Psychopathological and Neuropsychological Endophenotypes of Children with Disruptive Behaviour

by

Lee Propp

A thesis submitted in conformity with the requirements for the degree of Master of Arts
Graduate Department of Applied Psychology and Human Development
Ontario Institute for Studies in Education
University of Toronto

© Copyright 2018 by Lee Propp
Abstract

Psychopathological and Neuropsychological Endophenotypes of Children with Disruptive Behaviour

Lee Propp
Master of Arts
Graduate Department of Applied Psychology and Human Development
Ontario Institute for Studies in Education
University of Toronto
2018

Childhood Disruptive Behaviour Disorders (DBD) are a robust predictor of delinquency, academic failure, social difficulties, and persistent mental health disorders in adolescence and adulthood. Children with DBDs represent a highly heterogeneous group and the present study aimed to distill heterogeneity by identifying data-driven subgroups based on key dimensions of psychopathology and neuropsychological abilities. Data from children 6 – 12 years (n = 143), referred to a clinic for children with DBDs, were analyzed using Latent Class Analysis (LCA) to determine classes (i.e., subgroups) based on measures of children’s callous-unemotional (CU) traits, emotion regulation, and objective Executive Function (EF) task performance. Classes were compared on parent- and teacher-rated level of functional impairment, social functioning, and conduct problems. Classes with elevated levels of CU traits and poor selective attention had the highest conduct problems and impairment. This research may inform more refined assessment processes and future treatment innovation for children with DBDs.
Dedication

For my parents. The ones who believed in me, even when I didn’t believe in myself.
Acknowledgements

It occurs to me, as I write this, that an ‘acknowledgement’ doesn’t capture the gratitude I wish to express to the many people who have supported me on my academic journey thus far. First, and foremost, I would like to thank my supervisors Dr. Brendan Andrade and Dr. Anne-Claude Bedard. I consider myself one of the lucky few to have not one, but two, brilliant scientists supervising my graduate research. Your patience and guidance have helped instil in me an appreciation for the finer details necessary in conducting rigorous and impactful scientific research. You have both set examples that I will strive to follow.

I would also like to thank Dr. Rhonda Martinussen, my committee member, for her careful review of this manuscript. Thank you to Marcos Sanches and Dr. Madison Aitken for sharing with me their wealth of knowledge on all things statistics; I am a better scientist for it. I’d also like to recognize all of the undergraduate students and volunteers in the lab, especially Alina and Lucia, for the countless hours spent entering and checking the data that makes all of this research possible.

My family - my parents, to whom I have dedicated this thesis, my brothers and sister - deserve extra thanks. Mom and Dad, though we joke that pursuing higher education was ‘mandatory’ for the four of us, it is really our way of thanking you for inspiring within us a love of learning and for investing in our education(s) that will remain with us forever. I couldn’t have done it without you.

This work was supported by the Frederick Banting and Charles Best Canada Graduate Scholarship-Canadian Institutes of Health Research and the OISE Graduate Fellowships.

As I continue my training, it is my hope to always remain true to why we, as scientists, do this work. That I will always maintain the distinction between a pile of bricks and a true edifice and avoid contributing to the ‘chaos in the brickyard’ (Forscher 1963).
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>iv</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>vii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>List of Acronyms</td>
<td>x</td>
</tr>
<tr>
<td><strong>1 Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Disruptive Behaviour Disorders</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1 History and Definitions</td>
<td>2</td>
</tr>
<tr>
<td>1.1.2 Current Diagnostic Practices and Limitations</td>
<td>4</td>
</tr>
<tr>
<td>1.2 Heterogeneity</td>
<td>4</td>
</tr>
<tr>
<td>1.2.1 Endophenotypes</td>
<td>6</td>
</tr>
<tr>
<td>1.2.2 Data-Driven Approaches</td>
<td>7</td>
</tr>
<tr>
<td>1.3 Neuropsychological Abilities</td>
<td>9</td>
</tr>
<tr>
<td>1.3.1 Emotion Recognition</td>
<td>11</td>
</tr>
<tr>
<td>1.3.2 Selective Attention</td>
<td>12</td>
</tr>
<tr>
<td>1.3.3 Response Inhibition</td>
<td>13</td>
</tr>
<tr>
<td>1.4 Dimensions of Psychopathology</td>
<td>14</td>
</tr>
<tr>
<td>1.4.1 Emotion Regulation</td>
<td>14</td>
</tr>
<tr>
<td>1.4.2 Callous-Unemotional (CU) Traits</td>
<td>16</td>
</tr>
<tr>
<td>1.5 Study Objectives</td>
<td>17</td>
</tr>
</tbody>
</table>
1.5.1 Hypotheses ......................................................... 17

2 Method .......................................................... 19

2.1 Participants and Procedure ............................................ 19
  2.1.1 Participants' Demographic Characteristics .................... 20

2.2 Measures ....................................................... 22
  2.2.1 Exclusion Criteria ............................................. 22
  2.2.2 LCA Indicators ............................................... 22
  2.2.3 Outcome Variables ............................................ 25

2.3 Analytic Plan .................................................... 27
  2.3.1 Data Cleaning ............................................... 27
  2.3.2 Latent Class Analysis (LCA) ................................ 29

3 Results ............................................................ 31

3.1 Latent Class Analysis .............................................. 31
  3.1.1 Descriptive Statistics by Class ................................ 34

3.2 Problems and Impairment Across Classes ......................... 35

3.3 Age and Sex Differences Across Classes ........................ 36

4 Discussion ........................................................ 38

4.1 Main Findings .................................................... 38
  4.1.1 Relative Difficulties and Impairment ......................... 39
  4.1.2 CU and Emotion Recognition ................................ 40
  4.1.3 Dimensions of Psychopathology .............................. 43
  4.1.4 Measurement .................................................. 46

4.2 Limitations ...................................................... 47

4.3 Future Directions and Clinical Implications ...................... 48

Bibliography ......................................................... 51
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Participant Demographics</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Means and Correlations</td>
<td>28</td>
</tr>
<tr>
<td>3.1</td>
<td>Model Fit Statistics for Latent Class Analysis</td>
<td>32</td>
</tr>
<tr>
<td>3.2</td>
<td>Average Posterior Probabilities by Latent Class</td>
<td>33</td>
</tr>
<tr>
<td>3.3</td>
<td>Percentage of Scores in the Abnormally Elevated Range on SDQ by Class</td>
<td>34</td>
</tr>
<tr>
<td>3.4</td>
<td>Between Profile Comparisons on Total Difficulties, Impairment, and Age</td>
<td>37</td>
</tr>
</tbody>
</table>
List of Figures

3.1 Indexed Means within Latent Classes ................................. 33
3.2 Diagnostic Information by Class ...................................... 35
List of Acronyms

**ADHD**  Attention-Deficit Hyperactivity Disorder.

**AIC**  Akaike Information Criterion.

**ANOVA**  Analysis of Variance.

**ASD**  Autism Spectrum Disorder.

**AST**  Attention Switching Task.

**BDEFS-CA**  Barkley’s Deficits in Executive Functioning Scale-Children and Adolescents.

**BIC**  Bayesian Information Criterion.

**BLRT**  Bootstrap Likelihood Test.

**C-DISC**  Diagnostic Interview Schedule for Children, Computer Administered.

**CANTAB**  The Cambridge Neuropsychological Testing Automated Battery.

**CD**  Conduct Disorder.

**CU**  Callous-Unemotional.

**DBD**  Disruptive Behaviour Disorders.

**DLPFC**  Dorsolateral Prefrontal Cortex.
**DSM** Diagnostic and Statistical Manual of Mental Disorders.

**EF** Executive Functions.

**ERT** Emotion Recognition Task.

**ICU** Inventory of Callous-Unemotional Traits.

**IRS** Impairment Rating Scale.


**LCA** Latent Class Analysis.

**NIMH** National Institutes of Mental Health.

**OCD** Obsessive-Compulsive Disorder.

**ODD** Oppositional Defiant Disorder.

**PDD** Pervasive Developmental Disorder.

**SDQ** Strengths and Difficulties Questionnaire.

**SSRT** Stop Signal Reaction Time.

**SST** Stop-Signal Task.

**VLMRT** Vuong-Lo Mendell Rubin Test.
Chapter 1

Introduction

1.1 Disruptive Behaviour Disorders

Childhood Disruptive Behaviour Disorders (DBD) are a robust predictor of delinquency, academic failure, social difficulties, and persistent mental health disorders in adolescence and adulthood (Copeland, Shanahan, Costello, & Angold, 2009). They are also one of the most frequent reasons for referral to children’s mental health services (Frick & Nigg, 2012). DBDs are multidimensional in clinical presentation and include oppositional, disobedient and rule-breaking behaviour, aggression, and irritability (Bollini et al., 2017); all challenges that cut across diagnostic categories, including Oppositional Defiant Disorder (ODD), Conduct Disorder (CD), as well as features of several neurodevelopmental disorders (e.g., Attention-Deficit Hyperactivity Disorder (ADHD)) (Pardini, Frick, & Moffitt, 2010). Population prevalence estimates of DBDs range from 1.4% up to 12.3% (Demmer, Hooley, Sheen, McGillivray, & Lum, 2017), with some of the variability between studies attributed to non-representative samples and inconsistent diagnostic criteria (Nock, Kazdin, Hiripi, & Kessler, 2007). In addition to heterogeneity within clinical presentation and symptomology, comorbidity is the rule, rather than the exception, with some estimates as high as 94.4% that a child with a DBD will meet criteria for at least one other disorder in their lifetime.
Boys tend to be overrepresented in samples of children with DBDs, with studies citing prevalence ratios between 2:1 and 3:1 (boys:girls); however, part of this disparity may be attributable to the differences in externalizing disorder presentation in boys and girls (e.g., boys are often more rambunctious and overtly disruptive \cite{Demmer2017}). Left untreated, DBDs also place an enormous lifetime financial burden on public expenditures such as healthcare, education, and the criminal justice system, with estimates that up to $2-million can be saved for each high-risk individual who receives early intervention and support \cite{Christenson2016}. This underscores the importance of early and accurate identification and intervention for children with DBDs \cite{Luby2012}.

### 1.1.1 History and Definitions

Historically, disruptive behaviours have been viewed by both the medical community and society at large, at least in part, as a sociological phenomenon that occurs normatively across the population \cite{Moffitt2005}. As is the case with many mental disorders, it is when such traits and behaviours reach the extreme “tails” of the normal distribution and a certain level of impairment, that they may be flagged for diagnosis as a mental disorder. For example, irritability, a dimension of dysregulated emotion, occurs in typically developing children as a normally distributed developmental phenomenon; however, in its severe form it is also a symptom of DBDs and other childhood psychopathologies \cite{Copeland2015}.

The following are the current Diagnostic and Statistical Manual of Mental Disorders (DSM) (5th ed.; DSM-5; \cite{AmericanPsychiatricAssociation2013}) diagnostic definitions of DBDs. ODD is defined as ”a pattern of angry/irritable mood, argumentative/defiant behaviour, or vindictiveness”; CD is more serious than ODD and is defined as “an ongoing pattern of behaviour than breaks social rules and may include: aggressiveness toward people and animals, destroying others property, lying or stealing, or violation of important
rules” (American Psychiatric Association, 2013). ADHD, which is highly comorbid with ODD and CD and thus included here in the DBD grouping (Nock et al., 2007), is a neurodevelopmental disorder defined as “an ongoing pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development” (American Psychiatric Association, 2013). Prior to the DSM-III, previous editions categorized DBDs into reactionary subtypes, which included the unsocialized aggressive reaction and hyperkinetic reaction in the DSM (2nd ed.; DSM-II, American Psychiatric Association, 1968)). These ‘reactions’ laid the foundation for the diagnostic categories that are known and currently used (e.g., ODD, CD, ADHD), and it was in 1980, with the next publication of the DSM (3rd ed.; DSM-III, American Psychiatric Association, 1980), that these diagnostic labels entered into the psychiatric lexicon (Pardini et al., 2010). Since then there have been a number of changes in diagnostic requirements, categorization, and labeling; however, the fundamental premises and disorder phenotypes have remained the same (Frick & Nigg, 2012). For example, ODD was originally referred to as oppositional disorder and symptom groupings have since been added to the DSM-5, but the diagnostic label still describes a very similar phenotype (American Psychiatric Association, 2013).

The developmental sequence of psychopathology continues to remain a point of contention and disagreement in the field, with some arguing that ODD is a prodromal condition in the development of CD and further severe psychopathology (Nock et al., 2007); decades of developmental research have clarified that not all children with ODD will go on to develop CD, as reviewed by Christensen and colleagues (2016). However, these perceptions have the potential to stigmatize and marginalize children and their families, and can lead to a reluctance to refer for assessment and early intervention for fear of labeling children with a “fateful and unchangeable condition”, including a future of antisocial and criminal behaviour (Sterzer, 2010). Continued research in assessment, early intervention, and developmental trajectories stands to reduce stigma and change attitudes as to the benefits of resourcing prevention and early intervention initiatives.
1.1.2 Current Diagnostic Practices and Limitations

Unlike many other areas of medicine, where a clinician has many assessment tools at their disposal, including radiologic measures and both in vivo and in vitro diagnostics, psychiatric diagnoses mainly rely on the patient’s, or their caregiver’s, subjective account of symptoms and the clinician’s observational skills. The problem is not that these are uninformative or unimportant sources of information, rather that there are limited other sources from which a clinician can explore differential diagnoses or make well-informed decisions of which treatment(s) may work optimally for an individual patient (Insel & Cuthbert, 2015). Moreover, even when a diagnosis can be made with a high degree of certainty, clinicians have few means by which to predict outcomes or individual risk for developing more severe psychopathology, thus limiting the clinical utility of the diagnostic label (Garvey, Avenevoli, & Anderson, 2016). Taken together with the immense comorbidity across diagnoses and heterogeneity within them, it is becoming increasingly clear that current psychiatric diagnostic categories need to be revised to account for evidence-informed dimensions of psychopathology, including deeper layers of data that range from behavioural to genomic (NIMH, 2015).

1.2 Heterogeneity

As mentioned, a current problem in child psychiatry is that within individuals diagnoses often exist with a high degree of overlap (Stringaris, 2012). As a result, comorbidity has become the rule rather than the exception (Clark, Watson, & Reynolds, 1995). This has prompted a call from the National Institutes of Mental Health (NIMH) for researchers to begin studying mental disorders by integrating many levels of data to study dimensions, rather than categories, of mental illness (NIMH, 2015). Adopting this
transdiagnostic research approach may refine understanding of mental illness, and thus improve evidence-based assessment and intervention practices in the future (Wakschlag et al., 2015).

In large studies investigating comorbidity amongst psychiatric disorders, of children with a lifetime diagnosis of ODD, 35% and 42.3% will meet diagnostic criteria for ADHD or CD, respectively (Nock et al., 2007). Less common, albeit still prevalent, is the comorbidity between DBDs and mood and anxiety disorders in children and adolescents (Stringaris & Goodman, 2009). Additionally, phenotypic variation (i.e., heterogeneity of symptom presentation) within diagnostic categories points to another important factor limiting the clinical utility of the current diagnostic categories (Karalunas et al., 2014).

Heterogeneity in children with DBD exists over and above currently recognized diagnostic subtypes and specifiers. Despite a diagnosis of a DBD, these children show varying levels of important dimensions of psychopathology and neuropsychological abilities. With regards to psychopathology, studies have described that children with DBDs vary in reported levels of emotion dysregulation, including irritability and emotional lability (i.e., rapid, exaggerated changes in mood), and Callous-Unemotional (CU) traits (i.e., low empathy, lack of guilt and remorse, and poverty of interpersonal emotions) (Frick, Ray, Thornton, & Kahn, 2014; Steinberg & Drabick, 2015). The variation in severity across these dimensions has emerged as potential differential specifiers, with divergent developmental trajectories, in children with DBDs (Stringaris & Goodman, 2009). For example, children who exhibit elevated levels of CU traits in early- and middle-childhood tend to experience more severe and persistent conduct and antisocial problems (Hyde et al., 2013) and greater variation in temperament (Steinberg & Drabick, 2015). Emotion dysregulation, specifically irritability and emotional lability, has been associated with later depression and anxiety (Bolhuis et al., 2017; Steinberg & Drabick, 2015). With regards to neuropsychological abilities, including Executive Functions (EF), studies have shown that children with DBDs show both common and distinct patterns of performance
on task-based indices of EF (Shanmugan et al., 2016). Distinctions in EF abilities have been related to outcomes such as sensitivity to rewards and punishments (Matthys, Vanderschuren, & Schutter, 2013), and have been hypothesized as a moderator in the relationship between dimensions of psychopathology (e.g., CU traits) and externalizing behaviours (Sulik & Obradović, 2017). As such, these dimensions have garnered interest among researchers in their utility as predictors of outcome trajectories and indices by which to distill the heterogeneity among children with DBD (Aitken, Henry, & Andrade, 2017; Bolhuis et al., 2017; Eyberg, Nelson, & Boggs, 2008).

1.2.1 Endophenotypes

An endophenotype, or intermediate phenotype, is a measurable trait that more closely approximates the true biological etiology of a disorder as compared to its clinical phenotype (i.e., the observable signs and symptoms of a given disorder) (Gottesman & Gould, 2003). Moreover, endophenotypes cannot be seen with the unaided eye and exhibit a reasonable degree of heritability (Wong, Fox, Ng, & Lee, 2011). The concept of endophenotypes, first put forth in a 1966 paper describing the genomic variability of grasshoppers (John & Lewis, 1966), has since been adapted by researchers in areas such as neuropsychiatry to begin filling in the gaps between the behavioural phenotype of a disorder and its distal genotype (Lenzenweger, 2013). Endophenotypes have been studied at length in schizophrenia research and in other areas of neuropsychiatry research such as ADHD (McAuley, Crosbie, Charach, & Schachar, 2014) and Autism Spectrum Disorder (ASD) (Szatmari et al., 2007). Examples of candidate neuropsychological endophenotypes include sensorimotor gating (i.e., filtering out redundant neural stimuli) and working memory (i.e., temporarily storing and manipulating information) in schizophrenia (Gottesman & Gould, 2003); response inhibition (i.e., inhibiting a prepotent response) and working memory in ADHD (Doyle et al., 2005; McAuley et al., 2014); emotion recognition deficits in CD (Sully, Sonuga-Barke, & Fairchild, 2015); and CU traits in DBD more generally (Hyde et al., 2016). In one study
testing the persistence of cognitive deficits in children and adolescents with ADHD highly
comorbid with DBDs, researchers found that even among disorder remitters (i.e., they no
longer met criteria for ADHD diagnosis based on symptoms and impairment), response
inhibition still remained impaired as compared to typically-developing controls. As such,
researchers concluded that response inhibition represents a candidate endophenotype
for ADHD (McAuley et al., 2014). Additional research highlights the aforementioned
dimensions of psychopathology and neuropsychological abilities as promising candidate
endophenotypes for DBD (Hyde et al., 2016; Shanmugan et al., 2016; Wiggins et al.,
2016), based on their meeting heritability criteria (i.e., are found in unaffected siblings)
and the state-independence (i.e., manifests whether or not disorder is active) of these
constructs (Gottesman & Gould, 2003).

In contrast to other disease models in medicine, genetic findings have not yet improved
the precision in assessment or treatment response prediction in mental illness, which
may in part be due to immense phenotypic heterogeneity along with the polygenic and
multifactorial origins in the development of mental illness (Plomin, DeFries, Knopik, &
Neiderhiser, 2016; Szatmari et al., 2007). Taken together with the impact and interaction
of environmental variables on the manifestation of disease, there clearly exists a need to
use alternative methodologies to study the underlying mechanisms of psychiatric disorders.
Identifying key endophenotypes for children with DBDs may be a promising methodology
by which to distill the extensive heterogeneity in DBDs.

1.2.2 Data-Driven Approaches

Variable-centered statistical approaches have traditionally been the gold standard in clinical
research, where scientists posit a priori hypotheses to test for statistical significance
between predefined groups or time points (Haapasalo, Tremblay, Boulerice, & Vitaro,
2000). In contrast, data-driven approaches allow for a posteriori reasoning by using an
algorithmic approach to find patterns in data, thereby distilling highly heterogeneous data
into subgroups (Insel & Cuthbert, 2015). A data-driven approach accounts for individual variability and distills heterogeneity by establishing subtypes of a disorder based on key characteristics of the individuals within the subgroup. Once validated, these subgroups can be used to guide a clinician’s treatment decisions and potentially reduce uncertainty in prognosis and improve likelihood of treatment response (Saria & Goldenberg, 2015).

Popular techniques for this approach include cluster and latent variable modeling (L. M. Collins & Lanza, 2010; Marquand, Rezek, Buitelaar, & Beckmann, 2016) and more recently have begun to incorporate novel algorithms, such as machine learning approaches, that allow for the integration of multiple data types (LeCun, Bengio, & Hinton, 2015). Though novel in psychiatric research, working to understand phenotypic and biological diversity by using statistical techniques to delineate subgroups within heterogeneous diseases has been widely used in medicine with cancer research being one of precision medicine’s greatest success stories (F. S. Collins & Varmus, 2015). Data-driven approaches are now beginning to be applied to mental illness (Borsboom, 2017; Insel & Cuthbert, 2015). For example, in a study measuring EF performance among children with DBDs, the subset of the sample with higher levels of anxiety showed less impairment on a response inhibition task that was heavily loaded with affective stimuli, and study authors hypothesized that children with DBDs and elevated levels of anxiety may exhibit an overcontrolled style of self-regulation (Woltering, Lishak, Hodgson, Granic, & Zelazo, 2016). This study suggested a variation in the mechanisms leading to patterns of EF task performance and highlighted a latent subgroup of the sample not captured by diagnostic categorizations.

Other examples of data-driven approaches to distill heterogeneity in psychiatric disorders have included using dimensions of temperament in children with ADHD (Karalunas et al., 2014), combining neurobiological and phenotypic data to create subgroups amongst adults with psychosis spectrum disorders (Clementz et al., 2016), and using dimensions of CU traits and aggression to parse apart DBDs (Aitken et al., 2017). Aitken and colleagues
conducted a latent profile analysis (i.e., data-driven approach), using characteristics of DBDs thought to be associated with children’s social functioning, including indices of aggression, CU traits, and levels of prosocial behaviour. The emergent profiles characterized by elevated CU traits, reactive aggression (i.e., retaliatory reactions to blocked goals), and low levels of prosocial behaviour were associated with the greatest levels of problems in both social interaction and social status when compared with other profiles (Aitken et al., 2017). Although informative, this study used parent-report measures and did not include ratings from other sources such as teachers, self-report, or objective measures of children’s abilities. Investigation of additional variables, and utilization of broader methodology, is needed to build on this research.

Given the heterogeneity of DBDs, predefining groups for research using existing diagnostic categories or a case-control study design, as found in variable-centered approaches, may not capture the true expression of the disorder or differential impairment experienced by individuals with the disorder. Moreover, the high comorbidity of DBDs with other disorders creates challenges for using diagnostic categories as pre-defined groups for study (Marquand et al., 2016). Therefore, data-driven approaches provide a viable alternative, and may create opportunities to distill heterogeneity among children with DBDs.

1.3 Neuropsychological Abilities

Neuropsychological abilities refer to the broad range of brain structures and functions that are linked to specific psychological processes, i.e., behaviour and cognition (Luria, 1966). EFs are considered to be a subset of these abilities and include the higher-order cognitive processes that are necessary for behavioural regulation and goal-directed behaviour (Li et al., 2015). EFs include a number of abilities, and the coordination of these abilities underlie success in many areas of everyday functioning, such as decision making, planning, and navigating novel and complex tasks (Macdonald, Goines, Novacek, & Walker, 2016).
Deficits in EF abilities have been implicated in a number of neurodevelopmental disorders, including ADHD (Kofler et al., 2013), ASD (Kaufmann et al., 2013), Obsessive-Compulsive Disorder (OCD) (Hybel, Mortensen, Lambek, Thastum, & Thomsen, 2017), and DBDs (Wakschlag et al., 2018). Recent research in DBD populations has identified specific EF deficits, such as response inhibition (i.e., the ability to inhibit a prepotent response) and selective attention (i.e., the capacity to selectively attend to certain stimuli) (Friedman & Miyake, 2017). For example, a recent, large-scale study cutting across multiple dimensions of psychopathology found that individuals who cluster in “behavioural” dimensions, i.e., children and youth with DBDs, show consistent impairments on tasks loading heavily on cognitive control, particularly in maintenance and adjustment within task sets and error-monitoring (Shanmugan et al., 2016). Such tasks demand a participant’s “controlled attention”, hence the term cognitive control, where the individual must continually monitor and adapt behaviour in order to attain a certain goal (Friedman & Miyake, 2017). This is consistent with established findings citing the co-occurrence of DBDs and deficits in specific EFs, including response inhibition and selective attention, which both require significant self-monitoring (Sulik & Obradovic, 2017). Furthermore, meta-analyses have demonstrated that these deficits may contribute to other behavioural symptoms in children with DBDs, including decreased punishment sensitivity and sensation-seeking behaviours and poor treatment response (Matthys et al., 2013). One of the hypothesized mechanisms centres on response perseveration and a limited ability to adapt behaviour to changes in the environment, such that children with DBDs exhibit greater difficulties stopping ongoing behaviour in response to punishment (Matthys et al., 2013). Developing a greater understanding of the level of key EF deficits across children with DBD, and those that are characteristic of certain subgroups of individuals with DBDs, may provide a basis for further refined assessment practices and treatment innovation (Puzzo, Smaragdi, Gonzalez, Martin-Key, & Fairchild, 2016). Three key EF abilities associated with DBD include: emotion recognition, selective attention, and response inhibition.
1.3.1 Emotion Recognition

Emotion recognition is the ability to identify emotions from facial expressions and is an important aspect of social cognition that enables understanding and processing of affective information (Burke, Loeber, & Birmaher, 2002). A growing body of research has linked poor emotion recognition to a number of psychiatric disorders, including schizophrenia (Eddy, 2018), CD, and antisocial behaviour (Sully et al., 2015), and ADHD (Jusyte, Gulewitsch, & Schönenberg, 2017). Moreover, it has been suggested that impaired emotion recognition may be linked to poor interpersonal functioning, though the mechanisms remain unclear (Jusyte et al., 2017).

Recent studies have found that children with DBDs show marked deficits in emotion recognition abilities that are also seen in their unaffected immediate relatives (Sully et al., 2015). Sully and colleagues studied emotion recognition in individuals with an existing diagnosis of CD, first-degree relatives of CD probands that themselves screened negative for ODD or CD (lifetime or current), and healthy controls with no familial history of ODD or CD. Both the CD groups and the unaffected relatives showed marked deficits in emotion recognition abilities as compared to the healthy controls. This finding suggests that poor emotion recognition may be a candidate endophenotype for DBD given its high heritability (Sully et al., 2015). Moreover, deficits in emotion recognition have also been shown to approximate deficits in broader domains of social understanding and thinking (i.e., social cognition). In a recent study, Jusyte and colleagues (2017) used a dynamic emotion recognition task, where all stimulus faces started as neutral and morphed into one of the six basic emotions. Participants were instructed to press a button at the exact moment they were able to identify the emotion, thus also testing for perceptual sensitivity. Those with ADHD, compared to controls, showed lower accuracy across all emotions, however no group differences emerged in perceptual intensity across correct trials; in other words, the clinical group did not require greater intensity to correctly identify the emotion (Jusyte et al., 2017). As such, emotion recognition abilities (as assessed through objective
task performance) seem to be a representative construct of social cognition, i.e., a valid and reliable construct, and not a confounding impairment in perceptual sensitivity.

1.3.2 Selective Attention

Selective attention is the ability to attend to certain stimuli while ignoring others, tapping into an individual’s top-down cognitive control (Macdonald et al. 2016). Mental flexibility, or the ability to shift between different solutions in problem solving situations, is a key area of challenge for children with DBD. For example, Secruca and Silva (2016) studied EFs in individuals convicted for criminal and antisocial behaviour, i.e., either violent or non-violent crimes, and non-offender controls. Compared to non-offenders, the offender group showed marked impairments in tasks of mental flexibility and planning, and when the offenders were subdivided by type of crime the non-violent offenders had poorer mental flexibility while the violent offenders had worse planning abilities. The authors suggested that the deficits in mental flexibility in the non-violent offenders may be a mechanism limiting their selection of options to solve real-world problems (Seruca & Silva 2016). Similarly, children with DBDs show impairments in decision-making contexts that may partially result from a similar difficulty in weighing behavioural options and potential consequences of each choice (as reviewed by Nigg 2017). Other recent reviews and meta-analyses have found similar trends in selective attention deficits, suggesting practical outcomes of said deficits that include the inability to multitask, difficulties in capacity to weigh and switch between behavioural decision options, or to ignore extraneous stimuli irrelevant to the task at hand (Alegria, Radua, & Rubia 2016; Puzzo et al. 2016). This emerging body of research suggests that selective attention may be a key candidate endophenotype of DBD and, moreover, deficits in this specific EF dimension may closely approximate real-world impairments in decision making and task completion.
1.3.3 Response Inhibition

Response inhibition is perhaps one of the most thoroughly studied EFs, specifically in the context of DBD (Kofler et al., 2013). Response inhibition, often synonymous with inhibitory control, measures an individual’s ability to inhibit a prepotent response that may be irrelevant or inappropriate (Verbruggen & Logan, 2008). Poor response inhibition has been associated with higher levels of impulsive, aggressive, and overall externalizing behaviours among children with DBD. For example, Glenn and colleagues (2017) examined this relationship in children with DBDs, with and without comorbid ADHD, and found that deficits in response inhibition did not differ as a function of comorbid ADHD. They did, however, find that deficits in response inhibition were related to higher ratings of externalizing behaviour. These findings suggest that, in children with DBDs, comorbid ADHD may not be the primary mechanism of poor response inhibition (Glenn et al., 2017). The marked deficits that children with DBD exhibit on the number and estimated speed of successful stopping on response inhibition tasks, have been hypothesized to contribute to the frequent loss of temper and aggressive responses that are common in children with DBDs (Noordermeer, Luman, & Oosterlaan, 2016). Similarly, poor response inhibition is associated with elevated engagement in risky behaviours (Nigg, 2017). In a recent review, Puzzo and colleagues (2016) provide a distinction between the often interchanged constructs of impulsivity, risk-taking, and response (dis)inhibition. They proposed that the impulsivity and risk-taking behaviours that are related to poor response inhibition are cue-activated, automatic behaviours, and not those involved in decision making involving weighing costs and benefits of each possible behavioural response. As such, the behavioural outcomes of poor response inhibition are those that occur by way of failure to inhibit the most salient response option (Puzzo et al., 2016).

Children with ADHD also show consistent and persistent deficits in response inhibition, such that it has been hypothesized to be a valid, state-independent endophenotype of ADHD (i.e., persistent even when other disorder symptoms remit) (McAuley et al., 2014).
Moreover, there is a sizable literature demonstrating that individuals with ADHD not only perform worse than controls on such tasks, but also show greater individual variability within the task itself, a finding that is confirmed both clinically and neuropsychologically (Kofler et al., 2013). Given the high degree of overlap between ADHD and other DBDs, there is a paucity of research that accounts for said comorbidity; however recent meta-analyses suggest that these findings do generalize across the DBD population and that response inhibition may be a candidate endophenotype of DBD as well (Glenn et al., 2017; Noordermeer et al., 2016).

1.4 Dimensions of Psychopathology

In addition to measures of EF discussed above, two key dimensions of psychopathology have emerged from past research as candidate endophenotypes of DBD: 1) emotion regulation and 2) CU traits.

1.4.1 Emotion Regulation

Emotion regulation is the ability to exert control over one’s emotional experience and expression (Cavanagh, Quinn, Duncan, Graham, & Balbuena, 2017). Across development, children may show varying degrees of emotion dysregulation that manifest as irritable mood, low frustration tolerance, or emotional lability (Brotman, Kircanski, Stringaris, Pine, & Leibenluft, 2017). Frequent and pervasive dysregulated emotion that occurs across contexts and is developmentally inappropriate is often associated with a DBD (Copeland et al., 2015). Clinically, emotion dysregulation and irritability are among the most common concerns for which children are referred for psychiatric assessment and intervention (Brotman et al., 2017). Deficits in this area are closely linked with increased proneness to anger and outbursts, which often lead to difficulties in interpersonal interactions and relationships (Wakschlag et al., 2018).
Recent factor analyses suggest that emotion dysregulation, currently conceptualized as a disorder specifier in the DSM-5, may actually be better understood as a core deficit in children with DBD (Cavanagh et al., 2017). Cavanagh and colleagues conducted a factor analysis of a rating scale used to measure behavioural symptoms and found that the ODD diagnostic items of emotional lability and irritability loaded highly onto the emotion dysregulation subscale in particular. The authors suggest that rather than representing a diagnostic specifier, emotion dysregulation more closely represents a core deficit in children who would meet diagnostic criteria for ODD (Cavanagh et al., 2017). Further, emotional dysregulation in children with DBD has been related to a number of neurobiological and EF abnormalities that exist in children with DBD. These include brain areas such as the Dorsolateral Prefrontal Cortex (DLPFC), striatum, and amygdala, as well as poor response inhibition and mental flexibility (Bolhuis et al., 2017; Brotman et al., 2017; Wakschlag et al., 2018). Brotman and colleagues present a translational model of irritability, a dimension of emotion dysregulation, integrating brain-based abnormalities and executive dysfunction to conceptualize mechanisms of clinically significant irritability in children and youth. The proposed pathways highlight instrumental learning, i.e., the process of learning behaviours to obtain reward or avoid punishment, as a key concept in the model. One pathway, highly relevant in children with DBDs, involves aberrant processing of frustrative nonreward - the adaptive response to blocked goal attainment. In lab-based tasks, responses typically involve increased motor response following an unexpected reward omission and activity in neural circuits involving the prefrontal cortex, striatum, and amygdala. In children with DBDs, tasks involving unexpected reward omissions tend to elicit responses of increased magnitude and duration compared to typically developing peers (Brotman et al., 2017). Thus, pathological irritability, and emotion dysregulation more generally, may present another candidate endophenotype for DBD.
1.4.2 Callous-Unemotional (CU) Traits

CU traits are a set of dimensional characteristics that include low empathy towards others, lack of guilt or remorse over behavioural transgressions, and limited emotional expression (Frick et al., 2014; Lahey, 2014). CU traits exist along a continuum that influence the severity of childhood and adolescent conduct problems and disruptive behaviours, with emerging research suggesting that the presence of elevated CU traits reflect an important subgroup of children with DBDs (Frick et al., 2014; Pardini & Fite, 2010). In a large-scale longitudinal study of school-aged boys, Pardini and Fite (2010) found that elevated CU traits were predictive of future crime and violence. Further, elevated CU traits within DBDs are predictive of greater severity and frequency of aggressive and antisocial behaviours (Frick et al., 2014). In a review of CU traits, Frick and White (2008) found that elevated CU traits seem to be stable across childhood and adolescence, and represent a distinct subgroup of children with DBDs. Similarly, high levels of CU are associated with treatment resistance to first-line psychosocial treatments (Haas et al., 2011). For example, Hawes, Price, and Dadds (2014) reviewed 16 studies and found an association between CU traits and risk for poor treatment outcomes. However, treatment resistance is not uniform with evidence suggesting that CU behaviours may be malleable in early childhood and can be influenced by early parental sensitivity, among other environmental interactions (Waller et al., 2014). For example, Waller and colleagues (2014) used data from a cohort of high-risk toddlers and their mothers, and found that early parental warmth emerged as uniquely important in the protection against later development of CU traits rather than in the development of behaviour problems more generally (Waller et al., 2014). In addition, in a recent review, toddlers and preschool age children’s level of disregard for others emerged as a strong predictor for severe conduct problems later on, and the emergence of empathic concern predicted a reduction in risk for severe conduct problems (Wakschlag et al., 2018).

Given the growing body of research demonstrating the influence of CU traits on
the severity and type of DBD, there has been an increased interest in uncovering the neuropsychological and neurobiological correlates of DBDs in children with elevated CU traits specifically. Indeed, numerous studies have shown not only increased severity of behavioural symptoms, but poorer emotion recognition and abnormalities in fronto-amygdalar circuitry in children with DBD (Wakschlag et al., 2018). Taken together, this clinical utility, potential for change, and response to early intervention suggests that CU traits reflect an important phenotypic indicator and potential treatment target in individuals with DBD (Haas et al., 2011).

1.5 Study Objectives

The overall aim of the present study was to further distill heterogeneity and refine subgroup classification of DBD using dimensions of EF and psychopathology, and to determine core behavioural impairments within and across the identified classes. Latent Class Analysis (LCA) was used to identify classes (i.e., subgroups) of children with DBD based on parent-rated indicators of psychopathology, i.e., CU traits and emotion dysregulation, and objective neuropsychological (EF) task performance, i.e., emotion recognition, selective attention, and response inhibition. Resulting classes were then compared on parent- and teacher-rated levels of children’s conduct problems, peer problems, total emotional and behavioral difficulties, and level of functional impairment.

1.5.1 Hypotheses

In keeping with the data-driven nature of LCA, we did not identify specific hypotheses as to the precise nature of the resulting classes. However, based on previous research there were a number of indicators and outcome variables (i.e., comparators) that were expected to cluster together within and across identified classes:

1. It was expected that there would be one class that had relatively lower levels
of difficulties across all class indicator variables, i.e., relatively better EF task performance and relatively less CU traits and emotion dysregulation. This follows previous studies using similar methodology where one class emerged with lower levels of difficulties in all domains (Aitken et al. 2017; Andrade & Wade 2016). However, as with all hypotheses and findings, these are all relative levels as compared to other clinic-referred children with DBD.

2. It was expected that classes with children with elevated CU traits would have greater levels of conduct problems and functional impairment. This follows the literature on CU traits, where findings consistently state that CU traits are predictive of greater severity and frequency of aggressive and antisocial behaviour in children with DBD (Frick et al. 2014; Haas et al. 2011).

3. It was expected that classes that contain relatively lower performance on objective EF tasks, specifically response inhibition and selective attention, would also show relatively more emotional dysregulation. This follows previous research that has linked emotion dysregulation and irritability with impairments in response inhibition and cognitive flexibility (Bolhuis et al. 2017; Wakschlag et al. 2018).
Chapter 2

Method

2.1 Participants and Procedure

Participants were 143 children (aged 6 to 12 years) referred to a clinic for children with disruptive behaviour at an urban mental health hospital. This included cross-sectional, pre-treatment data from 40 participants entering a clinical trial comparing group and individualized treatments for children with DBD and 103 participants who were participating in a clinical-research assessment to determine fit for either individual parent-child or group-based psychosocial treatments. Assessments were conducted by trained undergraduate, graduate level students, and postdoctoral fellows in psychology. Following brief introductions, written and verbal consent and assent were obtained from primary caregivers and child participants, respectively, after being given a complete description of the study and having sufficient time to ask questions and clarify responsibilities. Caregiver(s) remained in the room with a graduate student or fellow and completed a semi-structured interview and a package of questionnaires. Children completed all testing in a separate room, usually next door, with two undergraduate students trained in the assessment procedures. The entire session lasted between two and two-and-a-half hours. See below for descriptions of the subset of measures and tasks, from the larger studies, that
were used in the present study. The hospital research ethics board approved both research studies. Participants received an honorarium; parents received $20 and children received a small toy or game of their choosing. Exclusion criteria included an estimated IQ < 70 on the Kaufman Brief Intelligence Test-2nd Edition (KBIT-2) \cite{Kaufman2013} and diagnosis of ASD or a Pervasive Developmental Disorder (PDD) as determined by a pre-admission phone screening. Additionally, participants missing > 3, out of the 5, LCA indicator variables (i.e., emotion recognition, selective attention, response inhibition, emotion regulation, and CU traits) were excluded from the analysis because they were missing over half of the model indicators. Accordingly, three participants were excluded from analyses; two did not meet the minimum estimated IQ standard score of 70 and one participant was missing > 3 indicator variables, leaving the final sample size at 140 participants ($M$ age = 9.64, $SD = 1.54$).

### 2.1.1 Participants’ Demographic Characteristics

The demographic characteristics of the sample (N=140) are summarized in Table 2.1. Children comprised an ethnically diverse sample, with 20.8% of the sample coming from families of mixed heritage. 46.3% of the sample was on prescription medication at the time of assessment, with most medications being stimulant medication for ADHD and a small portion including non-stimulants for ADHD (i.e., Atomoxetine) and atypical antipsychotics (e.g., Risperidone). Most participants came from mid- high-SES homes, with 73.8% of primary caregivers having completed post-secondary education.
Table 2.1: Participant Demographics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screener IQ&lt;sup&gt;a&lt;/sup&gt;</td>
<td>104.2</td>
<td>15.7</td>
</tr>
<tr>
<td>Sex (male)</td>
<td>106</td>
<td>77.9</td>
</tr>
<tr>
<td>Medication (yes)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>62</td>
<td>46.3</td>
</tr>
<tr>
<td>Positive Diagnoses&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODD</td>
<td>115</td>
<td>83.9</td>
</tr>
<tr>
<td>CD</td>
<td>40</td>
<td>29.2</td>
</tr>
<tr>
<td>ADHD</td>
<td>111</td>
<td>81.0</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian (East, South)</td>
<td>2</td>
<td>1.6</td>
</tr>
<tr>
<td>Black (African, North American)</td>
<td>12</td>
<td>9.2</td>
</tr>
<tr>
<td>Black (Caribbean)</td>
<td>3</td>
<td>2.3</td>
</tr>
<tr>
<td>First Nations</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>Latin American</td>
<td>5</td>
<td>3.8</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>White (European, North American)</td>
<td>77</td>
<td>59.1</td>
</tr>
<tr>
<td>Mixed Heritage</td>
<td>27</td>
<td>20.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parent Demographic</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest Level of Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up to grade 8</td>
<td>2</td>
<td>1.6</td>
</tr>
<tr>
<td>Grade 9-12</td>
<td>16</td>
<td>12.7</td>
</tr>
<tr>
<td>Some post-secondary</td>
<td>15</td>
<td>11.9</td>
</tr>
<tr>
<td>Completed post-secondary</td>
<td>93</td>
<td>73.8</td>
</tr>
<tr>
<td>Current Marital Status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>20</td>
<td>17.9</td>
</tr>
<tr>
<td>Married</td>
<td>66</td>
<td>58.9</td>
</tr>
<tr>
<td>Common law</td>
<td>9</td>
<td>8.0</td>
</tr>
<tr>
<td>Divorced</td>
<td>9</td>
<td>8.0</td>
</tr>
<tr>
<td>Widowed</td>
<td>1</td>
<td>0.9</td>
</tr>
<tr>
<td>Separated</td>
<td>7</td>
<td>6.3</td>
</tr>
</tbody>
</table>

<sup>a</sup> KBIT-2 IQ composite (standard score).
<sup>b</sup> Most medications included stimulants prescribed for ADHD.
<sup>c</sup> 68.6 % (N = 96) of overall sample met criteria for more than 1 diagnosis (based on C-DISC).
2.2 Measures

Diagnostic Interview Schedule for Children, Computer Administered (C-DISC) (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000).

The C-DISC is a clinician-administered, computerized, and validated semi-structured interview based on DSM-IV diagnoses. As this study uses dimensional indices of psychopathology, diagnostic categories were not used as exclusion criteria or for the main statistical analyses, but rather to describe the clinical characteristics of the sample. See Table 2.1 and Figure 3.2 for sample diagnostic info.

2.2.1 Exclusion Criteria


The KBIT-2 is a brief, individually administered measure of verbal and nonverbal intelligence. Two subscales, verbal and non-verbal intelligence, and one composite scale, estimated IQ, were computed. For the present study, a KBIT-2 composite score two standard deviations below the mean (i.e., Standard Score < 70) was used as an exclusion criteria. This was set to ensure that participants were able to comprehend the instructions for the performance-based measures and to exclude those participants with potential intellectual disabilities.

2.2.2 LCA Indicators

Barkley’s Deficits in Executive Functioning Scale-Children and Adolescents (BDEFS-CA) (Barkley, 2012).

The BDEFS-CA is a parent or teacher reported instrument that measures dimensions of children’s executive functioning in daily life. In this study, the 16-item self-regulation of emotions subscale was used. Parents rated the frequency of their children’s difficulties
on a 4-point likert scale, from ‘never or rarely’ to ‘very often’. Statements include, “has trouble calming him/herself down once he/she is emotionally upset” and “has a low tolerance for frustrating situations”. This measure was treated as a factor in the LCA, using total raw scores, and therefore reliability values for this subscale are not reported here (see Analytic Plan for further details).

**Inventory of Callous-Unemotional Traits (ICU) (Essau, Sasagawa, & Frick, 2006).**

The ICU is a 24-item questionnaire that provides a comprehensive assessment of CU traits. Parents rated how well each statement describes their child on a 4-point likert scale, from ‘not at all true’ to ‘definitely true’. Items include “Expresses his/her feelings openly”, “Easily admits to being wrong”, and “Shows no remorse when her/she has done something wrong”. The ICU consists of 3 subscales: uncaring, callousness, and unemotional and a total CU score; for the present study the total raw CU score was used. The total CU scale was found to be reliable (24 items; \( \alpha = .77 \)). Chronbach’s alphas for the scale in previous studies ranged from .74 to .85 (Kimonis et al., 2008).

**The Cambridge Neuropsychological Testing Automated Battery (CANTAB) (Cambridge Cognition, 2018).**

Participants completed three subtests from the tablet-administered CANTAB, which is a widely used testing platform for neuropsychological research across the lifespan, among other research and development initiatives (Luciana & Nelson, 2002; Robbins et al., 1994). These tests have been successfully used in both healthy and clinical populations, including children with neurodevelopmental disorders and DBDs (Delmonte et al., 2012; Glenthøj et al., 2016; Lin & Gau, 2017; Luciana & Nelson, 2002; Robbins et al., 1994), and research has supported the construct validity of the CANTAB in measuring brain-behaviour relationships in adult psychiatric and neurological patient populations (Luciana & Nelson,
Tests from the CANTAB battery were selected because they require minimal reading, are engaging for children as they are computerized and have a ‘game-like’ quality (participants use a press-pad or touchscreen), and once each task has been introduced and started it requires minimal interaction between tester and participant thereby reducing distraction and potential participant anxiety (Haring, Mottus, Koch, Trei, & Maron, 2015; Kaufmann et al., 2013). The following three subtests were used in the present study:

**Stop-Signal Task (SST).** The SST measures the ability to inhibit a prepotent response. Participants were asked to respond to an arrow stimulus by selecting one of two buttons, based on the direction of the arrow displayed on the screen, and to withhold their response when they heard an auditory signal. The task takes up to 20 minutes and consists of one practice sub-block followed by five assessed blocks (containing four sub-blocks each), for a total of 21 sub-blocks; one practice and 20 assessed. Based on recommended guidelines (Cambridge Cognition, 2018), the last half of the test, i.e., last 10 sub-blocks, were used in the analyses for the present study. Stop Signal Reaction Time (SSRT), in milliseconds, was used as the metric to estimate response inhibition. Participant scores > 1.5 SD from the mean were excluded from analyses, for a total of 13 individual scores, as per consultation with the CANTAB science team (Cambridge Cognition, 2018). In both clinical and normative populations, 1.5 SD from the mean on performance based measures, such as CANTAB, is thought to represent individuals who are impaired relative to age, gender, and education matched peers (Abbott et al., 2015). Thus, the rationale for excluding these individual scores in the present sample was that scores 1.5 SD from the mean may have represented difficulties with the task over and above poor response inhibition, such as those who did not understand task instructions or were not attending to the task.

**Emotion Recognition Task (ERT).** The ERT measures emotion recognition on six basic facially expressed emotions. Each face was displayed for 200ms, and then participants were asked to choose which emotion was displayed from six potential options:
happiness, sadness, surprise, anger, disgust, and fear. The task takes 6-10 minutes and includes two blocks of 90 trials each. Total percent correct was used as the metric of emotional recognition. This is a widely used metric to approximate emotion recognition abilities (Glenthøj et al., 2016) and was chosen because the construct of interest was how well participants could recognize basic emotions in faces. Participant’s who scored at less than what would have been expected by chance (i.e., < 16.67%, or \( \frac{1}{6} \), correct responses) were excluded from the analyses, for a total of two individual scores. This was done to exclude those who either did not understand task instructions or were not attending to the task.

**Attention Switching Task (AST).** The AST measures selective attention and top-down cognitive control. Participants were shown an arrow that could appear on either side of the screen and could point in either of two possible directions. Each trial displayed a cue at the top of the screen that indicated to the participant to press the button that corresponded with either direction that the arrow was pointing or the side of screen that it appeared on. Total percent correct was used as the metric of selective attention. Error score, i.e., percent correct or incorrect, is a widely used metric of selective attention (Corbett, Constantine, Hendren, Rocke, & Ozonoff, 2009) and was chosen because it was a close approximation of the construct of interest - how well an individual could successfully attend to certain stimuli while ignoring others. Participants who scored at less than what would have been expected by chance (i.e., < 50% correct responses) were excluded from the analyses, for a total of four individual scores. This was done in order to exclude those who either did not understand task instructions or who were not attending to the task.

### 2.2.3 Outcome Variables

**Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997).**

The SDQ is a 33-item (of which 25 items were used) parent and teacher report questionnaire that examines five behavioural dimensions (i.e., with five items per behavioural dimen-
Chapter 2. Method

sion) including: prosocial skills, emotional symptoms, conduct problems, hyperactivity-inattention, and peer problems, as well as a total difficulties composite score. Caregivers are asked to rate statements on a 3-point likert scale, from ‘Not True’ to ‘Certainly True’. The SDQ has been widely used in both research and clinical practice and, based on a recent meta-analysis, both parent and teacher forms show acceptable test-retest reliability and inter-rater agreement on subscales ($r = .26 - .47$) (Stone, Otten, Engels, Vermulst, & Janssens, 2010). In class comparisons, raw subscale scores (parent- and teacher-rated) were used and found to be reliable (Chronbach’s alpha reported as: parent, teacher); conduct problems ($\alpha = .53$, $\alpha = .43$), peer problems ($\alpha = .10$, $\alpha = .25$), and total difficulties score ($\alpha = .57$, $\alpha = .57$). These were comparable, though slightly lower, to previous studies where Chronbach’s alphas ranged from .54 to .73. (Andrade & Wade, 2016).

Impairment Rating Scale (IRS) (Fabiano et al., 2006).

The IRS is a widely used parent and teacher report measure of children’s functional impairment in seven (six for teachers) domains. Parent and teacher responses were recorded on a visual analogue scale that is later transformed to six-point numeric values ranging from “no problem” to “extreme problem”. Raters are asked to identify the degree to which children’s behavioural difficulties impair children’s relationships with peers, siblings, and parents, academic progress, self-esteem, family functioning, and the overall severity of the child’s behavioural impairment. The mean parent- and teacher-report scores of six of the ratings were used in the present study. The sibling relationship item was excluded because it was not applicable to many study participants. Chronbach’s alphas for the, parent- and teacher-rated, mean IRS score in the present study were .73 and .59, respectively.
2.3 Analytic Plan

2.3.1 Data Cleaning

Data for all variables, except for the BDEFS-CA (see below for explanation), were missing completely at random, Little’s $\chi^2$ (76) = 64.85, $p = .815$. Normality assumptions were violated for most variables as per tests of normality (Kolmogorov-Smirnov $p = .20$ to $p < 0.001$; Shapiro-Wilk $p = .16$ to $p < 0.001$); therefore, Spearman’s rho was used to examine correlations. Normality assumptions are not required for LCA (L. M. Collins & Lanza, 2010). Descriptive statistics and correlations are presented in Table 2.2.
Table 2.2: Means and Correlations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Emotion Rec.</td>
</tr>
<tr>
<td>CU</td>
<td>11 - 48</td>
<td>31.04</td>
</tr>
<tr>
<td>Emotion Rec.</td>
<td>25 - 82.22</td>
<td>53.01</td>
</tr>
<tr>
<td>Attention</td>
<td>.63 - 48.13</td>
<td>16.79</td>
</tr>
<tr>
<td>Inhibition</td>
<td>130.10 - 415.45</td>
<td>260.41</td>
</tr>
<tr>
<td>TDS</td>
<td>0 - 34</td>
<td>17.68</td>
</tr>
<tr>
<td>Impairment</td>
<td>1.5 - 6</td>
<td>4.31</td>
</tr>
</tbody>
</table>

CU Callous-Unemotional, Attention Selective Attention, Inhibition Response Inhibition (ms), TDS Total Difficulties (SDQ; parent-rated), Impairment Average Impairment (IRS; parent-rated)

Note. Spearman’s rho was used to calculate correlation coefficients

a % incorrect on task

**p < .01
2.3.2 Latent Class Analysis (LCA)

A LCA was computed, using Mplus version 8 (Asparouhov & Muthén, 2015), with five indicator variables; two parent-rated indices of psychopathology: CU traits and emotion regulation, along with scores from three neuropsychological performance-based measures: response inhibition, emotion recognition, and attention switching. LCA was used to determine the relationship between the indicator variables in order to identify latent classes (i.e., subgroups) in a data-driven manner and find the most parsimonious solution (L. M. Collins & Lanza, 2010). Using Bayes Theorem, the latent class model estimates the conditional probabilities of two parameters, (1) the probability that the individual fits into one of the classes and (2) the probability of a response given that class membership. The resulting classes may then allow for a more meaningful classification of individuals within the population of study and further analysis of class outcomes.

Forty-six percent of the participants were systematically missing (i.e., not at random) 12 out of the 16 items from the emotion regulation subscale of the BDEFS-CA. This was because the studies’ protocols changed partway from using the short-form of the BDEFS-CA, which only included four subscale items, to the long-form of the BDEFS-CA, which included the entire subscale. The approach taken was to use individual items, as opposed to the total score, as indicators of a latent variable to represent the subscale. Missing values for all indicator and predictor variables were dealt with by Mplus through Maximum Likelihood Estimation, which uses all available information in the data (Muthén & Muthén, 2002). Maximum Likelihood Estimation in the presence of missing data is one of the most sophisticated ways to handle missing values, comparable with Multiple Imputation (Muthén & Muthén, 2002).

Models with 1 through 5 classes were examined. Model fit was evaluated using the Vuong-Lo Mendell Rubin Test (VLMRT), Bootstrap Likelihood Test (BLRT), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), entropy, and interpretability of the classes, including size and theoretical and clinical relevance. The
VLMRT compares improvement of model fit from a $k-1$ to a $k$ class model, and the $p$ value indicates if there is a statistically significant difference in adding one more class to the model, which then provides one appropriate stopping point in the analysis (Nylund, Asparouhov, & Muthén, 2007). The BLRT provides another significance (i.e., $p$ value) based stopping point by using bootstrap samples to empirically estimate the log likelihood distribution, as opposed to assuming that it follows some already known distribution, which is then used to determine the statistical utility of adding one more class to the model (Asparouhov & Muthén, 2012; Nylund et al., 2007).

The 3-step solution for latent class predictor variables was used to compare parent- and teacher-rated social functioning, conduct problems, and overall impairment across the classes, using the parent- and teacher-rated SDQ and IRS. This approach was selected as it accounts for the conditional probability estimates from the LCA (Asparouhov & Muthén, 2014). Effect sizes (Cohen’s $d$) were estimated as the difference between the means across pairs of classes, divided by the pooled standard deviation (Kelley & Preacher, 2016). Fisher’s exact test was used to test for possible sex differences in class membership, and Analysis of Variance (ANOVA) was used to test for possible age differences across classes.
Chapter 3

Results

3.1 Latent Class Analysis

The first objective of this study was to identify classes (i.e., subgroups) of children with DBD based on parent-rated indicators of psychopathology, i.e., CU traits and emotion dysregulation, and objective neuropsychological (EF) task performance, i.e., emotion recognition, selective attention, and response inhibition. Resulting classes were then compared on parent- and teacher-rated levels of conduct problems, peer problems, total difficulties, and level of functional impairment. Model fit indices did not unanimously suggest a best-fitting class solution (Table 3.1). Model fit improved as indicated by decreases in the AIC and sample size adjusted BIC up to the 4-class solution. VLMRT $p$-values were non-significant past the 2-class solution; however, a number of guidelines advise against using the VLMRT as a primary stopping indicator due to susceptibility toward type-1 errors, as compared to the BLRT [Nylund et al., 2007]. BLRT $p$-values remained significant up until the 3-class solution and approached significance in the 4-class solution. Taken together with the AIC and BIC values, Classes 3 and 4 were reviewed and considered based on existing theory and a priori hypotheses. The 4-class solution was selected as the best model because all 4 classes were of adequate size (i.e., n > 5%).
and it provided a robust and theoretically and clinically relevant set of classes. Figure 3.1 displays class profiles with all scores indexed to the overall sample mean (i.e., each score is the difference between the class mean and the overall mean, divided by the class standard error). This was done so that all five indicator variables could be visually represented in the same figure, despite their scales being of different magnitudes. Table 3.2 shows average posterior probabilities for the 4-class model.

Class 1 (21% of sample) was characterized by relatively less CU and emotion dysregulation and moderate EF performance; it was labeled “less psychopathology/moderate EF”. Class 2 (15% of sample) was characterized by moderate CU and emotion dysregulation, moderate selective attention, and lower performance on emotion recognition and response inhibition tasks; it was labeled “moderate psychopathology/mixed EF”. Class 3 (11% of sample) was characterized by moderate CU, moderate emotion dysregulation and response inhibition, and lower performance on the selective attention and emotion recognition tasks; it was labeled “moderate psychopathology/lower selective attention”. Class 4 (53% of sample) was characterized as having more CU, moderate emotion dysregulation, and better EF performance; it was labeled “mixed psychopathology/better EF performance”. See Fig. 3.1

Table 3.1: Model Fit Statistics for Latent Class Analysis

<table>
<thead>
<tr>
<th>Classes</th>
<th>AIC</th>
<th>BIC</th>
<th>Adj. BIC</th>
<th>VLMRT p</th>
<th>BLRT p</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6026.515</td>
<td>6191.247</td>
<td>6014.070</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>5988.893</td>
<td>6171.275</td>
<td>5975.115</td>
<td>0.0234</td>
<td>&lt; 0.001</td>
<td>0.690</td>
</tr>
<tr>
<td>3</td>
<td>5978.227</td>
<td>6178.258</td>
<td>5963.115</td>
<td>0.7417</td>
<td>&lt; 0.001</td>
<td>0.724</td>
</tr>
<tr>
<td>4</td>
<td>5971.010</td>
<td>6188.692</td>
<td>5954.565</td>
<td>0.3631</td>
<td>0.0667</td>
<td>0.706</td>
</tr>
<tr>
<td>5</td>
<td>5973.141</td>
<td>6208.472</td>
<td>5955.363</td>
<td>0.3117</td>
<td>0.8333</td>
<td>0.748</td>
</tr>
</tbody>
</table>

AIC Akaike information criterion, BIC Bayesian information criterion, BLRT bootstrap likelihood ratio test, VLMRT Vuong-Lo-Mendell-Rubin test
Table 3.2: Average Posterior Probabilities by Latent Class

<table>
<thead>
<tr>
<th>Most Likely Latent Class</th>
<th>Latent Class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class 1</td>
</tr>
<tr>
<td>Class 1</td>
<td>0.759</td>
</tr>
<tr>
<td>Class 2</td>
<td>0.032</td>
</tr>
<tr>
<td>Class 3</td>
<td>0.042</td>
</tr>
<tr>
<td>Class 4</td>
<td>0.066</td>
</tr>
</tbody>
</table>

Figure 3.1: Indexed Means within Latent Classes

*Note.* Class 1: less psychopathology/moderate EF.
Class 2: moderate psychopathology/mixed EF.
Class 3: moderate psychopathology/lower selective attention.
Class 4: mixed psychopathology/better EF performance.
Table 3.3: Percentage of Scores in the Abnormally Elevated Range on SDQ by Class

<table>
<thead>
<tr>
<th>Subscale</th>
<th>% of Class (Parent-Rated)</th>
<th></th>
<th></th>
<th></th>
<th>% of Class (Teacher-Rated)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Emotional Problems</td>
<td>33.3</td>
<td>46.2</td>
<td>42.9</td>
<td>30.6</td>
<td></td>
<td>37.5</td>
<td>38.5</td>
<td>27.3</td>
</tr>
<tr>
<td>Conduct Problems</td>
<td>30.0</td>
<td>50.0</td>
<td>42.9</td>
<td>48.6</td>
<td></td>
<td>75.0</td>
<td>46.2</td>
<td>54.5</td>
</tr>
<tr>
<td>Hyperactivity</td>
<td>63.3</td>
<td>80.0</td>
<td>71.4</td>
<td>68.1</td>
<td></td>
<td>75.0</td>
<td>53.8</td>
<td>54.5</td>
</tr>
<tr>
<td>Peer Problems</td>
<td>30.0</td>
<td>55.0</td>
<td>50.0</td>
<td>45.8</td>
<td></td>
<td>62.5</td>
<td>38.5</td>
<td>63.6</td>
</tr>
<tr>
<td>Prosocial Difficulties</td>
<td>6.7</td>
<td>15.0</td>
<td>14.3</td>
<td>23.8</td>
<td></td>
<td>50.0</td>
<td>53.8</td>
<td>54.5</td>
</tr>
<tr>
<td>Total Difficulties</td>
<td>50.0</td>
<td>55.0</td>
<td>64.3</td>
<td>56.9</td>
<td></td>
<td>56.3</td>
<td>46.2</td>
<td>45.5</td>
</tr>
</tbody>
</table>

3.1.1 Descriptive Statistics by Class

Table 3.3 shows the percentage of children falling in the abnormal range on parent- and teacher-rated SDQ subscales by class. Figure 3.2 shows the proportion of children in each class in each diagnostic category. As can be seen, the percentage of the sample with abnormally elevated scores on the SDQ varied greatly depending on subscale and informant across the classes, ranging from anywhere between 6.7% to 80% of the sample on the parent rated form, and between 27.3% and 75.0% on the teacher rated form. Diagnostic categorization also varied greatly by class, with the proportion of ‘ADHD-only’ participants ranging from 0% to 23.33% of a class.
Chapter 3. Results

Note. Comorbid refers to > 1 diagnosis on C-DISC (i.e., 2 or 3 of: ADHD, ODD, and CD). 68.6% (N = 96) of overall sample fell into the comorbid category. None refers to children who had no positive diagnoses on C-DISC.

3.2 Problems and Impairment Across Classes

Class comparisons using the 3-step approach revealed a number of significant differences between classes on conduct problems, total difficulties, and overall impairment. Based on post-hoc comparisons, children in the mixed psychopathology/better EF performance
class (Class 4) had significantly higher parent-rated conduct problems than the less psychopathology/moderate EF class (Class 1), $\chi^2(1) = 6.356, p = 0.012$. Children in the moderate psychopathology/lower selective attention class (Class 3) had significantly higher parent-rated total difficulties than the less psychopathology/moderate EF class (Class 1), $\chi^2(1) = 3.717, p = 0.054$. Finally, children in the moderate psychopathology/lower selective attention class (Class 3), $\chi^2(1) = 6.738, p = 0.009$, and the mixed psychopathology/better EF performance class (Class 4), $\chi^2(1) = 4.650, p = 0.031$, had significantly greater parent-rated overall impairment than the less psychopathology/moderate EF class (Class 1). Parent rated peer-problems and all teacher-rated measures did not reach statistical significance for between class differences.

### 3.3 Age and Sex Differences Across Classes

See Table 3.4 for effect sizes (Cohen’s $d$) for class differences on parent-rated total difficulties (SDQ, total difficulties subscale) and average impairment (IRS), and age. There was no overall significant sex difference between the four classes, Fisher’s Exact Test, $p = .788$. However, there were significant age differences between classes, $F(3, 134) = 6.95, p < .001$. Post hoc comparisons using the Bonferroni correction revealed significant age differences between the less psychopathology/moderate EF class (Class 1; $M = 9.57, SD = 1.78$) and the moderate psychopathology/lower selective attention class (Class 3; $M = 8.34, SD = 1.47$); and the moderate psychopathology/lower selective attention class (Class 3; $M = 8.34, SD = 1.47$) and mixed psychopathology/better EF performance class (Class 4; $M = 10.07, SD = 1.29$). No other age differences reached statistical significance.
Table 3.4: Between Profile Comparisons on Total Difficulties, Impairment, and Age

<table>
<thead>
<tr>
<th>Class</th>
<th>Effect Size (Cohen’s $d$)</th>
<th>Mean Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class 1</td>
<td>Class 2</td>
</tr>
<tr>
<td>Class 1</td>
<td>–</td>
<td>0.42</td>
</tr>
<tr>
<td>Class 2</td>
<td>0.40</td>
<td>–</td>
</tr>
<tr>
<td>Class 3</td>
<td>0.72**</td>
<td>0.41</td>
</tr>
<tr>
<td>Class 4</td>
<td>0.55*</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Effect sizes above the diagonal are for total difficulties (SDQ; parent-rated); below the diagonal for overall impairment (IRS; parent-rated).

*p < .05; **p < .01; ***p < .001
Chapter 4

Discussion

4.1 Main Findings

The present study aimed to further distill heterogeneity and refine subgroup classification of DBDs, using dimensions of neuropsychology and psychopathology, and to examine differences in functional difficulties and impairments between the identified classes. This aim was met by using a LCA model that included two parent-rated indicators of psychopathology, i.e., CU traits and emotion dysregulation, and three objective neuropsychological (EF) task performance measures, i.e., emotion recognition, selective attention, and response inhibition. Resulting classes were then compared on parent- and teacher-rated levels of conduct problems, peer problems, total difficulties (as measured by the SDQ), and level of functional impairment (as measured by the IRS). Four classes of children with DBD were identified. Over half of the sample, the mixed psychopathology/better EF performance class (Class 4), presented with relatively moderate levels of emotion dysregulation, better performance on EF tasks, and elevated CU traits. Just over one-fifth of the sample, the less psychopathology/moderate EF class (Class 1), was comprised of children with relatively low scores on indices of psychopathology, and moderate EF performance scores (i.e., around the sample mean). The smallest of the identified classes, at 15% (the moderate
psychopathology/mixed EF class; Class 2) and 11% (the moderate psychopathology/lower selective attention class; Class 3) of the sample respectively, had relatively low-moderate ratings of psychopathology and EF performance scores that ranged between moderate and poor relative to the rest of the sample. Class differences emerged on measures of parent-rated conduct problems, total difficulties, and overall impairment, though few reached statistical significance. In line with the initial hypotheses, the class with the most elevated CU traits (Class 4) had the highest parent-rated conduct problems, and children with the poorest performance on the selective attention task (Class 3) had the highest parent-rated total difficulties. Additionally, children in both of these classes (Classes 3 & 4) had the greatest parent-rated overall impairment.

4.1.1 Relative Difficulties and Impairment

It was hypothesized that one class would be identified with lower levels of difficulties and impairment on all indicator variables in the LCA, i.e., relatively better EF task performance and relatively less CU traits and emotion dysregulation. This followed results from studies using similar methodologies, in which one class emerged with lower levels of difficulties on all domains (Aitken et al., 2017; Andrade & Wade, 2016). Consistent with this hypothesis, the less psychopathology/moderate EF class (Class 1) had the lowest levels of difficulties on all class indicators. Moreover, when compared to other classes, children in this class had fewer behavioural difficulties and less impairment than those in the other classes.

At first glance, this class (Class 1) does not seem to fit the profile of difficulties that may be expected in a child with DBD. However, one must be cautious in interpreting this relative lack of impairment as all participants presented to a specialized behaviour clinic, and therefore class membership reflects relative functioning and not the absence of difficulties and impairments given that all model indicators are compared to the overall sample and not a normative sample. As such, all descriptive comparisons made between
the classes are relative to other children with DBD (i.e., the study sample). Additionally, one of the primary tenets of LCA interpretation is that the individual variables within the class (i.e., model indicators or symptoms) are conditionally independent of one another. This means that the latent class(es) as a whole, and not any one indicator alone, is assumed responsible for creating the emergent associations between the model indicators (L. M. Collins & Lanza, 2010). Therefore, the comparisons between classes can be conceptualized as comparing between independent groups of children with DBD, i.e., distilling heterogeneity.

Though it is possible that Class 1 represents a relatively less impaired subset of children who present at a clinic for DBD, it is also possible that this is a subgroup of children whose core difficulties (i.e., endophenotypes) were not sufficiently captured by the model indicators included in the present study. For example, a large proportion of the sample met criteria for ADHD, and as such, other objective neuropsychological indicators such as working memory, spatial planning, or problem-solving abilities may have enhanced the precision of profiles as a “true” representation of core difficulties.

4.1.2 CU and Emotion Recognition

It was hypothesized that classes characterized by elevated CU traits would have greater levels of conduct problems and functional impairment. This followed the reliable findings that elevated CU traits in children with DBD are predictive of greater severity and frequency of aggressive and antisocial behaviour (Frick et al., 2014; Haas et al., 2011). Consistent with this hypothesis, when compared to other subgroups, the class of children with the highest levels of CU traits (Class 4), relative to the other 3 classes, also had the highest levels of parent-rated conduct problems and overall impairment. This class (Class 4) was the largest of the four classes, at over half of the overall sample, and was characterized by elevated CU traits and relatively better performance on EF tasks. The emergence of this class is consistent with current literature describing the high prevalence
of CU traits in clinical samples of children with DBD (Frick et al., 2014). In a review, Frick and colleagues (2014) found that children and adolescents with DBD who have elevated CU traits comprise a unique subset of that population who often do not respond positively to first-line psychosocial interventions. They concluded, however, that tailoring such interventions to specific emotional, cognitive, and motivational characteristics of this subgroup may provide not only effective reduction in aggressive and antisocial behaviours, but in CU traits as well (Frick et al., 2014). Moreover, recent findings from members of our research group suggest that there are certain classes within DBD that have elevated CU traits as a core feature, and these emergent classes show the most problematic social difficulties (Aitken et al., 2017).

Unexpectedly, the class with the relatively highest levels of CU (Class 4) had relatively better performance on all EF measures as compared to other classes. Existing studies have been mixed with regards to EF abilities in children and youth with elevated CU traits (Sulik & Obradović, 2017; Wakschlag et al., 2018; Waller, Hyde, Baskin-Sommers, & Olson, 2017). Waller and colleagues (2017) followed a cohort of high-risk children, collecting cross-informant and observational data, and found that CU traits at age 3 predicated aggression and peer rejection at age 10. Moreover, these effects were amplified in those with low levels of EF performance ratings, but at high levels of EF performance ratings, this effect was non-significant. Results from this study suggest that better EFs may serve as a protective factor for aggression and antisocial behaviour in children with elevated CU traits (Waller et al., 2017). However, other studies have found that elevated CU traits and deficits in EF are linked to more severe antisocial behaviour, including later violence and substance-abuse (Baskin-Sommers, Waller, Fish, & Hyde, 2015). Findings from the present study are consistent with studies suggesting that EF abilities are not impaired in all children with DBD who have elevated CU, and that EF abilities in these children are potentially a protective factor; however, these assertions are in need of much further investigation given that there was no typically-developing control group and that among
the four classes only one had markedly elevated CU traits, thus offering no other class for comparison.

Furthermore, and again inconsistent with expected clustering of class indicators, this class (Class 4) performed better than each of the other three groups as assessed by percentage of correct responses overall on the emotion recognition task, one of the specific EF domains measured. This finding is inconsistent with research findings suggesting that children with DBD who have elevated CU traits show poor emotional recognition of faces (Sully et al., 2015; Wakschlag et al., 2018). We posit that this unanticipated finding may have emerged for a few reasons. First, a person-centred analysis (i.e., LCA) was used in this study, as opposed to variable-centred approaches where the groups being compared are predefined. As such, all indicators should be considered in relation to the other identified classes, and as such do not indicate normatively ‘good’ or ‘poor’ performance. Additionally, there was no control group for which to compare performance to a typically-developing sample.

Next, the type of emotion recognition task that was used did not elicit emotional responses in participants while engaged in the task. Recent developmental research has started to conceptualize EF into dimensions that are related both in their neurobiology and clinical utility. One such distinction is between “cool”-cognitive (e.g., attention and response inhibition) EF and “hot”-motivational (e.g., affectively charged) EFs (Matthys et al., 2013). Though not a strict dichotomy, separating these two dimensions of EF this is a useful and important distinction to make, as many of the difficulties faced by children with DBDs involve regulation in emotionally charged situations such as temper tantrums and fights with other children (Woltering et al., 2016). Even though some studies have cited impaired emotion recognition on “cooler”-EF tasks, such as those similar to the facial recognition task used in the present study (Sully et al., 2015), children with DBD seem to show the greatest impairment on “hot”-EF emotion recognition tasks that elicit an emotional reaction during the task itself (Woltering et al., 2016), which is in line with
their difficulties in real-world, emotionally-charged situations. This follows the findings that traditional performance-based EF measures are somewhat disconnected with ratings of behaviour and that they lack generalizability to real-world outcomes \cite{Toplak2013}. Moreover, in the present study CU traits were measured by parent report and emotion recognition by objective task performance and this may have further contributed to the disconnect in expected indicator clustering, as the task-based measures may have been measuring an efficiency index rather than real-world, behavioural outcomes \cite{Toplak2013}. In keeping with this logic, it may then be reasonable to hypothesize that given consistent findings on the severe and persistent conduct and antisocial outcomes seen in children with elevated CU traits \cite{Frick2014}, along with emerging findings on the connection between CU traits and impaired emotion recognition \cite{Sully2015,Woltering2016}, CU and emotion recognition may cluster together when emotion recognition is measured using a “hot”-EF measure. This type of task may then more closely approximate the “true” difficulties expected in this class of children with DBD; i.e., elevated CU traits, impaired emotion recognition, and poorer caregiver rated problems and impairment. However, this latter assertion is in need of further investigation.

### 4.1.3 Dimensions of Psychopathology

It was also hypothesized that relatively lower performance on objective EF tasks, specifically response inhibition and selective attention, would cluster together with relatively more emotion dysregulation. This followed the literature linking emotion dysregulation and irritability in children with DBDs, with impairments in response inhibition and cognitive flexibility \cite{Bolhuis2017,Wakschlag2018}. Consistent with hypotheses, both the moderate psychopathology/mixed EF class (Class 2) and the moderate psychopathology/lower selective attention class (Class 3) showed relatively high levels of emotion dysregulation, and mixed performance on the EF tasks; with Class 2 defined by relatively poor response inhibition and Class 3 defined by relatively poor selective
This finding is consistent with existing literature linking emotion dysregulation and poor inhibition (and impulsivity) in children with DBD, both behaviourally and neurobiologically (Brotman et al., 2017; Wakschlag et al., 2018). In a recent review, Wakschlag and colleagues discuss a model of emotion dysregulation in children with DBD including related neurodevelopmental abnormalities. In this model, emotion dysregulation, specifically irritability, presents behaviourally in developmentally inappropriate contexts and frequencies and is related to impaired mental flexibility and response inhibition (Wakschlag et al., 2018). Brotman and colleagues present a similar model of irritability, though highlight the construct of frustrative nonreward, and its exaggerated manifestation in children with DBD in response to blocked goals (Brotman et al., 2017). Both of these complementary models underscore possible shared mechanisms between EF and emotion dysregulation; most notably influencing impaired abilities in children with DBD for considering alternative solutions, allocating attention to relevant parts of a problem, and inhibiting the most salient response option. Taken together, these difficulties may present behaviourally as a temper tantrum or a long-lasting irritable mood (Wakschlag et al., 2018). However, even though emotional dysregulation has been shown to be pervasive in children with DBDs (Cavanagh et al., 2017; Nigg, 2017), including in the present study, the specific impairments that encompass dysregulated emotion, such as irritability and emotional lability, have been studied less frequently. This is important given that irritability, compared to other dimensions of emotion dysregulation, has been shown to be a unique predictor of future mood and anxiety disorders (Stringaris & Goodman, 2009).

In line with the robust evidence base supporting early intervention for children with DBDs (Luby, 2012; Sonuga-Barke & Halperin, 2010; Waller et al., 2014), if severe irritability represents a prodromal dimension of later mood and anxiety problems, this subset of children with DBDs may benefit from early, targeted prevention measures.

Another important differentiation between Classes 2 and 3 lies in the value added by
adopting a dimensional approach, versus a categorical one, to understanding childhood psychopathology, and specifically DBDs. As previously mentioned, there has been a call from the NIMH for researchers to study mental illness using a dimensional, as opposed to categorical, lens and to integrate multilevel data that cuts across many existing diagnostic categories (NIMH, 2015). The distinction between Classes 2 and 3 provides an example of the potential utility in adopting these transdiagnostic models in research, and eventually clinical practice. In the same way that traits and behaviours, including those expressed in DBDs, occur normatively across the population and are only considered disorders when they reach the extreme “tails” of the distribution, dimensions of dysfunction vary across DBDs as well (Copeland et al., 2015). In the case of Classes 2 and 3, emotion dysregulation was relatively moderate, and EF performance mixed, in describing the two classes. However, upon closer examination, the levels of emotion dysregulation and relative EF task performance varied across the subgroups, providing a more nuanced understanding of the profiles of children within the identified classes. Furthermore, the moderate psychopathology/lower selective attention class (Class 3) had significantly higher parent-rated total difficulties and overall impairment than the less psychopathology/moderate EF class (Class 1). The moderate psychopathology/mixed EF class (Class 2) did not differ significantly from any other classes in levels of difficulties or functional impairment.

Additionally, when examining and interpreting the results of a data-driven analysis, such as a LCA, it is also important to consider classes not identified in the analysis. Often referred to as the antitype, or as “white spots” in the data, these refer to unidentified configurations of class indicators within the population, or sample, or study (Bergman & Magnusson, 1997). In the present study, no class emerged that had relatively good response inhibition as well as elevated emotion dysregulation. This hypothetical class would follow the literature on EF in children with DBD and comorbid internalizing symptoms (e.g., anxiety), where findings suggest that these children may perform better
on response inhibition tasks. These children may present as more behaviourally inhibited, as a result of their elevated anxiety (Woltering et al., 2016), but also present as highly dysregulated and irritable (Stringaris & Goodman, 2009). These children also tend to have a high degree of social inhibition and positive emotionality (Steinberg & Drabick, 2015) and thus may appear fine at school and with peers. It is possible that children who may fit in this hypothetical class did not present to the clinic as they are not traditionally “disruptive” and do not seem overtly impaired in some settings. Additionally, there is a child and youth mood and anxiety clinic within the same mental health service that the present sample was recruited from. It is possible that upon intake to the hospital, and initial consultation with psychiatry, some of the children belonging to this hypothetical “class” were referred to the mood and anxiety service as opposed to the service for children with disruptive behaviours. This underscores the importance of integrative and collaborative models of care, especially considering the high degree of comorbidity in this population.

4.1.4 Measurement

In comparing social and behavioural difficulties and impairment across classes, it is important to consider that many indices, such as questionnaires and task-based measures, don’t capture the entirety and nuance of difficulties that parents and teachers report. Children’s indicators of psychopathology and EF abilities can be measured in a number of different ways, including caregiver rating scales and objective task performance. Constructs measured by each of these sources of information are often intended to measure similar, or the same, mental constructs; however, studies comparing the two sources have demonstrated that weak correlations between the two measurement sources, and that they may actually be measuring two different, albeit complementary, constructs. It has been suggested that the performance-based measures provide an index of processing, whereas the rating scales more closely indicate goal pursuit in ‘real-world’ situations (Toplak et al., 2013). The
present study combined both performance-based and rating scales, tapping into both sets of constructs and accounting for individual variability in both complementary domains, thus providing deeper and more robust profiles of children with DBD.

4.2 Limitations

There were several study limitations that must be noted. First, there was a large amount of missing teacher data on our measures of functional difficulties and impairments (i.e., the SDQ and IRS). Although this data was determined to be missing completely at random, this missing data may have impacted why these teacher-rated measures, in contrast to the same parent-rated measures, did not emerge as significant in between class comparisons. Another source of missing data was on the BDEFS-CA, the measure of emotion regulation. As a result, there was not as much variability in scores as there may have been if all parents had completed the entire 16-item subscale of emotion regulation. However, missing data were dealt with using Maximum Likelihood Estimation in Mplus, which uses all available information, and was therefore the most robust way to include this measure given the circumstances [Muthén & Muthén, 2002].

Second, the task-based measure of emotion recognition used was not a “hot”-EF task, i.e., an affectively charged task. Children with DBDs seem to show the greatest impairments in emotion recognition when the task is a “hot”-EF task, where the task itself elicits an emotional reaction [Woltering et al., 2016], which may represent a more realistic measure of recognizing emotions in the faces of others during affectively charged situations. Another indicator-selection issue in the present study was that the measures of emotion regulation and CU traits were both parent-report measures, and not objective task-based measures that may more closely approximate the “true” construct, or endophenotype. Some examples of such tasks may have included “hot”-EF tasks combined with physiological recordings such as heart-rate variability and skin
conductance (Karalunas et al., 2014; Woltering et al., 2016), which may have also provided a more nuanced indication of emotional responsivity.

Thirdly, between class comparisons used global scales (i.e., SDQ and IRS) of difficulties and impairment. This was consistent with the overall aim of the study; however, using more domain specific and nuanced outcome measures to compare classes may have provided a more refined picture of areas of concern unique to emergent classes. Further, using multidimensional outcome measures, collected in different settings, may also provide a more holistic picture of specific areas of concern.

Lastly, the clinical sample studied limits generalizability because no typically developing control group was included on which to compare the classes. The sample age range was 6-12 years, which also limits generalizability to school-aged children, and does not allow inferences to be made for preschool aged children or adolescents with DBDs.

### 4.3 Future Directions and Clinical Implications

The results of the present study underscore the importance of investigating both neuropsychological and psychopathological endophenotypes of children with DBDs together to create more refined classes that distill heterogeneity. This study also highlights the importance of accurate identification of problems and impairments within subsets of this population, given the apparent differences that were found between the classes identified in this study. However, it should be noted that there may have been other latent factors that influenced the trait variability, and in turn the class differentiation, including parent characteristics and cognitions. Future projects might incorporate data on parent characteristics to further parse apart these relationships.

Research on evidence-based psychosocial interventions for DBDs consistently show only medium treatment effect sizes, suggesting that a concerning number of children fail to respond to these widely used treatments (Burke et al., 2002; Kazdin & Wassell...
As outlined earlier, the current evidence posits heterogeneity in neuropsychological abilities, neurobiological vulnerabilities, dimensions of psychopathology, and behavioural presentation as important factors limiting treatment effectiveness. In addition to pharmacological intervention for children diagnosed with ADHD, including those with comorbid DBDs, first-line interventions include cognitive and behavioural parent and child interventions that center around learning and improving social skills, coping mechanisms, and parent management of children’s behaviours (Eyberg et al., 2016). Additionally, many of these types of interventions target EF related life-skills such as impulsivity, emotion regulation, and perspective taking (Epstein, Fonnesbeck, Potter, Rizzone, & McPheeters, 2015; Sukhodolsky et al., 2016), including the Incredible Years Parenting Program (Webster-Stratton, 2005) and the Coping Power Program (Lochman, Wells, & Lisa, 2008). However, aside from clinical formulation, which may vary based on clinician orientation and training, there is a paucity of personalized approaches in selecting evidence-based interventions for children with DBD - and in childhood psychopathology more generally (Ng and Weisz, 2016). This often results in a trial-and-error approach to treatment selection which places undue burden on both families and the health care system. For example, in the present study Class 4, comprising over half of the sample, had markedly elevated CU traits but better EF relative to the other 3 classes. As such, children in Class 4 may benefit more from a tailored intervention centered on positive parent-child interactions and empathy toward others (Waller et al., 2014) than they would from one focused on self-regulation of emotions.

This precision-medicine approach would apply not only to the tailoring of existing interventions, but to treatment innovation as well. Presently, there exists a lack of effective treatments that target irritability (Leibenluft, 2017), even though it is a common clinical problem in children with DBDs (including some of the classes in the present study) and predicative of future mood and anxiety disorders (Stringaris & Goodman, 2009). Combined with findings on the neuropsychological impairments that have been related...
to irritability in children with DBDs and other mental disorders (Nigg, 2017; Wakschlag et al., 2018), findings from the present study may pave the way for the development of precisely targeted interventions for this population of children. As in assessment, improving treatment response prediction and adopting a precision medicine approach to treatment recommendations will rely on further refining diagnostic categories that blur boundaries and include neurobiological and psychosocial dimensions of dysfunction (Karalunas et al., 2014; Williams, 2016). Adopting this approach may also facilitate biomarker discovery, including novel treatment targets and further treatment innovation for children with DBD and their families.
Bibliography


BIBLIOGRAPHY


