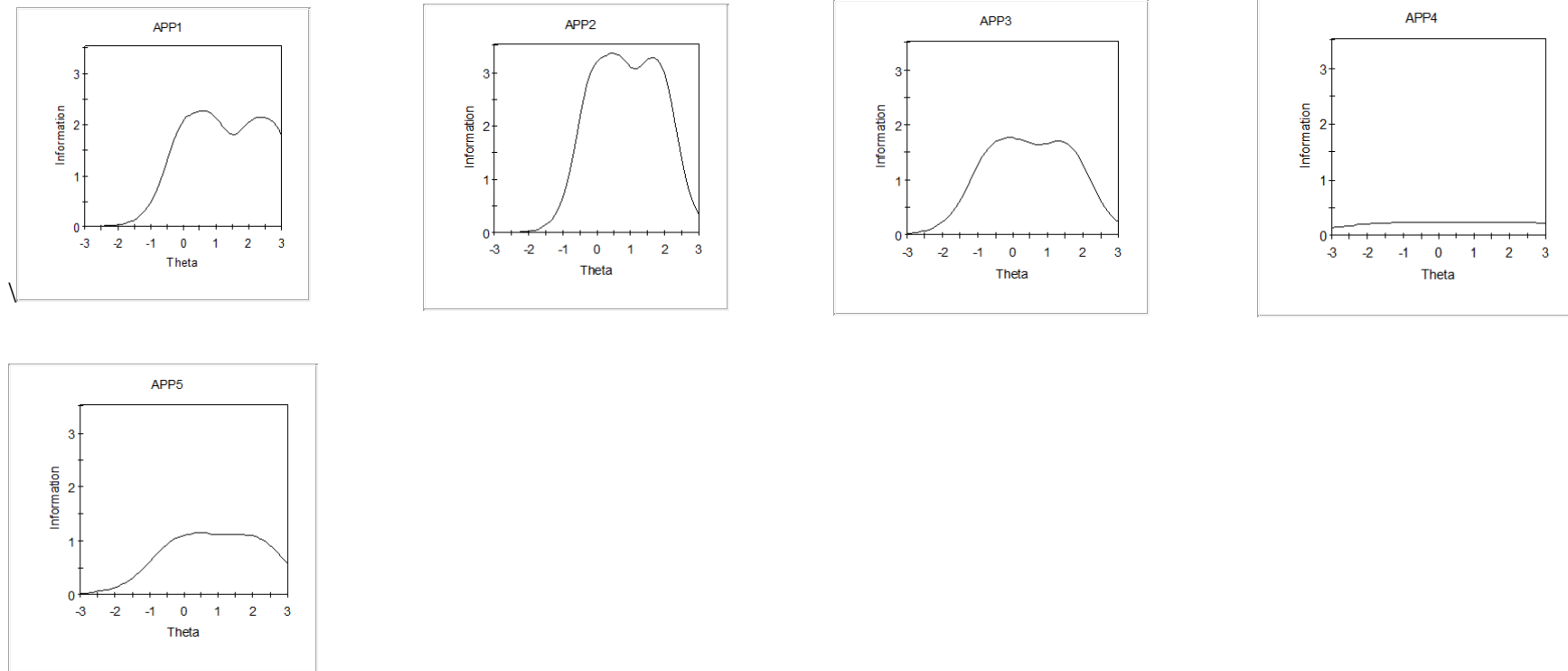


Figure 2. Information Functions for Approval of Physical Punishment Items (PDEP original version).

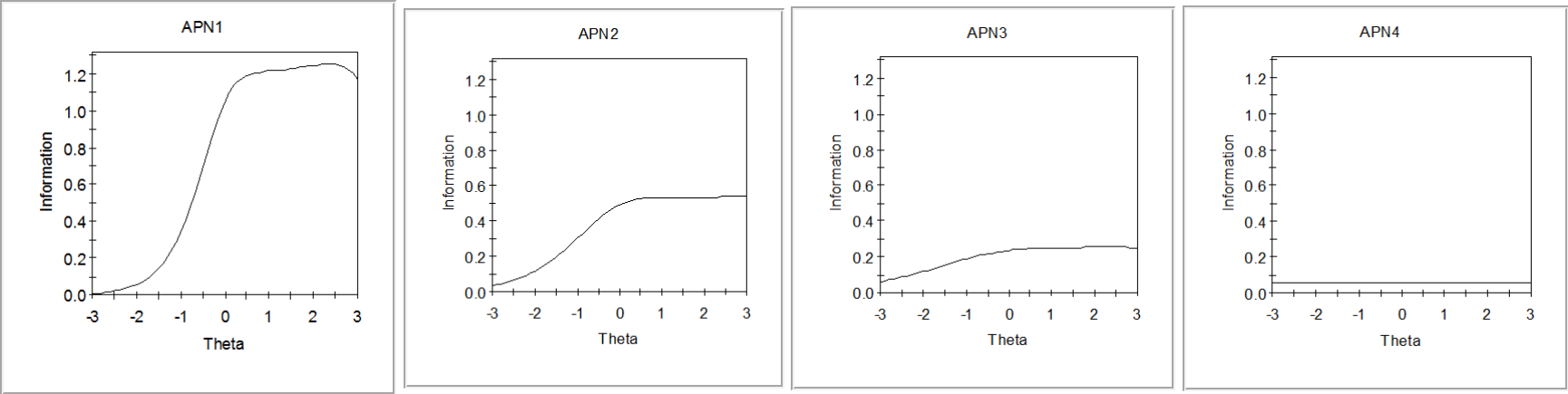


Figure 3. Information Functions for Approval of Non-physical Punishment Items (PDEP original version)

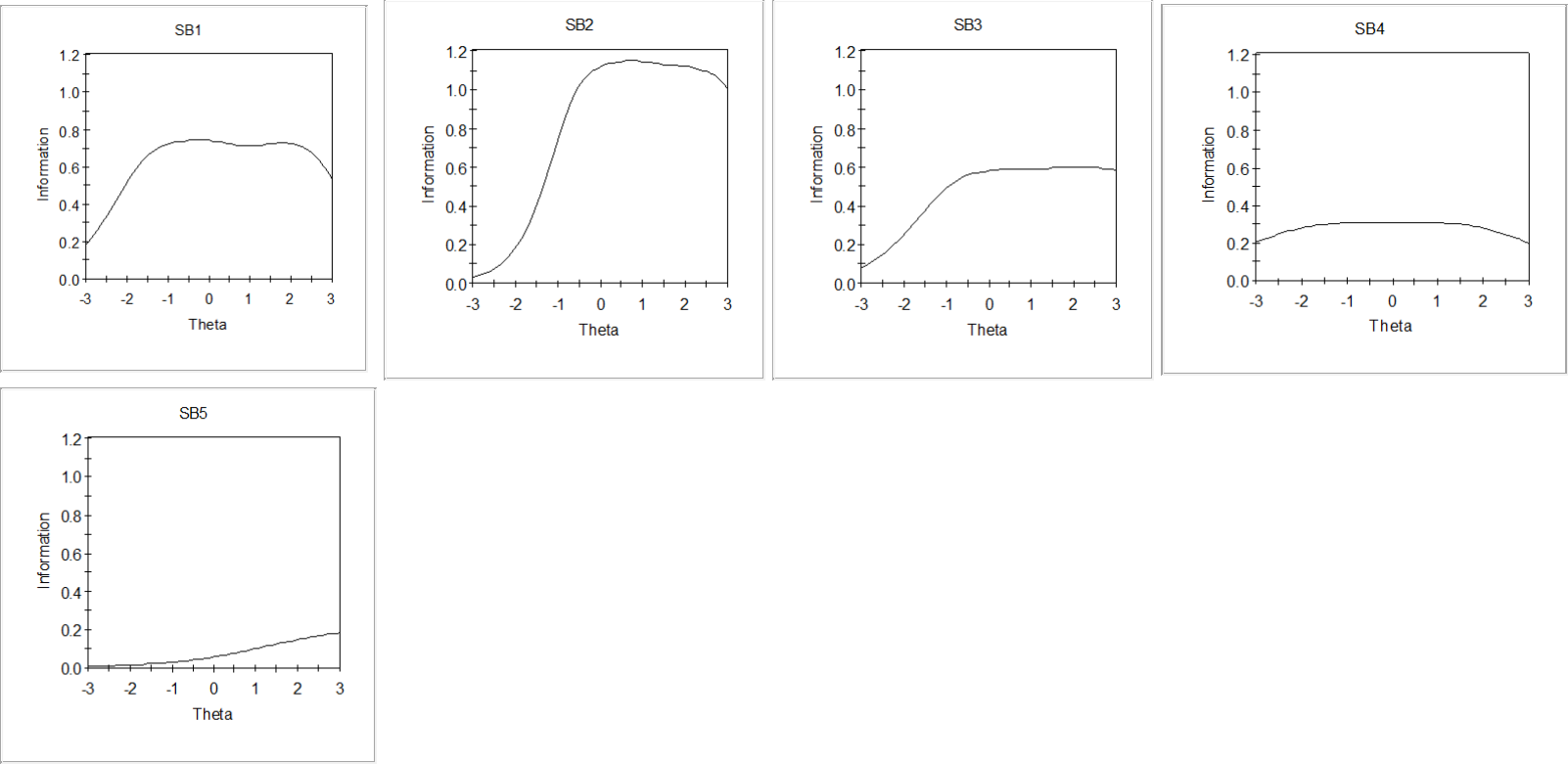


Figure 4. Information Functions for Subjective Norms Items (PDEP original version).

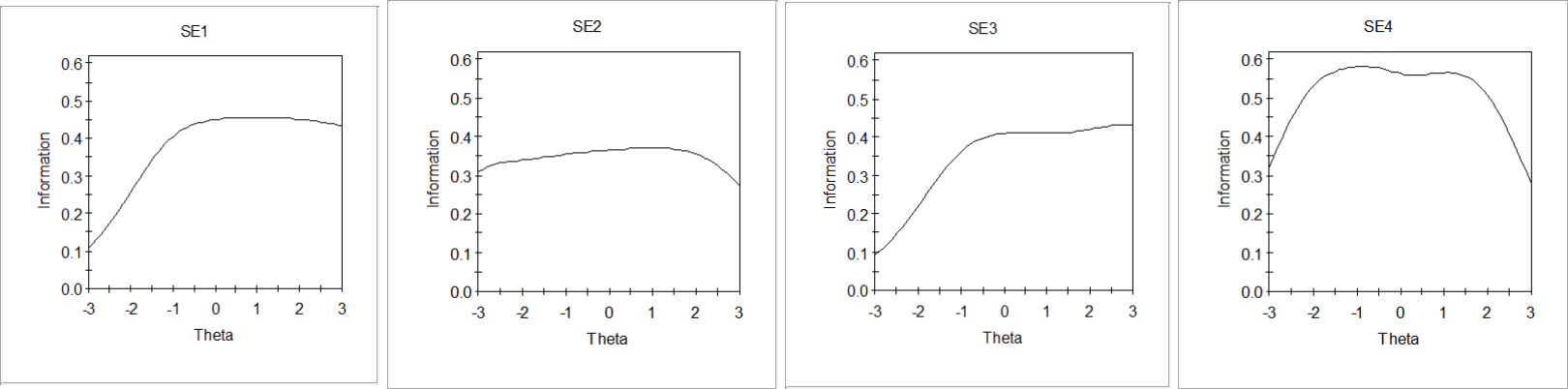


Figure 5. Information Functions for Self-Efficacy Items (PDEP original version)

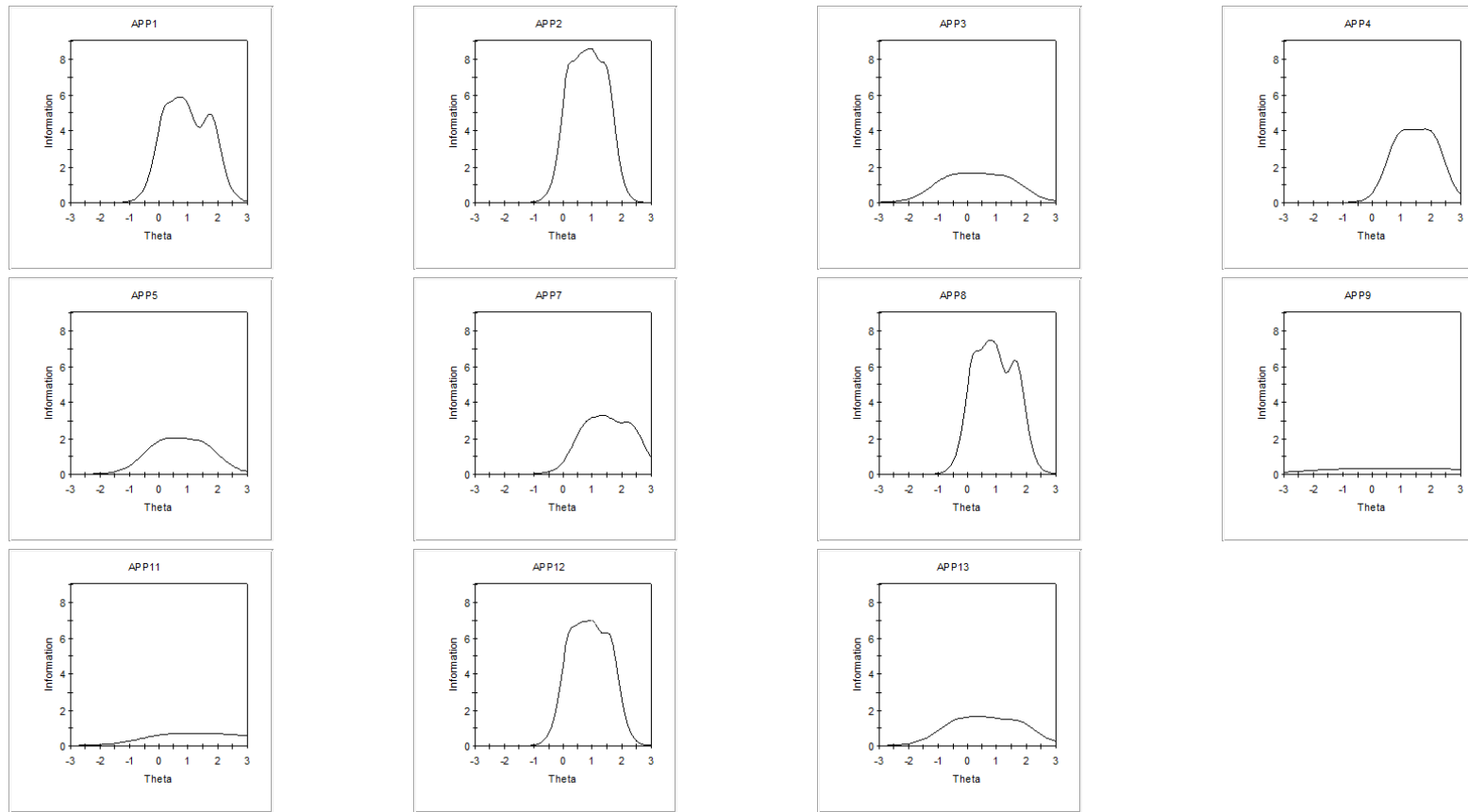


Figure 6. Information Functions for Approval of Physical Punishment Items (PDEP revised version).

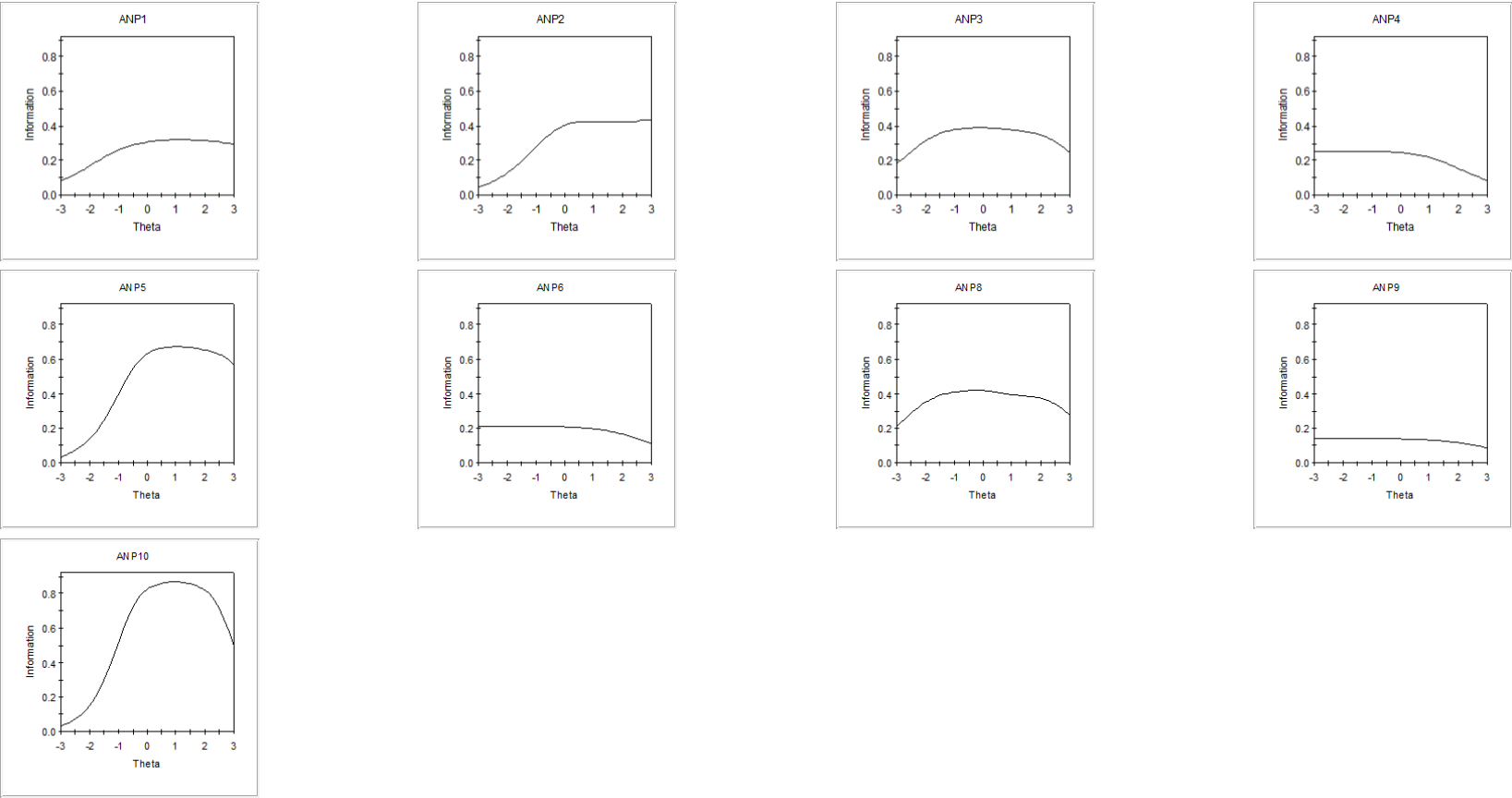


Figure 7. Information Functions for Approval of Non-Physical Punishment Items (PDEP revised version).

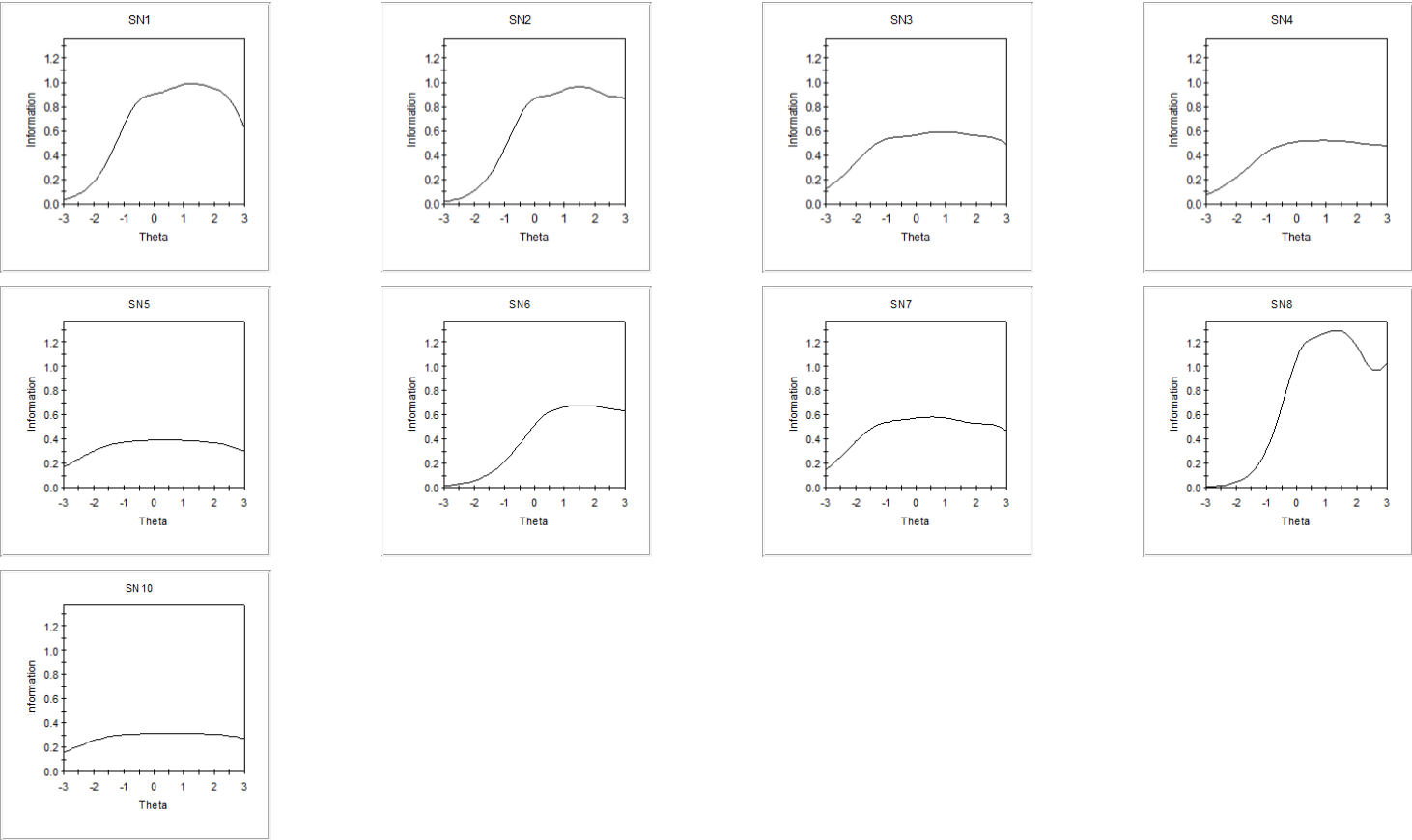


Figure 8. Information Functions for Approval of Subjective Norms (PDEP revised version).

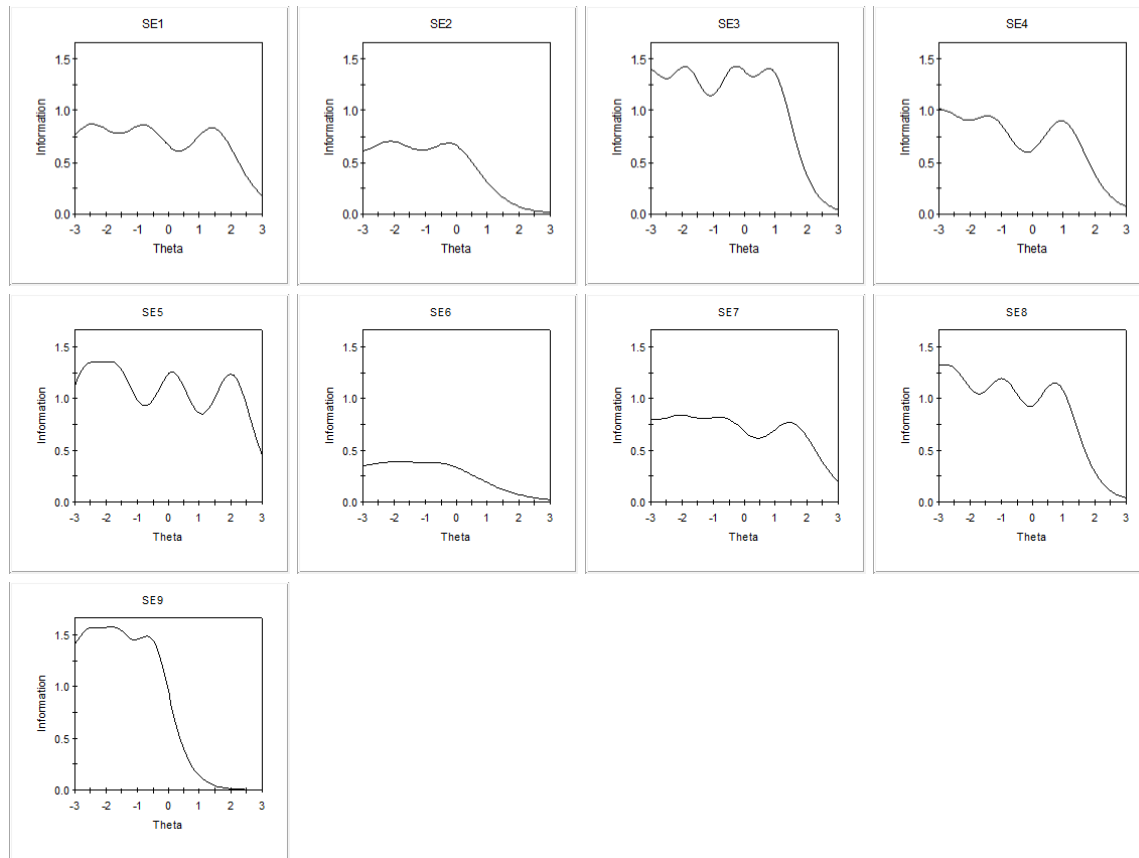


Figure 9. Information Functions for Self-Efficacy Items (PDEP revised version).

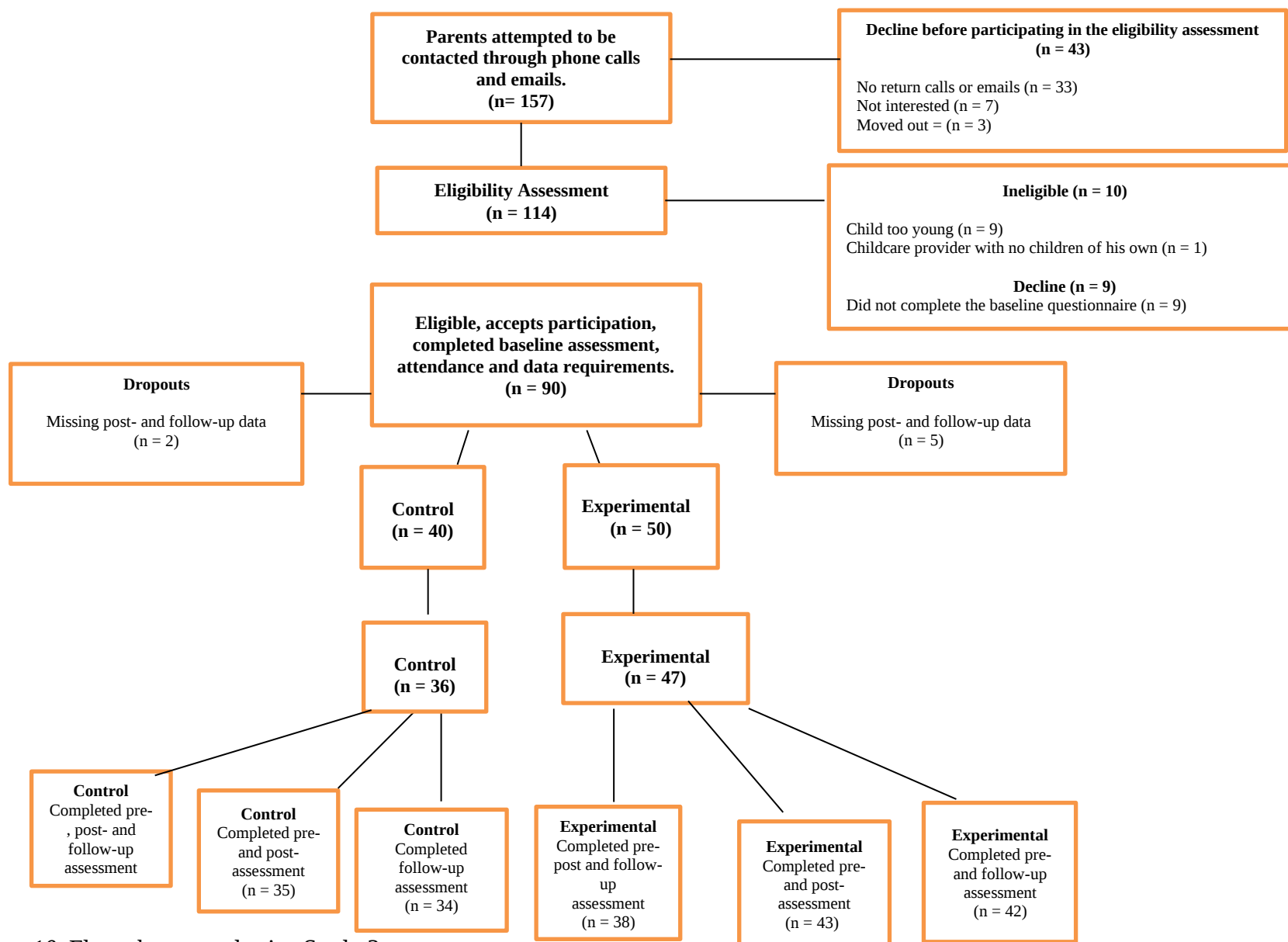


Figure 10. Flow chart sample size Study 3

Appendix A:

Ethics Approval Study 1 (Objective 1)

File Number: H11-15-09

Date (mm/dd/yyyy): 11/30/2015



Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Ethics Approval Notice Health Sciences and Science REB

Principal Investigator / Supervisor / Co-investigator(s) / Student(s)

<u>First Name</u>	<u>Last Name</u>	<u>Affiliation</u>	<u>Role</u>
Elisa	Romano	Social Sciences / Psychology	Supervisor
Elena	Gallitto	Arts / Others	Student Researcher

File Number: H11-15-09

Type of Project: Other

Title: Evaluation of the Factor Structure of the Positive Discipline in Everyday Parenting (PDEP) Questionnaire

Approval Date (mm/dd/yyyy)	Expiry Date (mm/dd/yyyy)	Approval Type
11/30/2015	11/29/2016	

(Ia: Approval, Ib: Approval for initial stage only)

Special Conditions / Comments:
N/A

Appendix B:

Ethics Approval Study 1 (Objective 2)

Université d'Ottawa

Bureau d'éthique et d'intégrité de la recherche



University of Ottawa

Office of Research Ethics and Integrity

CERTIFICAT D'APPROBATION ÉTHIQUE | CERTIFICATE OF ETHICS APPROVAL

Numéro du dossier / Ethics File Number

H-02-18-352

Titre du projet / Project Title

Evaluating the Validity and
Reliability of the Positive
Discipline in Everyday Parenting
(PDEP) Questionnaire

Type de projet / Project Type

Thèse de doctorat / Doctoral
thesis

Statut du projet / Project Status

Approuvé / Approved

Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)

05/03/2018

Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)

04/03/2019

Équipe de recherche / Research Team

Chercheur / Researcher

Affiliation

Role

Elena GALLITTO

École de psychologie / School of Psychology

Chercheur Principal / Principal Investigator

Elisa ROMANO

École de psychologie / School of Psychology

Superviseur / Supervisor

Conditions spéciales ou commentaires / Special conditions or comments

Appendix C:**Recruitment Notice Study 1, Objective 2 (Content Validity Experts)**

Hello,

I am a graduate student at the University of Ottawa under the supervision of Dr. Romano from the School of Psychology at the University of Ottawa. For my Ph. D. thesis, I am conducting a study to examine the content validity of a questionnaire that is currently being used to assess the outcomes of the Positive Discipline in Everyday Parenting program (PDEP). For more information about the program please visit the website: <http://www.positivedisciplineeveryday.com>.

Should you decide to participate, you will be invited to complete a questionnaire that will take approximately 20-30 minutes. The questionnaire is available on-line through Qualtrics and can be completed from any computer at a place of your choice. Once you begin the questionnaire, it will not be possible to close the study's webpage and return to it at a later point. However, you do have an 8-hour timeframe to complete the questionnaire.

You are free to refuse to participate or to withdraw from the study at any time without penalty.

We have contacted you because you have expertise in the area of parenting, child development, psychometrics, and/or the PDEP program.

If you are interested, please email me, and I will send you the link to the questionnaire along with the consent form.

Appendix D:

Study 1, Objective 2 (Consent Form Experts)

Study 1-Objective 1- Content Validity Forms Expert Panel)

Study title: Content Validity Assessment of the Positive Discipline in Everyday Parenting Questionnaire

The ethical aspects of the study have been approved by the Office of Research Ethics and Integrity of the University of Ottawa (REB # H-02-18-352, approved March 3, 2018).

Study Investigators

Elena Gallitto (Ph.D. Candidate, School of Psychology, University of Ottawa).
Dr. Elisa Romano (Professor, School of Psychology, University of Ottawa).

Should you have any questions or require additional information, please contact the Protocol Officer for Ethics in Research, Office of Research Ethics and Integrity, University of Ottawa.

Any questions about the current study may be addressed to Elena Gallitto.

Purpose

The purpose of this study is to examine the content validity of the Positive Discipline in Everyday Parenting (PDEP) questionnaire. The PDEP questionnaire measures four main constructs: 1) Approval of Physical Punishment; 2) Approval of Non-Physical Punishment; 3) Subjective Norms; and 3) Parenting Self-Efficacy. We are asking you to rate the representativeness and clarity of the questionnaire items and to complete a brief demographic questionnaire. It will take about 20-30 minutes to complete the study.

Potential risk/discomfort

There are no anticipated psychological or physical risks in this study. However, if you experience some anxiety or discomfort for any unanticipated reason, you may stop participating in the study.

Anonymity/confidentiality

Be assured that no identifiable information will be gathered. The information you provide during this study will be kept strictly confidential and anonymous. Access to the data you provide will be limited to the researchers listed above. Your name will not be revealed in any report regarding this research. To ensure the security and confidentiality of online communication in the study, the online study is equipped with encrypted communication features and a security certificate used for secure e-commerce. The data collected from this study will be kept for 10 years on a password-protected computer in Dr. Romano's laboratory. Data from the study will be analyzed and presented at a group level so there will not be

presentation of individual results.

Right to Withdraw

The study can be completed in a quiet place at a time that is convenient for you.

We would really appreciate your honesty in responding to the questions.

Participation in this online research is completely voluntary. You have the right to refuse to answer any question and to withdraw from the study at any time, for any reason, without explanation or penalty. Any data completed prior to withdrawing from the questionnaire will be included in the study. If you do not wish to complete the entire study, you can exit at any time by clicking "exit study" option. Clicking yes indicates that you have read the above description and agree to participate in this study. You must answer "yes" if you wish to continue.

____ YES

____ NO

Please print/save a copy of the consent form for your record.

Thank you!

Appendix E:**Recruitment notice Study 1, Objective 2 (Content Validity Parents)**

The study “Evaluating the Validity and Reliability of the Positive Discipline in Everyday Parenting (PDEP) Questionnaire” is being conducted by Elena Gallitto for her Ph.D. thesis, under the supervision of Dr. Elisa Romano from the School of Psychology at the University of Ottawa. The goal of this study is to examine the usefulness of a questionnaire being used to assess the outcomes of the PDEP parenting program. For more information about the program please visit the main website: <http://www.positivedisciplineeveryday.com>.

Should you decide to participate, you will be invited to take part in a focus group to discuss the items that are included in the PDEP questionnaire. You will also be asked to complete some forms that will ask specific questions about the PDEP questionnaire. This study will last approximately 2 hours. If you decide to participate, you will have the chance to win a \$40 gift card.

If you decide to participate, you are free to withdraw from the study at any time without penalty. However, we will use the information that you have provided to us.

To be eligible to participate in this study, you must be caring for at least one child under age 18 years and you will need to be able to read and speak English. We are looking for a mix of parents, some who have completed the PDEP program and some who have not. Please note that we will select parents on a first come first served basis until we reach our goal of 10.

If you are interested in learning more about this study, please contact us.

Appendix F:

Study 1 – Objective 2 (Consent Form Parents)
Informed Consent 2 (Study I-Research Question 2)



uOttawa

Faculté des sciences sociales
 Faculty of Social Sciences

École de psychologie
 School of Psychology

Study title: Focus Group with Parents to Assess Content Validity of the Positive Discipline in Everyday Parenting Questionnaire.

The ethical aspects of the study have been approved by the Office of Research Ethics and Integrity of the University of Ottawa (REB # H-02-18-352, approved March 3, 2018).

Study Investigators

Elena Gallitto (Ph.D. Candidate, School of Psychology, University of Ottawa).
 Dr. Elisa Romano (Professor, School of Psychology, University of Ottawa).
 Should you have any questions or require additional information, please contact the Protocol Officer for Ethics in Research, Office of Research Ethics and Integrity, University of Ottawa.

Any questions about the current study may be addressed to Elena Gallitto.

Purpose

The purpose of this study is to make changes to the Positive Discipline in Everyday Parenting (PDEP) questionnaire and improve its usefulness. We will ask you to complete a form where you will rate the items of the PDEP questionnaire. You will also be asked to answer a brief demographic questionnaire (e.g., gender, education). After completing the questionnaires, we will discuss the content of the items in a focus group. Please note that the group discussion will not be tape-recorded. However, two trained research assistants from the University of Ottawa will be present during the interviews to take notes and record the answers.

Duration, Location and Remuneration

This study will take about 2 hours. The study will take place in Room (TBD) at the University of Ottawa (or name of the community centers). If you agree to participate in this study, you can be entered into a draw to win a \$40 gift card. Should you decide to participate in the study, you will be invited to enter a draw for a \$40 gift card. The draw is open to all research participants who enter their name in the draw, regardless of whether they decide to withdraw from further participating in the research project. Upon completion of the study, one name will be randomly selected among those who have entered and this participant whose name is drawn will be awarded the prize. The odds of winning a prize will depend on the number of eligible entries received. The prize must be accepted as awarded or forfeited and cannot be redeemed for cash.

The name that you provide when you enter the draw is collected for the purposes of contacting you if your name is selected in the draw. The name you have provided will be kept confidential and then destroyed once the prizes have been awarded. We reserve the right to cancel the draw or cancel the awarding of the prize if the integrity of the draw or the research or the confidentiality of participants is compromised. The draw is governed by the applicable laws of Canada.

Potential risk/discomfort

There are no known or anticipated psychological or physical risks in this study. However, if you experience anxiety or discomfort for any unanticipated reason, you may stop participating in the study. You have the option to contact the Counselling Services at the University of Ottawa and the Distress Centre of Ottawa and Region.

Anonymity/confidentiality

Be assured that no identifiable information will be gathered. The information you provide during this study will be kept strictly confidential. Your name or other identifying information will not be associated with your interview and will be removed during transcription (notetaking). As a focus group involves a group discussion it is impossible to provide complete confidentiality or anonymity, but guidelines will be presented to the group to ensure that the focus group space is a safe space for open discussion. The data collected from this study (e.g., notes, interview files, demographic information) will be kept for 10 years on a password-protected computer in Dr. Romano's laboratory. Data from the study will be analyzed and presented at a group level so there will not be presentation of individual results.

Right to Withdraw

Participation in this study is completely voluntary. You have the right to refuse to answer any question and to withdraw from the study at any time, for any reason, without explanation or penalty. However, because the interdependent nature of focus group data, individual data cannot be removed once the focus group has been completed. Selecting yes indicates that you have read the above description and agree to participate in this study.

☐ YES

☐ NO

Please note that there are two copies of consent form, one of which is yours to keep.

Thank you!

Appendix G:**Interview Protocol for the Focus Group**

First, we would like to thank you for agreeing to meet with us. We are here today to ask you some questions about a questionnaire that is being currently used to understand the effects of the Positive Discipline in Everyday Parenting (PDEP) program. We'd like to start by briefly reviewing a few important guidelines to keep in mind for today's discussion.

Please feel comfortable to say what you think and feel. We are here to gather information and opinions, positive or negative, and your responses today will remain anonymous. Your answers won't affect your services with the (NAME of community center) now, or in the future. We will not reveal the identity of any of you who have spoken today, and we ask that you refrain from talking about other people's opinions outside of the focus group.

We would like to hear from everyone as much as possible, so we will do our best to make sure that everyone's thoughts are considered, and everyone is provided a similar amount of time to contribute. At times, we may also summarize the main ideas addressed by one of you to ensure that it reflects your opinions/comments.

Please note we will not be audio-recording our discussion. However, two trained research assistants will record your answers on paper. Note that this will not be shared with anyone outside of our research team. Any identifying information recorded during the discussion will also be removed so your responses remain anonymous.

First, you will be asked to rate each item using a form. In this form, you will be asked to rate each item to see if there is anything that can be improved. After you have finished completing the form, we will gather together to have a discussion about the PDEP questionnaire.

Is there any question that was particularly confusing for you? If so, what would make this question less confusing?

How would you rate the language used to formulate the questions? Was it easy or difficult to understand?

How did you arrive at that answer?

What does (a particular word/concept) mean to you?

In your opinion, are there any items that do not related to any of the PDEP program's goals? If so, what would make this question more in line with the PDEP program's goals? *

This question was asked only to parents who participated in the PDEP program or had prior knowledge of the PDEP program.

Appendix H:**Sociodemographic Questionnaire (Experts)**

The following questions include background information that will only be used descriptively to understand the characteristics of participants. All information will be kept strictly confidential and will be presented for the sample as a whole.

1. What is your gender?

- ☐ Male
- ☐ Female
- ☐ T
- ☐ You don't have an option that applies to me. I identify as (please specify) _____

2. How old are you? _____ years

3. What is the highest level of education that you completed?

- ☐ Less than high school
- ☐ High school diploma or equivalent (GED)
- ☐ Community college
- ☐ Bachelor or undergraduate degree
- ☐ Master's degree (e.g., M.A., M.Sc., M.Ed.)
- ☐ Doctorate degree or equivalent
- ☐ Other (Please specify) _____

4. Please indicate which options best represent your ethnic background: (Select all that apply)

- ☐ White /European
- ☐ Aboriginal (North American Indian, Métis or Inuit)
- ☐ Arab (e.g., Saudi, Egyptian, Iraqi, Lebanese, Palestinian, Syrian etc.)
- ☐ Black (e.g., African, African American, African Canadian, Caribbean)
- ☐ East Asian (e.g., Chinese, Japanese, Korean, Polynesian)
- ☐ Latin American
- ☐ South Asian (e.g., Indian, Pakistani, Sri Lankan, Bangladeshi)
- ☐ Southeast Asian (e.g., Burmese, Cambodian, Filipino, Indonesian, Laotian, Malaysian, Thai, Vietnamese)
- ☐ West Asian (Afghan, Armenian, Iranian, Israeli, Turks etc.)
- ☐ Other (Please specify) _____
- ☐ Prefer not to answer

5. List your top 3 areas of expertise _____

6. How would you rate your level of knowledge about child development?

- ☐ Very low
- ☐ Low
- ☐ Medium
- ☐ High
- ☐ Very high

7. How would you rate your level of knowledge about measurement development?

- ☐ Very low
- ☐ Low
- ☐ Medium
- ☐ High
- ☐ Very high

8. How would you rate your level of knowledge about parenting research?

- ☐ Not at all
- ☐ A little
- ☐ Moderately
- ☐ Very
- ☐ Extremely

Appendix I:**Sociodemographic Questionnaire (Parents)**

The following questions include contact and background information. All information provided will be kept strictly confidential. The background information will be used only for data analyses (e.g., correlation) and group comparisons (e.g., gender differences). Your name or other identifying information will be coded in the final dataset and will not be associated with the data you provide on the questionnaires.

1. What is your gender?

- ☐ Male
- ☐ Female
- ☐ T
- ☐ You don't have an option that applies to me. I identify as (please specify) _____

2. How old are you? ____ years

4. What is your current marital status?

- ☐ Married
- ☐ Living with a partner
- ☐ Single (never married)
- ☐ Widowed
- ☐ Separated
- ☐ Divorced

5. What language do you speak most often at home? (choose one)

- ☐ English
- ☐ French
- ☐ Other: (please specify) _____

6. How would you rate your English reading comprehension?

- ☐ Beginner
- ☐ Intermediate
- ☐ Advanced
- ☐ Superior
- ☐ Native

7. How would you rate your writing skills in English?

- ☐ Very low
- ☐ Low
- ☐ Medium
- ☐ High
- ☐ Very high

8. Please indicate which options best represent your ethnic background: (Select all that apply)

- ☐ White /European

POSITIVE DISCIPLINE

- ☐ Aboriginal (North American Indian, Métis or Inuit)
- ☐ Arab (e.g., Saudi, Egyptian, Iraqi, Lebanese, Palestinian, Syrian etc.)
- ☐ Black (e.g., African, African American, African Canadian, Caribbean)
- ☐ East Asian (e.g., Chinese, Japanese, Korean, Polynesian)
- ☐ Latin American
- ☐ South Asian (e.g., Indian, Pakistani, Sri Lankan, Bangladeshi)
- ☐ Southeast Asian (e.g., Burmese, Cambodian, Filipino, Indonesian, Laotian, Malaysian, Thai, Vietnamese)
- ☐ West Asian (e.g., Afghan, Armenian, Iranian, Israeli, Turks etc.)
- ☐ Other:
- ☐ Prefer not to answer

9. What is the highest level of education that you completed?

- ☐ Less than high school
- ☐ High school diploma or equivalent (GED)
- ☐ Community college
- ☐ Bachelor or undergraduate degree
- ☐ Master's degree (e.g., M.A., M.Sc., M.Ed.)
- ☐ Doctorate degree or equivalent
- ☐ Other (Please specify) _____

9. If you are currently living with a partner, what is the highest level of education that he/she completed?

- ☐ Less than high school
- ☐ High school diploma or equivalent (GED)
- ☐ Community college
- ☐ Bachelor or undergraduate degree
- ☐ Master's degree (e.g., M.A., M.Sc., M.Ed.)
- ☐ Doctorate degree or equivalent
- ☐ Other (Please specify) _____

10. What is your employment situation?

- ☐ Not employed
- ☐ Employed outside the home on a part-time basis
- ☐ Employed outside the home on a full-time basis
- ☐ Other (please specify) _____

11. If you are currently living with a partner, what is his/her employment situation?

- ☐ Not employed
- ☐ Employed outside the home on a part-time basis
- ☐ Employed outside the home on a full-time basis
- ☐ Other (please specify) _____

12. In which of the following groups does your household income fall in Canadian funds (prior to income tax)?

- ☐ Less than \$10,000
- ☐ \$10,000 - \$19,999
- ☐ \$20,000 - \$29,999
- ☐ \$30,000 - \$39,999

POSITIVE DISCIPLINE

- ☐ \$40,000 - \$49,999
- ☐ \$50,000 - \$59,999
- ☐ \$60,000 - \$69,999
- ☐ \$70,000 - \$79,999
- ☐ \$80,000 - \$89,999
- ☐ \$90,000 - \$99,999
- ☐ \$100,000 - \$109,999
- ☐ \$110,000 - \$119,999
- ☐ \$120,000 - \$129,999
- ☐ \$130,000 - \$139,999
- ☐ \$140,000 - \$149,999
- ☐ Over \$150,000

13. How did you hear about the study?

- ☐ Community organization
- ☐ My friends
- ☐ Email
- ☐ Social Media (e.g., Facebook, Twitter, Readit, etc.)
- ☐ Internet (other websites)
- ☐ Other

Please specify: _____

Did you participate in the Positive Discipline in Everyday Parenting (PDEP) program?

- ☐ Yes
- ☐ No

Appendix J:**Content Validity Forms (Experts and Parents)***Content Validity Form for Approval of Physical Punishment***Component Description/Definition:** The degree to which parents believe that physical punishment is acceptable, effective and/or necessary.

Items measuring “Approval of Physical Punishment”	Q1 ^a Representativeness	Q2 ^b Clarity	Q3c Suggestion or alternative wording
Sometimes a spank or a slap is the best way to get a child to listen.	1 2 3	1 2 3	
Spanking or slapping is fine as long as the parent is not angry.	1 2 3	1 2 3	
Parents should have the right to decide whether to spank or slap their children.	1 2 3	1 2 3	
If parents don’t spank or slap, their children will be spoiled.	1 2 3	1 2 3	
It’s ok to spank a 5-year-old’s bottom if she does something dangerous.	2 3	1 2 3	
It’s ok to slap a teenager’s face if she uses bad words in front of her parents.	1 2 3	1 2 3	
Young children who aren’t spanked will get into trouble when they are teenagers.	1 2 3	1 2 3	

Spanking teaches children right from wrong.	1 2 3	1 2 3	
If a 4-year-old won't leave a store, his parent should pull his arm and drag him out.	1 2 3	1 2 3	
It's ok to shake a 3-year-old to make her stop screaming.	1 2 3	1 2 3	
If a 2-year-old boy bites another child, it's ok to bite him to teach him how it feels.	1 2 3	1 2 3	
Children who are spanked learn to respect their parents.	1 2 3	1 2 3	
It's ok to slap a 2-year-old's hand for touching dangerous objects.	1 2 3	1 2 3	

Note:

a Q 1(Representativeness). Please rate how well each item represents the construct as defined in the appropriate space above.

[Rating Scale: **1 = Slightly representative; 2 = Moderately representative; 3 = Highly representative**]

b Q2 (Clarity). Please rate each item's clarity. Clarity refers to the degree to which the item is readily understood.

[Rating Scale: **1= Low; 2= Moderate; 3=High**]

c Q3 (Suggestion). If you have rated the representativeness or clarity of item low, please change the wording of the item to make it clearer and more representative of the construct.

Content Validity Form for Approval of Non-Physical Punishment

Component Description/Definition: The degree to which parents believe that non-physical punishment is acceptable, effective and/or necessary.

Items measuring “Approval of Non-Physical Punishment”	Q1 ^a Representativeness	Q2 ^b Clarity	Q3c Suggestion or alternative wording
If a 14-year-old is failing in school, his parent should make him do extra chores.	1 2 3	1 2 3	
If a 12-year-old gets into trouble at school, her parent should punish her before she has a chance to give excuses.	1 2 3	1 2 3	
If a 16-year-old wears a hairstyle that his parent disapproves of, he should have to stay inside until he changes it.	1 2 3	1 2 3	
If children break the rules, their parents should take away privileges.	1 2 3	1 2 3	
If a 3-year-old won’t eat, his parent should make him stay at the table until he finishes his food.	1 2 3	1 2 3	
If two sisters won’t stop fighting, the parent should threaten to take away their phones	1 2 3	1 2 3	
A good way to teach a 13-year-old to dress properly is to embarrass her in front of her friends.	1 2 3	1 2 3	
If a 6-year-old won’t get dressed, his parent should threaten to take him out in his underwear.	1 2 3	1 2 3	
A 10-year-old who hurts another child should be forced to apologize.	1 2 3	1 2 3	

If a 3-year-old won't share his toys, he should be made to sit on a chair for 3 minutes.	1 2 3	1 2 3	
--	-------	-------	--

Note:

a Q 1(Representativeness). Please rate how well each item represents the construct as defined in the appropriate space above.

[Rating Scale: **1 = Slightly representative; 2 = Moderately representative; 3 = Highly representative**]

b Q2 (Clarity). Please rate each item's clarity. Clarity refers to the degree to which the item is readily understood.

[Rating Scale: **1= Low; 2= Moderate; 3=High**]

c Q3 (Suggestion). If you have rated the representativeness or clarity of item low, please change the wording of the item to make it clearer and more representative of the construct.

Content Validity Form for Subjective Norms

Component Description/Definition: The degree to which parents perceive age-typical behaviour as 'misbehaviour'.

Items measuring "Subjective norms"	Q1 ^a Representativeness	Q2 ^b Clarity	Q3c Suggestion or alternative wording
Two-year-olds who say "No!" to their parents are being defiant.	1 2 3	1 2 3	
Usually, 3-year-olds have tantrums because they are spoiled.	1 2 3	1 2 3	
Four-year-olds who interrupt adults are rude.	1 2 3	1 2 3	
A 13-year-old who does not want to be seen with his mother should be ashamed of himself.	1 2 3	1 2 3	
If an 8-year-old uses bad words in front of his parents, this is a sign of disrespect.	1 2 3	1 2 3	
A 16-year-old who questions her parents' religious beliefs is being disrespectful.	1 2 3	1 2 3	

A 5-year-old child who won't go to sleep is being stubborn.	1 2 3	1 2 3	
A 6-year-old child who won't show affection to his grandmother is rude.	1 2 3	1 2 3	
Boys should learn not to cry by the time they are 12 years old.	1 2 3	1 2 3	
Teenagers who live at home should always follow their parents' rules, even if they don't agree.	1 2 3	1 2 3	

Note:

a Q 1(Representativeness). Please rate how well each item represents the construct as defined in the appropriate space above.

[Rating Scale: **1 = Slightly representative; 2 = Moderately representative; 3 = Highly representative**]

b Q2 (Clarity). Please rate each item's clarity. Clarity refers to the degree to which the item is readily understood.

[Rating Scale: **1= Low; 2= Moderate; 3=High**]

c Q3 (Suggestion). If you have rated the representativeness or clarity of item low, please change the wording of the item to make it clearer and more representative of the construct.

Content validity form for the Self-Efficacy scale

Component Description/Definition: The degree to which parents believe that they are able to respond to conflict constructively.

Items measuring "self-efficacy"	Q1 ^a Representativeness	Q2 ^b Clarity	Q3c Suggestion or alternative wording
I can solve challenging situations without punishing this child.	1 2 3	1 2 3	
I believe I have the skills to be a good parent to this child.	1 2 3	1 2 3	

I feel like I just don't know what to do with this child.	1 2 3	1 2 3	
I can control my anger in challenging situations with this child.	2 3	2 3	
I feel badly about how I handle challenging situations with this child.	1 2 3	2 3	
I feel hopeful that I can improve my relationship with this child.	1 2 3	2 3	
In challenging situations, I end up shouting at this child.	1 2 3	1 2 3	
I believe that I will be able to handle future challenges as this child gets older.	2 3	2 3	
I feel hopeless about my relationship with this child.	1 2 3	1 2 3	

Note:

a Q 1(Representativeness). Please rate how well each item represents the construct as defined in the appropriate space above.

[Rating Scale: **1 = Slightly representative; 2 = Moderately representative; 3 = Highly representative**]

b Q2 (Clarity). Please rate each item's clarity. Clarity refers to the degree to which the item is readily understood.

[Rating Scale: **1= Low; 2= Moderate; 3=High**]

Appendix K:**Recruitment Notice Study 1 -Construct Validity Assessment (Convergent Validity)**

I am a PhD student at the University of Ottawa. I am inviting you to take part in a study. To participate in the study, you need to have at least one child who is between the ages of birth and 18 years. You will need access to a computer, and you will need to be able to read and write English.

My goal is to find out the best way to learn what people think about parenting. I am studying several questionnaires to find out which ones provide the most useful information.

If you decide to participate, you will fill out several questionnaires on-line. This will take 30 to 40 minutes of your time. You can fill out the questionnaires from any computer or smart phone. When you complete the questionnaires, you will be entered into a draw to win one of four \$40 Walmart gift cards.

When you have finished the questionnaires, you will be asked if you would like to participate in a second on-line study. If you say “Yes”, you will be asked for your email address.

We will contact you by email about 4 weeks later. At that time, you will be asked to fill out a questionnaire that will take 10 to 15 of your time. When you fill out that questionnaire, you will be entered into a new draw to win one of two \$40 Walmart gift cards.

I will be very careful to protect your privacy. I will not know your name and I will not be able to connect your questionnaires to you in any way. We have systems to ensure that everything you say on the questionnaires will be kept private.

You are free to not participate. You also are free to leave the study at any time, even if you have started filling out the questionnaires.

If you wish to participate in this study, please click on this link that will take you to the consent form and the questionnaires.

Appendix L:**Consent Form- Validity and Reliability Assessment****uOttawa**Faculté des sciences sociales
Faculty of Social SciencesÉcole de psychologie
School of Psychology

The ethical aspects of the study have been approved by the Office of Research Ethics and Integrity of the University of Ottawa (REB # H-02-18-352, approved March 3, 2018).

Study Investigators

Elena Gallitto (Ph.D. Candidate, School of Psychology, University of Ottawa).

Dr. Elisa Romano (Professor, School of Psychology, University of Ottawa).

Should you have any questions or require additional information, please contact the Protocol Officer for Ethics in Research, Office of Research Ethics and Integrity, University of Ottawa.

Any questions about the current study may be addressed to Elena Gallitto.

Purpose

The goal of this study is to examine several questionnaires used to understand parenting. You will be asked to answer some questions regarding your parenting behaviours as well as a brief demographic questionnaire (e.g., gender, education). Should you decide to participate, you will be invited to complete a questionnaire which will take approximately 40-50 minutes. Given the length of the questions, it is possible that some participants may require additional time to complete the questionnaire. The questionnaire is available on-line and can be completed from any computer at a place of your choice. Once you begin the questionnaire, it should be noted that it will not be possible to close the study's webpage and return to it at a later point. However, you will be allocated an 8-hour time-frame to complete the questionnaire. At the end of the questionnaire, you will have the option to leave your email address if you are interested in being contacted about a second research study.

The goal of the second study is to collect reliability data on a measure of parenting behaviours. If you leave your email address, a research assistant will contact you within four weeks to complete a 10-15-minute questionnaire about your parenting behaviours. You will also be asked to enter a unique 8-character code (i.e., your initials, day, month and last two years of birth) which will be used to link your responses at both time points. If you do not want to participate in the second study, answer "No, I am not interested" and proceed to the draw page.

Potential risk/discomfort

Some of the questions are sensitive in nature in that they will ask you about your current psychological functioning. As such, it is possible that you may experience some emotional distress. The on-line study provides a link to a list of resources, such as telephone distress lines and counselling services

available at the University of Ottawa. Please do not hesitate to contact these resources should you feel distressed. Any data that you have completed up until that point will be retained, and you will have the option to return to the questionnaire.

Anonymity/confidentiality

Be assured that no identifiable information will be gathered. The information you provide during this study will be kept strictly confidential and anonymous. Access to the data you provide will be limited to the researchers listed above. Your name will not be revealed in any report regarding this research. To ensure the security and confidentiality of online communication in the study, the online study is equipped with encrypted communication features and a security certificate used for secure e-commerce. The data collected from this study will be kept for 10 years on a password-protected computer in Dr. Romano's laboratory. Data from the study will be analyzed and presented at a group level so that there will not be presentation of individual results.

Copyright ©

This survey includes copyrighted materials. It may not to be translated, stored, recorded, used, transmitted, or reproduced in whole or in part without prior permission in writing from the principal researchers.

Compensation

Should you decide to participate in the study, you will be invited to enter a draw for one of four \$40 gift cards. The draw is open to all research participants who enter their name in the draw, regardless of whether they decide to withdraw from further participating in the research project.

Right to Withdraw

The study can be completed in a quiet place at a time that is convenient for you. We would really appreciate your honesty in responding to the questions. Participation in this online research is completely voluntary. You have the right to refuse to answer any question and to withdraw from the study at any time, for any reason, without explanation or penalty. Any data completed prior to withdrawing from the questionnaire will be included in the study. Once you complete the questionnaire, or, if you do not complete the entire study, when you select the "exit study" option, you will be invited to be entered for a draw for one of four \$50 Visa gift cards. If you choose to enter your name for the draw, you will be forwarded to a different window and will be required to enter your email address. Please note that your e-mail address cannot be linked to the answers you have provided during the study. Clicking yes indicates that you have read the above description and agree to participate in this study. You must answer "yes" if you wish to continue.

☐ YES

☐ NO

Appendix M:
Parent Cognition Scale

Please rate how much you would agree, in general, that the following reasons for misbehaviour are true for the target child and his/her behavior for the past two months:

1. I was not as firm as I usually am.
2. My child won't listen.
3. I'm not structured enough with my child.
4. My child cannot understand the rules.
5. My child thinks that he/she is the boss.
6. I don't know how to handle my child.
7. I don't give my child enough attention.
8. My child is headstrong.
9. It's hard for me to set limits.
10. My child is in a stage.
11. My child wants what he/she wants when he/she wants it.
12. I was tired at the time.
13. I handle my child in a non-confident way.
14. My child purposely tries to get me angry.
15. My child feels like there is no time for him/her.
16. I'm not patient.
17. My child tries to get my goat or push my buttons.
18. My child wants things his/her way.
19. It's difficult for my child to do what I want.
20. I can't control my child.
21. I couldn't respond quickly enough at the time.
22. I'm not able to be clear.
23. My child is very demanding.
24. I handled things in an unusual way.
25. My child likes to see how far he/she can push me.
26. I was busy with something at the time.
27. I don't do the right thing.
28. My child tires easily.
29. I have a hard time really listening to my child.
30. My child refuses to do what I think he/she should do.

Response categories:

- 1 = Always True
- 2 = Frequently True
- 3 = Sometimes True
- 4 = Occasionally True
- 5 = Rarely True
- 6 = Never True

Appendix N:**Discipline subscale of the Self-Efficacy for Parenting Tasks scale**

Please read each of the following statements, then check the box that best describes how you feel about it.

1. I am a good enough disciplinarian for my child.
2. I have trouble deciding on appropriate rules for my child.
3. My child never seems to listen to me when I want him/her to do something.
4. I have more trouble with discipline than any other aspect of parenting.
5. My disciplinary skills are at least as good as an average parent.
6. I spend too much time with ineffective attempts to discipline my child.
7. I really don't have much trouble disciplining my child.
8. I thought I was a good parent until I started struggling so much with discipline.

Response categories:

- 1 = strongly disagree
- 2 = I mostly disagree
- 3 = I somewhat disagree
- 4 = I somewhat agree
- 5 = I mostly agree
- 6 = I strongly agree

Appendix O:

Ethics Approval Study 3

04/07/2018

Université d'Ottawa

Bureau d'éthique et d'intégrité de la recherche

University of Ottawa

Office of Research Ethics and Integrity

CERTIFICAT D'APPROBATION ÉTHIQUE | CERTIFICATE OF ETHICS APPROVAL

Numéro du dossier / Ethics File Number

H-05-18-630

Titre du projet / Project Title

Evaluation of the Positive
Discipline in Everyday Parenting
(PDEP) Program

Type de projet / Project Type

Recherche de professeur /
Professor's research project

Statut du projet / Project Status

Approuvé / Approved

Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)

04/07/2018

Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)

03/07/2019

Équipe de recherche / Research Team

**Chercheur /
Researcher**
Affiliation**Role**

Elisa ROMANO

École de psychologie / School of
Psychology

Chercheur Principal / Principal Investigator

Elena GALLITTO

École de psychologie / School of
PsychologyCoordonnateur de recherche / Research
Coordinator

Joan DURRANT

University of Manitoba

Collaborateur / Collaborator

George HOLDEN

Southern Methodist University

Collaborateur / Collaborator

Christine ATEAH

University of Manitoba

Co-chercheur / Co-investigator

Carl NEWTON

École de psychologie / School of
Psychology

Assistant de recherche / Research Assistant

Conditions spéciales ou commentaires / Special conditions or comments

Appendix P:
Recruitment Notice Study 3

Do you have a child between the ages of 0 and 6?

Would you like to help researchers at the University of Ottawa to better understand and support families?

The PDEP program will be offered at different locations including the following: (1) Centre des services communautaires Vanier (Vanier); (2) Andrew Fleck Child Care Services (Ottawa-South); (3) Western Ottawa Community Resource Centre (Western Ottawa/Carleton); (4) Pinecrest-Queensway/Ottawa West-Nepean (Pinecrest-Queensway Health and Community Services) and (5) Orléans-Cumberland Community Resource Centre (Orleans).

This parenting program is:

A disciplinary approach that aims at promoting positive and non-violent discipline of children

When?

- This 8-week parenting program will be offered at a group of 20 parents between September and December 2018 and between February and April 2019

Why to participate at this parenting program?

- To obtain information on child development, understand the reasons and feelings of children and develop strategies to resolve conflicts that may occur from birth to adolescence.
- To help the researchers at the University of Ottawa to evaluate the efficacy of the program in a way that will inform public programs on how to support parents in their disciplinary practices.

The participation at this study is free and voluntary.

If you accept to participate in the study, you will be asked to complete anonymous online questionnaires.

Are you interested? Do you have any questions?

Please contact:

Dr. Elisa Romano (613-562-5800, ext. 4403) of the Children Well-Being Laboratory at the University of Ottawa.

Appendix Q:

Consent Form- Study 3



uOttawa

Faculté des sciences sociales
Faculty of Social SciencesÉcole de psychologie
School of Psychology**Title:** Evaluation of the Positive Discipline in Everyday Parenting (PDEP)**Principal Investigator:** Elisa Romano, Ph.D.School of Psychology,
Faculty of Social Sciences, University of Ottawa**Research Team:** Christine Ateah (Co-investigator; University of Manitoba, Winnipeg, Manitoba)

Leslie Barker (Collaborator; PDEP Master Trainer, Calgary, Alberta)

George Holden (Collaborator; Southern Methodist University, Dallas, Texas)

Joan Durrant (Collaborator; University of Manitoba, Winnipeg, Manitoba)

Elena Gallitto, David Romano, and Stephanie Quinn (Research Assistants; University of Ottawa, Ottawa, Ontario)

Invitation to Participate: I am invited to participate in this research project by Dr. Romano and colleagues. This project is being done with the help of the Early ON Child and Family Centres in Ottawa, in surrounding areas, and in southern Ontario. This project is funded by the Social Sciences and Humanities Research Council of Canada.**Purpose of the Study:** The purpose of the project is to understand how the Positive Discipline in Everyday Parenting (PDEP) program might be helpful for families. The research project is separate from the PDEP program. I can still participate in the program even if I do not want to participate in the research.**Inclusion/Exclusion criteria:** To participate in the research study, I must be completing the PDEP program and have a child between 2 and 6 years old. I need to feel comfortable with the English language and I should not be completing another parenting program while I am in the PDEP group (or while I am waiting to start the PDEP group).**Participation:** This research examines how the PDEP program might be helpful for families, so it has to follow specific steps where two PDEP groups are given every year. The first group will happen fairly soon and then the second group will happen a few months after the first group ends. If I participate in the research, I understand that I could be randomly placed in the first group or that I might have to wait a few months for the second group.

To make sure that the PDEP program is being properly followed, a research assistant may be present at the meetings to complete a checklist. This checklist is to make sure that the facilitator is delivering the PDEP program the way that it was meant to be delivered – this is not about you, but about

the facilitator's training in PDEP. It is also important for the research assistant to know which PDEP sessions I have attended and how many sessions in all. I understand that a research assistant will collect this information toward the end of the PDEP program.

If I decide to participate, I will be asked to complete the following:

Questionnaire package. The questionnaires will be completed four times – the first time is before the start of the PDEP program; the second time is soon after PDEP has ended; the third time is 1 month after PDEP has ended; and the fourth time is 3 months after PDEP has ended. Depending on which PDEP group I am in and when it starts, I may not be asked to complete the questionnaires the fourth time.

These questionnaires will take about 60 minutes to complete, and they ask about such things as my background, my thoughts about discipline, and the ways I handle discipline with my child. There are different ways to fill out the questionnaires depending on what works best for me. I can either complete them online on a computer, or I can complete the questionnaires by hand by myself or with the help of a research assistant.

If I complete the questionnaires online, it would be important to minimize the risk of security breaches and to help ensure my confidentiality. As such, it is recommended that I use standard safety measures – for example, signing out of my account, closing my browser, and locking my screen or device when I am no longer using them when I have finished completing the questionnaires.

Interview. Soon after the PDEP program ends, I may be asked to complete an interview with a research assistant who will either come meet me at the agency or by phone. The interview will take about 30-60 minutes and will ask about what I found helpful and unhelpful and what I might suggest to improve the program. The interview will be audiotaped and then typed up.

Risks: My participation requires quite a bit of time so I may sometimes feel tired, but I understand that I can take breaks when I need them. Because of the time, my family's routines may be affected but I understand that the researchers will try to schedule meetings at a time that works best for me. I will also be talking about personal things, and this may be difficult emotionally. However, I can choose what to talk about, and I do not have to talk about anything that might make me feel uncomfortable.

Benefits: By participating in this study, I will be providing information to the researchers to help them better understand the ways that parents and families are helped by the PDEP program and what might be important to change to make the program better.

Confidentiality and anonymity: For the research study, I understand that no information will be gathered that could identify me. My answers are strictly confidential, and no one will be told of my answers. My answers on the different questionnaires will be matched through a code, but this code will not allow researchers to identify me personally. Only the research team will have access to my information, which will be kept for 10 years after the study ends on a password-protected computer in Dr. Romano's laboratory at the University of Ottawa and in a locked filing cabinet in the laboratory.

For the information that I provide, such as questionnaires and interviews, it will be coded so that my identity will be protected. For example, my name will not appear on any of these documents

but instead, a code will be used. Also, I understand that some of the information may be used for student projects and will be presented at conferences and in journal articles. For example, Ms. Quinn will be using some of the information for her fourth year undergraduate Honour's thesis project. When this happens, there will never be mention of my name, the names of my family members, or any other information that could point to who my family is.

I understand that the information I share will remain confidential. The information from the questionnaires and interview, for example, will not be shared with the agency where I completed the PDEP program. However, I understand that there are certain situations where it may not be possible to protect confidentiality. These include situations where a child may be experiencing maltreatment, where I may be at risk of hurting myself or another person, where I may be experiencing sexual harm by a health professional, or where there is legal involvement with my family. In such cases, the appropriate authorities will need to be contacted to make sure that individuals stay safe.

Conservation of data: The information collected will be safely stored in Dr. Romano's laboratory at the University of Ottawa. The laboratory is a locked room which has a locked cabinet for such things as my questionnaires and transcribed interview. There also is a password-protected computer for the information from my questionnaires and interview. For the online questionnaires, data will be stored on Qualtrics. The Qualtrics website is encrypted, and your data will be stored on a server protected by a firewall. Once the study is complete, your information will be stored on Qualtrics for 30 days and then it will be stored electronically on a password-protected computer in Dr. Romano's laboratory. Results will be examined in group form meaning that no individual results will be presented. Information will be kept for 10 years after the study ends and then it will be deleted. Only Dr. Romano and her research team will have access to the data.

Compensation: Every time I fill out the questionnaire package, I will receive a \$50 gift card. If I complete the interview, I will receive a \$20 gift card. If I miss one of these sessions and it cannot be rescheduled, I will not receive any compensation. However, if I begin one of the sessions and do not totally complete it, I will still receive the compensation.

Voluntary Participation: I understand that I am not obliged to participate in this study. Whether I choose to participate or not will have no impact on me being able to complete the PDEP program. If I choose to participate, I can withdraw from the research study at any time and/or refuse to answer any questions. There are no consequences if I decide to stop participating. If I want to continue participating in the PDEP program but not in filling out the questionnaires, this is also possible. If I choose to stop participating in the study, all information gathered up until that point can be kept for the researchers to use. However, I have the right to ask the researchers not to use any of the information that I have provided.

Acceptance (paper-pencil version): I, _____ (name of the participant), agree to take part in this research by Dr. Elisa Romano and the research team.

I understand that by accepting to participate, I still have the right to withdraw from the study anytime. If I have any questions about the study, I may contact Dr. Romano at the telephone number or email at the top of this consent form.

If I have any questions regarding the ethical conduct of this project, I may contact the Protocol Officer for Ethics in Research
University of Ottawa

There are two copies of the consent form, one of which is mine to keep.

Participant's signature:

Date:

Researcher's signature:

Date:

Acceptance (online): If you understand all the statements above and consent to participate in the online study to evaluate the parenting group, click on the Yes button below. If you understand all the statements above and do not want to participate in the online study to evaluate the parenting group, click on the No button below.

Do you wish to participate in the research study?

Yes *Please note that by clicking yes, you will be directed to the first questionnaire to be completed on-line.

No

_____ Click here to print a copy of this consent form.

Appendix R:

Facilitator Fidelity Forms

Session 1: Introduction of the Program

RA	Component	Topic	Done(✓)	Comments
	Welcome	Establishes welcoming atmosphere as people arrive; facilitates introductions of parents and self.		
	Opening	Addresses housekeeping issues (parking, washrooms, refreshments, phones, etc). Introduces self, establishes credibility as parenting educator. Provides a short activity for parents to introduce themselves.		
		Builds rapport with and among parents.		
	Pre-Program Questionnaire	Administers Pre-Program Questionnaire by first explaining: Rationale		
		Anonymity		
		Confidentiality		
		Freedom to participate or not		
		Provides quiet atmosphere		
		Confidential collection of forms		
		Secures completed forms		
	Group Agreements (GA)	Asks parents for ideas on how this group will work together, facilitates discussion to draw out key elements (e.g., start/end on time, right to participate/pass, respect for self/others, confidentiality). Posts the GA in a visible location.		
	Agendas	Provides an overview of 8-week Agenda and Session 1 Agenda.		
	Program Introduction	Introduces the program's approach - explains non-prescriptive nature of program; reassure that with time and practice, it will become automatic way of thinking/approaching problems.		
	Gathers parent's questions	Acknowledges that parents have many questions about parenting and that most parents have a specific concern.		
		Invites parents to record concern on slip of paper, and put in "Problem Box". Reassures that many parents will solve their concerns during the program once they understand the		

RA	Component	Topic	Done(✓)	Comments
		process, but if not, they will be brought back for resolution toward the end of program.		
		Reinforces PDEP not a prescriptive program; it helps parents find their own solutions so they can continue to solve their parenting challenges in the future.		
		Conveys confidence in parents' competence.		
	Experience of parenting	Carries out <i>'Experience of Being a Parent'</i> activity as described in the Manual (p. 24).		
		Normalizes parenting emotions and mixed feelings about parenting; renew focus on positive experiences.		
		Sets a warm and welcoming atmosphere; non-judgmental responses; asks questions for clarification; encourages responses from all participants.		
	How we learn to parent	Leads group discussion on how we learn to parent as described in Manual (p. 26).		
		Acknowledges that parents do the best they can with the knowledge/skills/supports they have. Explains that how we were parented influences, but does not determine how we parent, and that we can learn new ways.		
		Reassures parents that with time, knowledge and patience, their abilities will grow, and they'll be more and more confident in finding solutions.		
	Introducing the PD Approach	Explains PD as an approach; a way of thinking.		
		Explains the "3 -circle" diagram (p. 28) so parents realize program is based in research (healthy child development and effective parenting) and child rights.		
	Introduce the UN CRC (p.30)	Asks how many people are familiar with the CRC.		
		Explains CRC's purpose is to guarantee children's basic human rights.		

RA	Component	Topic	Done(✓)	Comments
		Shares the year when this training's country ratified the CRC.		
		Explains that CRC is also about supporting and meeting parents' needs (to be able to meet their children's basic needs, help them develop to full potential, and support in learning how to parent effectively).		
		Directs parents to the CRC handout in their workbook and answers questions.		
	Physical and Emotional Punishment	States that PD does not include either physical or emotional punishment.		
		Acknowledges that PD is based on fundamental belief in non-violent conflict resolution.		
	What PD is and is not	States what PD IS NOT, and what PD IS.		
		Distinguishes PD from permissive parenting.		
		Distinguishes PD from non-physical punishment; is clear that 'time-out' is NOT a PD strategy.		
		Establishes definition of 'discipline' is 'to teach'. PDEP is about effective teaching.		
	The PD Model	Explains the house metaphor (p. 44) to illustrate the components of the PD Model (Long term Goals, Warmth and Structure, How Children Think and Feel, Problem Solving).		
	The Parent Book "Treasure Hunt"	In small groups, has parents explore the Parent Book with guided exploration questions.		
		Emphasizes how the Parent Book reflects the PDEP model, and that each level is equally important. Note that effective problem solving is based on a thorough understanding of all components.		
	Closing	Provides formal closing with summary of key points, answers questions; Offers 'between session' suggestions for things to try at home; may provide a mini evaluation.		

Summary Note: Day 1:

Session 2: Identifying Long-Term Parenting Goals

RA	Component	Topic	Done(✓)	Comments
	Welcome and Opening	Is prepared when parents arrive; establishes welcoming atmosphere; asks about parents' thoughts about or experience with PD over the intervening week; provides a brief opener activity.		
		Introduces session topic and process.		
	Understanding short-term goals (STGs)	Conducts imagination scenario (p. 50) to help parents identify their STGs (asks parents to close eyes if comfortable, clearly reads scenario);		
		Asks parents to identify their immediate parenting goals, then share with the group.		
		Records parents' responses on flipchart.		
	The Stress Response	Reads part 2 of Scenario (p. 52).		
		Helps parents identify the impact of stress on their body/mind – particularly the body, voice and feelings/emotions.		
		Asks parents to write on post-it notes, what a parent might say or do in this situation. Reads out responses while posting them on flipchart.		
		Reassures parents these feelings/responses are typical for most people; identifies connection to physical/emotional punishment.		
	Understanding the Stress Response	Uses ppt slides to explain basic brain structures, and the stress response in the brain. Shows Dan Siegel Video (The Hand Model of the Brain), and leads parents in practice.		
		Stresses the relationship between the 'the thinking brain' and 'the feeling brain', the concept of 'flipping our lid' and the process of self-regulation through reconnecting the 'thinking' and 'feeling' parts of the brain.		
	How the stress response affects children	Uses ppt slide to explain how child's behaviour can trigger parent's stress response, and how parent's negative response can escalate the situation.		
		Reinforces that stress response is automatic AND controllable with		

RA	Component	Topic	Done(✓)	Comments
		learned skills, and parents can learn to prevent ‘flipped lids’.		
	Identifying parents’ stress triggers	Helps parents identify their own stress triggers.		
		Helps parents identify things they can do to manage their stress.		
		Helps parents understand that when they can identify their stress, they can learn to regulate their reaction to it.		
	Understanding long-term goals (LTGs)	Conducts scenario (pg. 59) imagining child as a young adult and generates responses from the parents.		
		Records parents responses about the kind of person and relationship they hope for their child. NOTES that these are parents Long Term Goals.		
		Labels the Flipchart (Long Term Goals) and posts it in a prominent location. Stresses that these LTG are the foundation of PDEP.		
	The slow route or the fast route?	Explains the difference between quick responses to short term stresses vs staying conscious of their LTGs.		
		Helps parents see short term challenges as opportunities to build skills and knowledge.		
		Reinforces that parental self-regulation is key.		
	Closing	Provides formal closing with summary of key points, answers questions; Offers ‘between session’ suggestions for things to try at home; may provide a mini evaluation.		

Summary Note Day 2:

Session 3: Providing Warmth and Structure

RA	Component	Topic	Done(✓)	Comments
	Welcome and Opening	Establishes welcoming atmosphere; asks about parents’ thoughts or experience re: PD over the intervening week; provides a brief opener activity.		
		Introduces session topic and process.		
	The Importance of Warmth	Parents complete exercise “ Why is Warmth Important? ”.		

RA	Component	Topic	Done(✓)	Comments
		Leads group in discussion about Teacher A or Teacher B. Discusses the reasons for their choices.		
		Paraphrases, clarifies, summarizes parent responses; keeps discussion on track.		
		Identifies that Warmth = safety, security, trust		
		Helps parents see the link between learning and feeling safe and secure.		
	What is Warmth?	Using ppt slide, explains the components of “Warmth”		
		Draws parents’ attention to how they felt with Teacher A and Teacher B to help them understand the elements of warmth.		
	Providing Warmth	Uses ppts to illustrate examples of Warmth, expands on several concepts, links back to Warmth poster.		
		Conducts “Giving Warmth” exercise; leads discussion based on parents’ responses. Draws attention to the variety of ways parents can provide Warmth. Links back to Warmth poster.		
	The Importance of Structure	Explains that Warmth and Structure must ALWAYS go together – neither is sufficient on their own.		
		Parents complete exercise “ Why is Structure important? ”		
		Leads group in discussion about Teacher A or Teacher B. Discusses reasons for their choices.		
		Paraphrases, clarifies, summarizes parent responses; keeps discussion on track.		
		Identifies that Structure = Information, Guidance, Teaching and Support		
		Helps parents see the link between providing structure to interest, motivation and learning.		
	What is Structure?	Using ppt slide, explains the components of “Structure”.		
		States that punishment is NOT structure.		
		Draws parents’ attention to how they felt with Teacher A and Teacher B to		

RA	Component	Topic	Done(✓)	Comments
		help them see the elements of structure.		
	Structure as Scaffolding	Uses ppt to explain ‘scaffold’ and relates it to providing guidance/help with providing enough information to help, but not take over.		
	Providing Structure	Uses ppts to illustrate examples of Structure, expands on several concepts, links back to Structure poster.		
		Conducts “Giving Structure” exercise; leads discussion based on parents’ responses. Draws attention to the variety of ways parents can provide Structure. Links back to Structure poster.		
	Closing	Provides formal closing with summary of key points, answers questions; Offers ‘between session’ suggestions for things to try at home; may provide a mini evaluation.		

Summary Note Day 3:

Session 4: Understanding how Children Think and Feel (B to 2 years)

RA	Component	Topic	Done(✓)	Comments
	Welcome and Opening	Establishes welcoming atmosphere; asks about parents’ thoughts or experience re: PD over the intervening week; provides a brief opener activity.		
		Introduces session topic and process.		
		Provides guidelines for parents’ comfort and safety during imagination scenarios: Encourages anyone feeling distress to open eyes, be aware of their surroundings, and know that they are safe in this space. Encourages parents to talk to leader if they are uncomfortable with feelings that arise.		
	Land of the Giants	Conducts Part A of Land of the Giants Scenario . Keeps parents focused on their own experience during the exercise.		

RA	Component	Topic	Done(✓)	Comments
		Gathers parents' responses (feelings, behaviour, relationship and need) and lists them clearly on Flipchart #1.		
		Conducts Part B of Land of Giants Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents' responses (feelings, behaviour, and relationship) and lists them clearly on Flipchart #2.		
		Compares and contrasts the responses from the two scenarios, focusing on who provided warmth/structure and what impact that had on behaviour and the relationship.		
		Emphasizes the importance of attachment and trust.		
	Development Birth to 6 months	Makes transition to child development (e.g., "this is a similar situation for young babies...")		
		Covers key child development information for this stage, focusing on safety, security, trust and attachment.		
		Reviews basics of brain development at this age (brainstem at birth, development of senses, importance of consistent, nurturing care to developing sense of trust)		
		Introduces the Building Blocks of Healthy Development and posts the foundational blocks: Trust and Attachment .		
	Trying to Communicate	Conducts Part A of Trying to Communicate Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents' responses (feelings, behaviour, relationship and need) and lists them clearly on Flipchart #1.		
		Conducts Part B of Trying to Communicate Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents' responses (feelings, behaviour, and relationship) and lists them clearly on Flipchart #2.		
		Compares and contrasts the responses from the two scenarios, focusing on		

RA	Component	Topic	Done(✓)	Comments
		who provided warmth/structure and what impact that had on behaviour and the relationship.		
		Emphasizes the different responses to stress (fight, flight, freeze).		
	Development 6 – 12 months	Makes transition to child development (e.g., "this is a similar situation for older babies...")		
		Covers key child development information for this stage, focusing on increasing abilities, development of the senses and ability to communicate.		
		Reviews basics of brain development at this age (attachment, language centre, importance of frequent experiences, connection between all the sense receptors)		
		Reviews the Building Blocks of Healthy Development (notes foundational blocks Trust and Attachment), and places next block "Respectful Communication Skills" noting how healthy respectful communication develops out of a trusting relationship and positive role modeling and how, in turn, this strengthens the attachment/trust.		
	The Joy of Independence	Conducts Part A of The Joy of Independence Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents' responses (feelings, behaviour, relationship and need) and lists them clearly on Flipchart #1.		
		Conducts Part B of The Joy of Independence Scenario . Keeps parents focused on their own experience during the exercise		
		Gathers parents' responses (feelings, behaviour, and relationship) and lists them clearly on Flipchart #2.		
		Compares and contrasts the responses from the two scenarios, focusing on who provided warmth/structure and what impact that had on behaviour and the relationship.		

RA	Component	Topic	Done(✓)	Comments
		Emphasizes the importance of recognizing and supporting child's growing sense of independence for healthy physical and emotional development.		
	Development 1 to 2 years	Makes transition to child development (e.g., "this is a similar situation for young toddlers...")		
		Covers key child development information for this stage, focusing on: increasing mobility growing independence the relationship between trust/security and exploration learning through experimentation/safety		
		Reviews basics of brain development at this age (primed for learning, need lots of chances to explore, "thinking brain" can NOT yet understand danger)		
		Reviews the Building Blocks of Healthy Development (Trust and Attachment, Respectful Communication Skills), and notes how development is a continuous process with each step building on and strengthening the previous. Adds "Independence" and "Non-violent Conflict Resolution" blocks noting that children learn what is respected and modeled.		
	Closing	Provides formal closing with summary of key points, answers questions; Offers 'between session' suggestions for things to try at home; may provide a mini evaluation.		

Summary Note Day 4:

Session 5: Understanding how Children Think and Feel (2 to 5 years)

RA	Component	Topic	Done(✓)	Comments
	Welcome and Opening	Establishes welcoming atmosphere; asks about parents' thoughts or experience re: PD over the intervening week; provides a brief opener activity.		
		Introduces session topic and process.		

RA	Component	Topic	Done(✓)	Comments
	Fear in the Night	Conducts Part A of Fear in the Night Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents' responses (feelings, behaviour, relationship and need) and lists them clearly on Flipchart #1.		
		Conducts Part B of Fear in the Night Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents' responses (feelings, behaviour, and relationship) and lists them clearly on Flipchart #2.		
		Compares and contrasts the responses from the two scenarios, focusing on who provided warmth/structure and what impact that had on behaviour and the relationship.		
		Emphasizes the importance of listening to and respecting others' feelings for healthy emotional development/behavioural expression and strong relationships.		
	Development 2 to 3 years	Makes transition to child development (e.g., "this is a similar situation for older toddlers who are just learning about feelings...")		
		Covers key child development information for this stage, focusing on: wants to be independent; needs to explore easily frustrated when they can't yet explain what they need Says NO – a lot! just starting to learn about feelings Fear is big emotion		
		Reviews basics of brain development at this age - intense development time for "feeling brain"; helps to give feelings a name; takes time to learn to show emotions in ways that won't hurt self or others; building pathway between 'feeling' and 'thinking' brain		
		Reviews the Building Blocks of Healthy Development (Trust and Attachment, Respectful Communication Skills, Independence,		

RA	Component	Topic	Done(✓)	Comments
		Non-conflict Resolution), and notes how development is a continuous process with each step building on and strengthening the previous. Adds “Respect for Feelings” block noting that when toddlers learn to understand their feelings, they can start to learn how to cope.		
	A New Job	Conducts Part A of A New Job Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents’ responses (feelings, behaviour, relationship and need) and lists them clearly on Flipchart #1.		
		Conducts Part B of A New Job Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents’ responses (feelings, behaviour, and relationship) and lists them clearly on Flipchart #2.		
		Compares and contrasts the responses from the two scenarios, focusing on who provided warmth/structure and what impact that had on behaviour and the relationship		
		Emphasizes that when people feel respected and supported, they are motivated to learn.		
	Development 3 to 5 years	Makes transition to child development (e.g., “this is a similar situation for preschoolers who are just learning to take initiative...”)		
		Covers key child development information for this stage, focusing on: curiosity confidence wanting to help in own way vivid imagination		
		Reviews basics of brain development at this age – development of neural pathways; stronger connection between the ‘feeling’ and ‘thinking’ brain, fostered by strong, loving, respectful relationships, information and opportunities to practice.		

RA	Component	Topic	Done(✓)	Comments
		Reviews the Building Blocks of Healthy Development (Trust and Attachment, Respectful Communication Skills, Independence, Non-conflict Resolution, Respect for Feelings), and notes how development is a continuous process with each step building on and strengthening the previous. Adds “ Seeking Information ” block noting that scaffolded experiences, opportunities to practice, and getting questions answered are critical elements for learning.		
	Closing	Provides formal closing with summary of key points, answers questions; Offers ‘between session’ suggestions for things to try at home; may provide a mini evaluation.		

Summary Note Day 5:

Session 6: Understanding how Children Think and Feel (middle childhood)

RA	Component	Topic	Done(✓)	Comments
	Welcome and Opening	Establishes welcoming atmosphere; asks about parents’ thoughts or experience re: PD over the intervening week; provides a brief opener activity.		
		Introduces session topic and process.		
	Reading Together	Conducts Part A of Reading Together Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents’ responses (feelings, behaviour, relationship and need) and lists them clearly on Flipchart #1.		
		Conducts Part B of Reading Together Scenario . Keeps parents focused on their own experience during the exercise.		
		Gathers parents’ responses (feelings, behaviour, and relationship) and lists them clearly on Flipchart #2.		
		Compares and contrasts the responses from the two scenarios, focusing on who provided warmth/structure and		

RA	Component	Topic	Done(✓)	Comments
		what impact that had on behaviour and the relationship.		
		Emphasizes that when people feel respected and supported, they are motivated to learn and having one's individuality respected has a positive impact on emotions, behaviour and relationships.		
	Introduces Temperament	Makes transition to Temperament by letting parents know: temperament is a part of personality; present from birth makes individuals unique no 'good' or 'bad' temperament – it just 'is'. present from birth can cause challenges for children as they are exposed to more adults.		
	Seven dimensions of Temperament	Using PPT slides, explains 7 dimensions of temperament (Activity, Regularity, Response to New Situation, Adaptability, Distractibility, Persistence, Intensity).		
	Understanding Your Temperament	Asks parents to complete “ Your Temperament ” exercise. Conducts temperament activity: has parents align with their specific temperament on each dimension draws parents' attention to how they can be high on one dimension and low on another normalizes all temperaments provides clarification, answers questions, and keeps discussion respectful		
	Advantages and Challenges of Your Temperament	During discussion, asks parents to consider times when their temperament can be an advantage for them and times when it can be a challenge and how they have learned to cope. Keeps discussion respectful.		Note: Both the longer version outlined in the manual, OR the shorter version outlined on this form are acceptable.
	Understanding Your Child's Temperament	Has parents complete “ My Child's Temperament ”		
	Advantages and Challenges of Your Child's Temperament	Has parents consider the advantages or challenges that their children may face as a result of their temperament. Discusses 'labeling' and 'potential for		

RA	Component	Topic	Done(✓)	Comments
		punishment' when temperament is misunderstood. Leads discussion on how parents can help children learn to manage their temperamental challenges.		
	The Match	Asks parents to complete the exercise "The Match". Guides parent discussion on how similarities and differences can lead to conflict or getting along.		
	Development Between 5 and 9 Years	Makes transition to child development. Children's unique temperament can be source of conflict at school if they are poor match with teachers/other adults.		
		Covers key child development information for this stage, focusing on: now exposed to more friends and other adults temperament can create challenges Warmth (acceptance) and Structure (gradually building coping strategies) can help manage challenges		
		Reviews the Building Blocks of Healthy Development (Trust and Attachment, Respectful Communication Skills, Independence, Non-conflict Resolution, Respect for Feelings, Seeking Information), and notes how development is a continuous process with each step building on and strengthening the previous. Adds " Confidence " block, stressing their confidence in their abilities grows as they learn more about themselves and how to solve issues in their lives.		
	Closing	Provides formal closing with summary of key points, answers questions; Offers 'between session' suggestions for things to try at home; may provide a mini evaluation.		

Summary Note Day 6:

Session 7: Understanding how Children Think and Feel – (late childhood and adolescence)

RA	Component	Topic	Done(✓)	Comments
	Welcome and Opening	Establishes welcoming atmosphere; asks about parents' thoughts or experience re: PD over the intervening week; provides a brief opener activity.		
		Introduces session topic and process.		
	Nowhere to Turn	Conducts Part A of Nowhere to Turn . Keeps parents focused on their own experience during the exercise.		
		Gathers parents' responses (feelings, behaviour, relationship and need) and lists them clearly on Flipchart #1.		
		Conducts Part B of Nowhere to Turn . Keeps parents focused on their own experience during the exercise.		
		Gathers parents' responses (feelings, behaviour, and relationship) and lists them clearly on Flipchart #2.		
		Compares and contrasts the responses from the two scenarios, focusing on who provided warmth/structure and what impact that had on behaviour and the relationship		
		Emphasizes that when people feel respected and supported, they are motivated to learn; and that empathy and understanding can strengthen relationship.		
		Draws parents' attention to the power of joint problem-solving.		
	Development 10 and 13 years	Conducts "The Imaginary Child (Fred)" exercise: Divides parents into small groups Provides each group with prepared flipchart as described on p. 164 Introduces and leads group through the stages of the activity stating with Introductions, then exploration of physical and psychological changes typical of this age.		
	Understanding Family Conflict at this Stage	Explores potential areas of family conflict at this age and relates to developmental changes.		

RA	Component	Topic	Done(✓)	Comments
	Building a Safety Net with Warmth	Helps parents explore ways to reduce conflict with warmth.		
	Encouraging Independent Decision-Making with Structure	Helps parents explore ways to encourage independent decision-making with structure.		
	Demonstrates respect of children	Shows respect for children throughout discussion by normalizing developmental stage, speaking respectfully of children, and by respectfully gathering up posters at end of session.		
	Reviews the Building Blocks of Healthy Development (BBHD)	Refers to BBHD wall charts (Trust and Attachment, Respectful Communication Skills, Independence, Non-conflict Resolution, Respect for Feelings, Confidence), and notes how development is a continuous process with each step building on and strengthening the previous. Reinforces that practicing decision-making in a safe, supportive environment builds confidence.		
	Remembering Adolescence	Leads discussion on ‘Remembering Adolescence’ : Asks parents to recall conflict with parents from own adolescence Debriefs by examining teen and parents’ perspectives Reframes behaviour in terms of adolescent’s need for making own decisions as a developmental norm Shows respect for opinions/choices of both participants and their parents.		
	Development 15 to 18 years	Makes transition to child development for this age/stage.		
		Covers key development information for this stage, focusing on: Competing desire for both independence and acceptance Autonomy and forming own identity Risk taking Sexuality		
		Reviews basics of brain development at this age:		

RA	Component	Topic	Done(✓)	Comments
		Neural pathways strengthening, faster conductivity Synaptic pruning makes brain work more efficiently Pre-frontal cortex is now the area of high development – not yet able to fully comprehend risk.		
		Refers to BBHD wall charts (Trust and Attachment, Respectful Communication Skills, Independence, Non-conflict Resolution, Respect for Feelings, Confidence), and adds last building block “ Identity ”. Notes how development is a continuous process with each step building on and strengthening the previous. Parent role to provide safe environment/support for teens to make independent decisions.		
	Supporting Adolescents’ Development with Warmth and Structure	Explains Warmth in terms of respectful/non-intrusive monitoring; Structure in terms of information, guidance and support; listening, encouraging and collaborative problem-solving.		
	Closing	Provides formal closing with summary of key points, answers questions; Offers ‘between session’ suggestions for things to try at home; may provide a mini evaluation.		

Summary Note Day 7:**Session 8: Problem Solving and Responding with Positive Discipline**

RA	Component	Topic	Done(✓)	Comments
	Welcome and Opening	Establishes welcoming atmosphere; asks about parents’ thoughts or experience re: PD over the intervening week; provides a brief opener activity.		
		Introduces session topic and process.		
	Being a Problem Solver	Using the ppt slides, briefly reviews what parents have already learned by reviewing the main points of each of the blocks in the PDEP model that have already been covered (Long Term Goals, Warmth, Structure, Understanding How Children Think and Feel).		

RA	Component	Topic	Done(✓)	Comments
		Helps parents understand that while PD is important for all interactions, it has a great impact in conflict situations.		
		Reviews that the first step is for the parent to self-regulate – calm themselves, then help their child get to calm, then they can effectively problem solve.		
	Remembering Long Term Goals	Demonstrates how to use the PD Model when there are problems: Identify the problem (just the facts, without blaming or judging) Think about all the possible reasons for the child's behaviour (How Children Think and Feel) Remember your Long-Term Goals and how to eventually get there using Warmth and Structure Use what you know about Warmth and Structure to think through possible solutions to the problem (Problem Solving)		
		Reminds parents when they model problem solving they are teaching their children a skill for life.		
	Shifting the Focus from Children's Behaviour to its Underlying Reasons	Reminds parents that children's behaviour most often reflects their brain development, their developmental stage, and their individual temperament.		
		Has both the PD Model and the Building Blocks of Healthy Development chart clearly visible.		
	Practicing Using Positive Discipline	Encourages parents to practice with simple problems first to get comfortable with the process.		
	Problem Solving Exercises	Conducts Problem-solving Exercise #1 with appropriate handout: Asks parents to define the problem (without blaming or judging) Asks parents to consider what they know/have learned about children this age – reminds to consider child's developmental stage and temperament. Lists reasons parents generate.		

RA	Component	Topic	Done(✓)	Comments
		Asks parents to remember their LTGs. Discusses potential solutions that provide warmth and structure in this situation. Reflects that there are MANY reasons for behaviour and solution depends on what the problem is. Reinforces that it isn't defiance or being bad – behaviour is simply a reflection of how children think and feel. Draws a link to the parents' experiences in Part A/Part B of the practice scenarios.		
	Working Through the Exercises	Repeats process above for two other age groups: 5 to 9 years old 14 to 18 years old		
	Involving Children in Problem Solving	Explains to parents that children need to be increasingly involved in the problem-solving process in accordance with their developmental abilities so that they eventually learn to problem solve on their own.		
	Problem Solving for Parents' Unique Challenges	Uses a couple of parents' own questions from Question Box (Session 1) and leads group through problem solving process.		
	Ending the formal program	Provides encouragement about parents' growth over the program. Reminds them of Session 9 (practice session) if applicable.		
	Post Program Questionnaire	Explains: Rationale		
		Anonymity		
		Confidentiality		
		Freedom to participate or not		
		Provides quiet atmosphere		
		Confidential collection		
		Secures completed forms		
	Closing	If Session 9 is planned , invites parents to follow-up session in 2 weeks. Provides information on timing and location. Asks them to: read Parent Book choose one minor real-life challenge to work through, and try to apply the PD		

RA	Component	Topic	Done(✓)	Comments
		model making notes on what they still have questions about OBSERVER: If Session 9 is not planned, enter N/A in the checklist		
		Formally closes the program with a small activity to encourage parents' final thoughts. Reminds parents of other community resources and encourages them to continue on learning journey in parenting.		

Summary Note Day 8:**Session 9: Follow Up Session (Optional)**

RA	Component	Topic	Done(✓)	Comments
	Welcome and Opening	Establishes welcoming atmosphere; asks about parents' thoughts or experience re: PD over the intervening weeks; provides a brief opener activity.		
		Introduces session topic and process.		
		Reviews the PD Model and the Problem-Solving steps.		
	Sharing Positive Discipline Plans	Encourages parents to work together in pairs to review their homework.		
		Leads a large group discussion of sharing parents' individual plans. Encourages additional perspectives and suggestions.		
	Ending the Program	Reminds parents of personal growth throughout the program.		
	Closing	Provides formal closing with summary of key points, answers questions; Offers suggestions for additional supports such as follow-up parenting programs, drop in programs, etc.; may provide a mini evaluation.		

Appendix S:**Socio-Demographic Questionnaire – Study 3**

You will be asked to complete this questionnaire several times. To link your answers over time while also making sure that your answers remain anonymous, you will use a unique 6-character code based on the following:

Month of birth

Day of birth

First 2 letters of your first name given to you at birth

For example, if Mary was born on January 8, her code would be 0108MA

Please enter your code here: _____

What agency are you involved with for the PDEP program?

- (1) Vanier Community Service Center
- (2) Haldimand-Norfolk R.E.A.C.H
- (3) Pinecrest-Queensway Health and Community Services
- (4) Orléans-Cumberland Community Resource Centre
- (5) Lowertown Community Resource Center
- (6) Mothercraft Ottawa
- (7) Pembroke Early ON Child & Family Centre
- (8) Community Resource Center Killaloe Toy Bus
- (9) Arnprior Family Preschool Resource Centre

When are you completing this questionnaire?

- ___ Before PDEP has started
- ___ I just finished the PDEP group
- ___ I am still waiting to take PDEP but the first group just finished
- ___ I finished the PDEP group about 1 month ago
- ___ I am still waiting to take PDEP but the first group finished about 1 month ago
- ___ I finished the PDEP group about 3 months ago
- ___ I am still waiting to take PDEP but the first group finished about 3 months ago

What is your gender?

- ___ Female
- ___ Male
- ___ You don't have an option that applies to me. I identify as (please specify) _____

What is your age? _____ years old

Provide the following information about individuals who currently live in your home.

Age (Years) Relationship to you

_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

What is your current marital status?

Married
 Living with a partner
 Single (never married)
 Widowed
 Separated
 Divorced

What is your racial or ethnic background?

Caucasian (e.g., European descent)
 African Canadian
 Asian (e.g., China)
 South and Southeast Asian (e.g., India)
 First Nations
 Middle Eastern
 More than one race/ethnicity (Please specify which ones) _____
 Other (Please specify) _____

What language is most often spoken at home?

English
 French
 Other (please specify) _____

What is your highest education?

- (a) Less than high school
- (b) High school diploma or equivalent (GED)
- (c) Community college
- (d) Bachelor or undergraduate degree
- (e) Master's degree (e.g., M.A., M.Sc., M.Ed.)
- (f) Doctorate degree or equivalent
- (g) Other (Please specify) _____

If you are currently living with a partner, what is their highest education?

- (a) Less than high school
- (b) High school diploma or equivalent (GED)
- (c) Community college
- (d) Bachelor or undergraduate degree
- (e) Master's degree (e.g., M.A., M.Sc., M.Ed.)

POSITIVE DISCIPLINE

307

(f) Doctorate degree or equivalent

(g) Other (Please specify) _____

What is your employment?

(a) Not employed outside the home

(b) Employed part-time

(c) Employed full-time

(d) Other (please specify) _____

If you are currently living with a partner, what is their employment?

(a) Not employed outside the home

(b) Employed part-time

(c) Employed full-time

(d) Other (please specify) _____

Which of the following groups does your household income fall into (before income tax)?

Less than \$10,000

\$10,000 - \$19,999

\$20,000 - \$29,999

\$30,000 - \$39,999

\$40,000 - \$49,999

\$50,000 - \$59,999

\$60,000 - \$69,999

\$70,000 - \$79,999

\$80,000 - \$89,999

\$90,000 - \$99,999

\$100,000 - \$109,999

\$110,000 - \$119,999

\$120,000 - \$129,999

\$130,000 - \$139,999

\$140,000 - \$149,999

Over \$150,000

Is your partner (or other person responsible for your children) taking the PDEP program with you (or planning to take PDEP)?

_____ No

_____ Yes

If you answered yes and they are also participating in the research study, what is their code (6 digits)
Month of birth/Day of birth/First 2 letters of the first name given to them at birth

Have you participated in any other parenting programs either at your Early ON Child and Family agency or elsewhere?

_____ No

_____ Yes

If yes, check all the programs that you have participated in.

_____ Headstart nursery school

_____ Canadian prenatal nutrition program

_____ Play and learn drop-ins

POSITIVE DISCIPLINE

- _____ Dad & me spaghetti and gym night
- _____ Infant massage
- _____ Under one club
- _____ Game On youth program
- _____ Nobody's Perfect
- _____ Mini Mind Masters
- _____ Kids Have Stress Too
- _____ Stir it up
- _____ How to talk so kids will listen and listen so kids will talk
- _____ Terrific Toddlers
- _____ Setting Consequences
- _____ Make the Connection
- _____ Parent Child Mother Goose
- _____ Parenting books
- _____ Parenting websites
- _____ Parent outreach program (home visiting support)
- _____ Shelter respite programs
- _____ Early ON playgrounds
- _____ Early ON parent workshops (please list _____)
- _____ Circle of Care parenting program
- _____ Buns in the oven (Ça Mijote)
- _____ In Home Intervention
- _____ Beyond The Basics
- _____ COPE
- _____ Mom's Talk
- _____ Canada Prenatal Nutrition Program- Healthy
- _____ Community Action Program for Children-
- _____ Young Parents Program
- _____ Moms Eating Well for 2
- _____ Other (please list _____)

Appendix T:**DERS -SF- Reduced version****Impulse**

When I'm upset, I become out of control.

When I'm upset, I lose control over my behavior.

When I'm upset, I have difficulty controlling my behavior.

Strategies

When I'm upset, it takes me a long time to feel better.

When I'm upset, I believe there is nothing I can do to make myself feel better.

When I'm upset, I believe that I will end up feeling very depressed.

Awareness

I care about what I am feelings.

When I'm upset, I acknowledge my emotions.

I pay attention to how I feel.

Response categories:

1 = Almost never 0-10%

2 = Sometimes 11-35%

3= About half the time 36-65%

4 = Most of the time 66-90%

5 = Almost always 91-100%

Appendix U:

Parenting Daily Hassles Scale

The statements below describe a lot of events that routinely occur in families with young children.

These events sometimes make life difficult.

Please read each item and circle how often it happens to you (rarely, sometimes, a lot, or constantly) FOR THE PAST MONTH.

Also, how much of a 'hassle' or problem do you feel that it has been for you FOR THE PAST MONTH?

If you have more than one child, these events can include any or all of your children.

1. Continually cleaning up messes of toys or food
2. Being nagged, whined at, complained to
3. Meal-time difficulties with picky eaters, complaining etc.
4. The kids won't listen or do what they are asked without being nagged
5. Baby-sitters are hard to find
6. The kids schedules (like pre-school or other activities) interfere with meeting your own household needs
7. Sibling arguments or fights require a 'referee'
8. The kids demand that you entertain them or play with them
9. The kids resist or struggle with you over bed-time
10. The kids are constantly underfoot, interfering with other chores
11. The need to keep a constant eye on where the kids are and what they are doing
12. The kids interrupt adult conversations or interactions
13. Having to change your plans because of unprecedented child needs
14. The kids get dirty several times a day requiring changes of clothing
15. Difficulties in getting privacy (eg. in the bathroom)
16. The kids are hard to manage in public (grocery store, shopping centre, restaurant)
17. Difficulties in getting kids ready for outings and leaving on time
18. Difficulties in leaving kids for a night out or at school or day care
19. The kids have difficulties with friends (eg. fighting, trouble, getting along, or no friends available)
20. Having to run extra errands to meet the kids needs

Response categories (Frequency):

- 1 = Rarely
- 2 = Sometimes
- 3 = A lot
- 4 = Constantly

Response categories (Intensity):

Hassle (Low to High)

- 1 2 3 4 5

Appendix V:**Problem Solving Inventory**

Please respond to the items as honestly as possible. Your responses should reflect what you actually do to solve problems, not how you think you should solve them. When you answer these questions, think about parenting problems with your child or children.

1. When a solution to a problem did not work, I do not examine why.
2. When I have a complex problem, I do not develop a strategy to collect information, so I can figure out exactly what the problem is.
3. After I have solved a problem, I do not analyze what went right or what went wrong.
4. After I have tried to solve a problem in a certain way, I take time and compare what happened to what I thought should have happened.
5. When I have a problem, I think up as many possible ways to handle it as I can.
6. When I have a problem, I consistently examine my feelings to find out what is going on in the situation.
7. When I have a problem, I tend to do the first thing that I can think of to solve it.
8. When deciding on a solution to a problem, I do not take time to consider the chances of each possible solution being successful.
9. When I have a problem, I stop and think about it before deciding on a next step.
10. When I have a problem, I usually go with the first good idea that comes to my mind.
11. When making a decision for a problem, I weigh the consequences of each alternative and compare them against each other.
12. When trying to solve a problem, I try to predict the overall result of carrying out a particular solution.
13. When I try to think up possible solutions to a problem, I do not come up with very many ideas.
14. When I have to solve a problem, I have a systematic (organized) way for comparing alternatives and making decisions.
15. When I have a problem, I do not usually think about how things in my environment might be affecting the problem.
16. When I am confused by a problem, one of the first things I do is examine the situation and consider all the important pieces of information.

Response categories:

- 1 = I strongly disagree
- 2 = I mostly disagree
- 3 = I somewhat disagree
- 4 = I somewhat agree
- 5 = I mostly agree
- 6 = I strongly agree

Appendix W:
Parenting Young Children Scale -Modified

For each question rate, how often you were able to use each parenting strategy during the LAST MONTH.

	Not at all						Most of the time
1. Play with your child in a way that was fun for both of you?							
2. Speak calmly with your child when you were upset with him or her?							
3. Warn your child before a change of activity was required?							
4. Stand back and let your child work through problems s/he might be able to solve?							
5. Explain what you wanted your child to do in clear and simple ways?							
6. Plan ways to prevent problem behavior?							
7. Invite your child to play a game with you or share an enjoyable activity?							
8. Tell your child what you wanted him							

or her to do rather than tell him/her to stop doing something?							
9. Give reasons for your requests?							
10. Teach your child new skills?							
11. Tell your child how you expected him or her to behave?							
12. Make a game out of everyday tasks so your child followed through?							
13. Involve your child in household chores?							
14. Break a task into small steps?							
15. Notice and comment to your child when s/he did something well or showed a new skill?							
16. Prepare your child for a challenging situation?							