What is a Good Outcome of Treatment for Adolescent Depression? A Mixed-Methods Exploration of Measurement, Concepts, and Priorities

Rose Karolin Krause

University College London

Thesis submitted for the degree of Doctor of Philosophy (PhD)

March 2020

Declaration

I, Rose Karolin Krause, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated.

Signature:

12/03/2020

Abstract

Background and aim: Evidence-based care places outcome measurement at the core of mental health practice and research. But there is no consensus on what constitutes a 'good outcome'. This thesis aimed to advance the debate specifically for adolescent depression, through a mixed-methods exploration of outcome concepts, priorities, and measurement.

Methodology: (1) A narrative review of the outcomes literature in mental health identified an initial taxonomy of outcomes. (2) A systematic review considered outcomes reported in recent treatment effectiveness studies. (3) Qualitative content analysis explored outcome perspectives amongst youth, parents, and clinicians following therapy. (4) A Q-methodological study assessed the relative importance given to different outcomes by youth and practitioners. (5) Quantitative analysis examined the convergence of reliable change ratings across selected outcome domains and measures in a naturalistic dataset.

Results: (1) based on all five studies an outcome taxonomy for adolescent depression was developed, consisting of 32 outcomes across seven domains. (2) Treatment effectiveness studies primarily reported on change in symptoms (94%), followed by global functioning (55%). (3) Symptom change was a key focus for youth, parents, and clinicians post treatment, but closely followed by coping, family functioning, and academic functioning. (4) Four distinctive viewpoints emerged on what outcomes matter the most: *Symptoms – feeling better*; *Self-management – resilience through coping skills; Parental support – resilience through family support*; and *Functioning – less interference with daily life*. (5) Symptom change appears to be an imperfect proxy for change in functioning and progress towards personal goals.

Conclusion: Change in depressive symptoms emerged as a core outcome. However, youth, parents, and clinicians endorsed additional outcomes, conveying a multidimensional picture that is inadequately captured by measuring symptoms alone. To promote outcome assessment that is streamlined, person-centred, and can illuminate treatment mechanisms, the consideration of additional outcomes, beyond symptom change, may be beneficial.

Impact Statement

Depression is a common mental health problem in adolescence, with a prevalence of around 6% in the United Kingdom (UK). Globally, depression is one of the most common causes of disability in adolescence, and associated with considerable adversity over the life course. In light of the high burden, a commission of health experts has stressed the urgent need to deliver high-quality treatments for depression. In this context, understanding what treatments deliver the best possible outcomes is imperative. However, there is currently no consensus on what constitutes a ‘good outcome’ for youth seeking treatment for depression.

Two initiatives are under way to address this lack of consensus for youth depression. They aim to identify core sets of outcomes that will be recommended for routine use in clinical practice or research, in an attempt to streamline measurement. One such set is being devised for use in clinical practice by an international working group under the lead of the International Consortium of Health Outcomes Research (ICHOM), while another is under development at the Hospital for Sick Children in Toronto, Canada for use in clinical trials. Both initiatives are faced with the challenge of scarce conceptual guidance and evidence on the topic.

This PhD thesis contributes to filling this gap. It has produced a taxonomy that provides a conceptual framework for appraising different treatment outcomes. It has further shown that reduced depressive symptoms are a core outcome in the eyes of young people, parents, and clinicians – a marker of success that all three groups can agree on. Symptom change is also by far the most frequently measured outcome in treatment effectiveness research, which has rarely covered other outcomes (with the exception of functional impairment). In addition, youth, parents, and clinicians, value a range of other outcomes, such as improved *coping and resilience*, *family functioning*, *therapeutic alliance*, *academic and vocational functioning*, *social functioning*, *friendships*, and *self-confidence*. There is considerable diversity in how youth and clinicians prioritise these additional outcomes. Lastly, the thesis provides evidence that symptom-change is an inadequate proxy for change in functioning or progress towards personal treatment goals. Approaches to measurement that focus solely on symptom change are likely to yield an incomplete picture of the changes achieved, and may need to be enhanced through measurement of additional outcomes.

Over the course of this thesis, the doctoral candidate has become associated with both of the above-mentioned core outcome set initiatives, and the finding from this doctoral thesis have impacted on their outputs. As a research fellow, the candidate has been closely involved in the development of the ICHOM set since its inception in October 2018, and was able to transfer conceptual and methodological learning, such as tools and processes for the systematic extraction of outcomes from recent treatment effectiveness studies, and for the conceptual mapping of the extracted concepts. The ICHOM core set has sparked strong interest amongst clinicians worldwide, with an external consultation survey yielding more than

450 responses from health professionals in 45 countries (the largest response that ICHOM have ever received to a consultation). Collaboration with the developers of the Canadian core outcome set has involved presenting emerging findings from this thesis at a joint symposium on core outcome sets at the Anxiety and Depression Association of America's annual conference in March 2019, and a lunchtime presentation at the Cundill Centre for Child and Youth Depression in Toronto in October 2019. Given the growing interest in outcome measurement, it can be expected that further initiatives will emerge in the future. They will be able to draw on the conceptual framework and empirical findings provided by this thesis, to strengthen outcome measurement in clinical research and practice.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ADHD	Attention deficit hyperactivity disorder
AFNCCF	Anna Freud National Centre for Children and Families
APA	American Psychiatric Association
ASD	Autism spectrum disorder
BPI	Brief Psychosocial Intervention
CAMHS	Child and adolescent mental health services
CBT	Cognitive behavioural therapy
CGAS	Children’s Global Assessment Scale
CGI	Clinical Global Impression scale
CHIP	Children's Health Insurance Program
CL	Clinicians
CMHI	Child Mental Health Initiative
CORC	Child Outcomes Research Consortium
CORS	Child Outcome Rating Scale
COS	Core Outcome Set
CYP IAPT	Children and Young People's Improving Access to Psychological Therapies programme
DD	Dysthymic disorder
GBO	Goal Based Outcomes
HEDIS	Healthcare Effectiveness Data and Information Set
IMPACT	Improving Mood with Psychoanalytic and Cognitive Therapies
MDD	Major depressive disorder
MI	Meaningful improvement
NCINQ	National Collaborative for Innovation in Quality Measurement
NCQA	National Committee for Quality Assurance
NHS	National Health Service
NIMH	National Institute of Mental Health
OMS	Outcome Measurement Review
ORS	Outcome Rating Scale
PA	Parent
PTOP	Post-Therapy Outcome Perspectives
RC	Reliable change
RCADS	Revised Children’s Anxiety and Depression Scale

RCI	Reliable change index
RI	Reliable improvement
SAMHSA	Substance Abuse and Mental Health Services Administration
SDQ	Strengths and Difficulties Questionnaire
SSRI	Selective Serotonin Reuptake Inhibitor
STPP	Short-Term Psychoanalytic Psychotherapy
TADS	Treatment for Adolescents with Depression Study
TORDIA	Treatment of Resistant Depression in Adolescents
TP	Therapist
UCL	University College London
UK	United Kingdom
WHO	World Health Organization
YP	Young person

Publications and Presentations Associated with This Thesis

Journal Articles

Krause, K. R., Bear, H. A., Edbrooke-Childs, J., & Wolpert, M. (2019). What Outcomes Count? A Review of Outcomes Measured for Adolescent Depression Between 2007 and 2017. *Journal of the American Academy of Child and Adolescent Psychiatry*, *58*(1), 61–71. https://doi.org/10.1016/j.jaac.2018.07.893

Krause, K., N. Midgley, J. Edbrooke-Childs., M. Wolpert (2020). *"It woke me up to how my brain works": A comprehensive mapping of outcomes described following psychotherapy for adolescent depression*. Manuscript submitted for publication.

Krause, K., J. Edbrooke-Childs., M. Wolpert (2020). *Are we comparing apples with oranges? Assessing reliable improvement across symptoms, functioning, and goal progress in depressed adolescents.* Manuscript submitted for publication.

Presentations

Krause, K., Chung, S., The Working Group, Wolpert, M. (2019, October 27-28). *Devising a global core outcome set for young people with anxiety, depression, OCD and PTSD* [Paper presentation]. 5th International Conference on Youth Mental Health, Brisbane, Australia.

Krause, K., Chung, S., The Working Group, Wolpert, M. (2019, September 06). *Developing a core outcome set in children and young people's mental health: Depression & Anxiety* [Paper presentation]. Cundill Centre Speaker Series, Centre for Addiction and Mental Health, Toronto, Canada.

Krause, K., Chung, S. (2019, September 05). *Developing a core outcome set in children and young people's mental health: Depression & Anxiety* [Paper presentation]. NIHR Children and Families Policy Research Unit Child Maltreatment Core Outcome Development Workshop, Great Ormond Street Institute of Child Health, London, United Kingdom.

Krause, K. (2019, July 16). *What outcomes count? A systematic review of treatment outcomes for adolescent depression measured from 2007-2017* [Poster presentation]. University College London Institute of Mental Health Launch Event, London, United Kingdom.

Krause, K. (2019, March 30). What constitutes a ‘good’ outcome of treatment for adolescent depression? [Paper presentation] In P. Szatmari (Chair), *Counting what Counts: Selection and Reporting of Outcomes in Major Depressive Disorder* [Symposium]. Anxiety and Depression Association of America 39th Annual Conference, Chicago, IL, United States.

Krause, K., Bear, H., Childs, J., Wolpert, M. (2018, May 4). *What outcomes count? A systematic review of treatment outcomes for adolescent depression measured from 2007-2017* [Paper presentation]. British Psychological Society Annual Conference, Nottingham, United Kingdom.

Krause, K., Wolpert, M. (2017, July 3). *What counts as a good outcome in child mental health?* [Paper presentation]. Workshop: What Constitutes a Good Outcome in Child Mental Health? Anna Freud National Centre for Children and Families, London, United Kingdom.

Declaration of the Candidate’s Role in Each of the Studies

Guidance was provided throughout by my supervisors Prof. Miranda Wolpert and Dr. Julian Edbrooke-Childs.

Chapter 1. General Introduction

All work is the candidate’s own.

Chapter 2. Research Design and Methodology

All work is the candidate’s own.

Chapter 3. Outcome Taxonomy Review

All work is the candidate’s own.

Chapter 4. Outcome Measurement Review

The candidate conceptualised the review and undertook all the work. Holly Bear cross-validated the full text screening for 10% of the studies identified during title and abstract screening, replicated the mapping of outcome measures onto outcome concepts for 25% of the measures extracted from included studies, and cross-validated the data quality assessment for 25% of the included studies.

Chapter 5. Post-Therapy Outcome Perspectives

The conceptualisation of the research question, all data coding, and all analysis are the candidate’s own. The qualitative data for this study was collected as part of the IMPACT-ME study, led by Nick Midgley (see Midgley, Ansaldo, & Target, 2014 for a detailed description of the wider study). The candidate’s involvement began after data collection had been completed, with the transcription of 15 interviews. Nick Midgley provided feedback on a paper that was developed based on this chapter.

Chapter 6. Viewpoints on Outcome Priorities

All work is the candidate’s own. Prof. Paul Stenner provided advice on comparative Q-study design at the stage of conceptualising the study. Peter Schmolck advised on a specific methodological question relating to by-person factor analysis. Ana Calderon provided feedback on a draft version of the Q-set.

Chapter 7. Comparing Change Across Outcome Domains

The candidate conceptualised this study, computed all analytical variables, and conducted all analysis. The data utilised in this study was collected by child and adolescent mental health services in the United Kingdom (UK) between 2011 and 2015. Data collation, processing and cleaning was led by the Child Outcomes Research Consortium.

Chapter 8. General Discussion and Conclusions

All work is the candidate’s own.

Acknowledgements

I would like to extend my deepest gratitude to my supervisors Miranda Wolpert and Julian Edbrooke-Childs. Ever since my first day at Evidence Based Practice Unit (EBPU), your enthusiasm has been infectious. Having crossed over to child mental health from sociology, I could not have hoped for a more open, encouraging, and inspiring welcome to the field. I am extremely grateful for your empathetic guidance, and the many discoveries and encounters you have enabled. I would also like to extend my thanks to Jess Deighton and Dan Hayes for their mentorship over the past three years, and to Nick Midgley and Ana Calderon for their helpful inputs on specific studies. I gratefully acknowledge the support provided by the Anna Freud National Centre for Children and Families alongside the UCL Division of Psychology and Language Science in funding and enabling this PhD.

The completion of this thesis would not have been possible without the support of all those who took active part in the research or facilitated data access. I would like to thank the young people and professionals who donated their time and energy to the Q-study, as well as those who helped with recruitment. I am indebted to Nick Midgley and Sally O'Keeffe for facilitating access to the IMPACT-ME data, and to my colleagues at the Children's Outcome Research Consortium (CORC) for helping me navigate the depths of UCL's Data Safe Haven.

It is often said that PhDs are a lonely experience. I had the privilege of traversing PhD life with Holly Bear, Bettina Moltrecht, Shaun Liverpool, and Mia Eisenstadt, and am most grateful for their companionship, advice, and emotional support over the course of these three years. I would also like to thank my colleagues at EBPU and CORC for ensuring that my PhD experience was anything but lonely. I would like to extend a thanks to collaborators beyond the Anna Freud Centre and UCL, for the many inspiring conversations: Sophie Chung, Astrid Chevance, Nancy Butcher, Peter Szatmari, Martin Offringa, Suneeta Monga, Edward Harcourt, and the ICHOM working group for youth depression and anxiety.

I cannot begin to express my thanks to my friends from Hanover, Göttingen, Paris and London – scattered as you are and yet always close – for your support, encouragement and patience. A special thanks goes to Carrie Baptist and Spencer Wilton, and Y-ling Chi and Hector Page, for giving me wonderful homes in London during the final stages of the PhD. Another special thanks goes to Claire Weiß for being always there.

I am infinitely grateful for the support of my family: to my mother (Ulrike Hessel-Krause) and my brother (Dino Krause) for being beacons of comfort, laughter, warmth, and love; to my father (Jochen Krause) for being my north star; and to my partner Florian Poli for walking steadily beside me on this winding path, never tiring to point out discoveries left and right, always ready to take an unknown turn, and unwavering in your encouragement to keep going.

Statement of Funding

The PhD was funded through an Impact Studentship awarded jointly by University College London and the Anna Freud National Centre for Children and Families.

Chapter 1. General Introduction

1.1 Background

Over the past three decades, attention to outcome measurement in child mental health has grown significantly. In the context of spiralling costs and strained resources for health care generally, providers are increasingly held to account for delivering the best possible outcomes at the lowest cost (Kleinert & Horton, 2017; Porter, 2010). Within the evidence-based practice paradigm, there is an ongoing drive to identify the most efficacious treatment approaches based on randomised controlled trials and meta-analyses, and to recommend and justify their funding with reference to scientific evidence (APA Presidential Task Force on Evidence-Based Practice, 2006; Sackett et al., 1996). In parallel, mental health services are increasingly accountable for effectively delivering these treatments in practice, with performance indicators shifting from a focus on inputs, costs, or processes to a focus on outcomes (Chee et al., 2016; Garralda, 2009; Gray, 2017; Porter, 2010; Values-Based Child and Adolescent Mental Health System Commission, 2016). With evidence-based practice and data-driven management, the measurement of treatment outcomes has moved to the core of policy making, service planning, and commissioning in child mental health.

A growing emphasis on patient-centred care means that a ‘good outcome’ is increasingly defined from the perspective of service users and their families (Coulter, 2017; Mulley et al., 2017). A broad alliance of governments (Australian Commission on Safety and Quality in Healthcare, 2010; Department of Health, 2010; The Scottish Government, 2009; US Department of Health and Human Services, 2008), international organisations (World Health Organization [WHO], 2007), and service user and health policy organisations (International Alliance of Patients’ Organizations, 2007; National Institute for Health and Care Excellence, 2011; Picker Institute, 2004; The Health Foundation, 2016; The King’s Fund, 2012) have stressed the importance of placing service users and their needs in the centre of health care decision-making (Kitson et al., 2013). Patient-centred care has been defined by the US Institute of Medicine as care that is “respectful of and responsive to individual patient preferences, needs, and values and ensuring that patient values guide all clinical decisions” (Wolfe, 2001, p. 243). As part of this ambition, it is crucial that outcomes measured reflect what truly matters to service users and their families. If they are genuinely at the core of mental health care delivery, decisions must be made based on outcome data that reflects their priorities and needs. However, there is currently no consensus and limited data on what outcomes matter most to young people and families.

In a parallel development, the impact of depression on health services has become a growing concern. As defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association [APA], 2013), depression is characterised by the core symptoms of persistent depressed mood and/or loss of interest or pleasure in daily activities (i.e., anhedonia), as well as secondary symptoms such as sleeping difficulties, fatigue or loss of energy, changes in weight or appetite, slowing-down of thoughts and a reduction of physical movements, issues with concentration, excessive feelings of guilt or worthlessness, and suicidal thoughts or behaviour.

The World Health Organization (WHO, 2012) has identified depression as a “leading cause of disability worldwide in terms of total years lost due to disability” (p.6), and estimates the number of people affected at over 300 million (WHO, 2017a). Depression is a chronic and recurrent disorder (Hardeveld et al., 2010, 2013; Kessler & Bromet, 2013; Torpey & Klein, 2008), associated with significant functional impairment (Wells et al., 1989), morbidity, and mortality (Angst et al., 2002; De Hert et al., 2011; Keller, 2003; Pan et al., 2011; Rutledge et al., 2006; Van der Kooy et al., 2007; Whooley et al., 2008). A recently assembled clinical commission on depression involving public and mental health experts has stressed the urgent need to deliver evidence-based interventions and high-quality care to address the global burden (Herrman et al., 2019).

Depression frequently has its onset in adolescence (Kessler & Bromet, 2013; Thapar et al., 2012), defined here as the period from the age of 12 to the age of 21 years (Eunice Kennedy Shriver National Institute of Child Health and Human Development, 2019; Williams et al., 2012). Depressed children and adolescents may display more anxiety and anger, and fewer changes in sleep, energy levels, appetite and other biological symptoms than depressed adults (Fonagy et al., 2015). In general, however, depression is considered to have broadly similar characteristics across the life course (Costello et al., 2002; 2003). In the United Kingdom, the estimated prevalence rate is 5.6% (Costello et al., 2006). Around 50% of young people presenting to child and adolescent mental health services (CAMHS) in England show symptoms of depression (Wolpert et al., 2016). According to the WHO (2019), depression is one of the most common causes of disability in adolescence, and suicide is the third most-common cause of death in 15-19-year-olds. Adolescent-onset depression is associated with adverse outcomes across the life span, such as increased rates of recurrence (e.g., Costello et al., 2002; Fombonne et al., 2001; Kovacs et al., 2016; Patton et al., 2014), suicide, other mental and physical health problems (Gould et al., 1998; Jonsson et al., 2011; Rudolph & Klein, 2009), school drop-out, low educational attainment, unemployment, and reduced income (Clayborne et al., 2019; Fletcher, 2008, 2013; Holsen & Birkeland, 2017; Kessler et al., 2001). Effective treatment of adolescent depression is thus a key aspect of any comprehensive effort to reduce the burden from depression.

It has been argued that in health care, outcomes tend to be inherently problem-specific, as both the service users’ needs and the treatment offered tend to be at least partly shaped by the main presenting problems (Kazdin, 1999b; Porter, 2010).Therefore, it is suggested that value be defined and measured specifically for each condition, while accounting for the presence of comorbid difficulties (Porter, 2010). In child mental health, young peoples’ developmental stage is another aspect to consider when assessing what might constitute a ‘good outcome’ of treatment (Hoagwood et al., 1996).

In the past year, two efforts have been launched to develop standard sets of recommended outcome measures for depression in young people. One set is being developed for use in clinical practice under the lead of the International Consortium of Health Outcomes Measurement (ICHOM, see www.ichom.org) and focusses on anxiety, depression, obsessive-compulsive disorder, and post-traumatic stress disorder in children and young people (K. Krause et al., 2020). The doctoral candidate has been involved in this initiative as a research fellow since its inception in October 2018. An outcome set for clinicial trials focussing more specifically on adolescent depression is being developed by Monga and colleagues (2019) at Toronto Outcomes Research in Child Health (Hospital for Sick Children Research Institute) and the Cundill Centre for Child and Youth Depression (Centre for Addiction and Mental Health) in Toronto (see http://www.comet-initiative.org/studies/details/1122). These timely initiatives are currently facing scarce evidence on outcome priorities amongst depressed young people and their families, as well as limited conceptual guidance.

1.2 Thesis Aims and Organisation

This doctoral thesis aimed to advance the debate on what constitutes a ‘good outcome’ in relation to adolescent depression, both conceptually and empirically. Any evaluation of ‘good outcome’ is likely to vary based on (a) how outcome is conceptualised, and what types of outcome are considered relevant; (b) the perspective, priorities, and values of those who are asked; and (c) aspects of how these outcomes are then measured empirically. This doctoral thesis used a mixed-methods research design to examine these issues from different angles within a pragmatist research paradigm (see Chapter 2 for further details). The thesis is organised as follows:

- **Chapter 1 – General Introduction:** The remainder of this chapter provides an overview of key outcome definitions and concepts in mental health; critically reviews the existing evidence base on outcome perceptions and priorities amongst adolescents, parents, and clinicians; and introduces key issues related to outcome measurement.
- **Chapter 2 – Research Design and Methodology**: The second chapter outlines the methodological approach underpinning this doctoral thesis.
- **Chapter 3 – Outcome Taxonomy Review**: This chapter presents findings from an integrative narrative review of existing outcome taxonomies, which were synthesised into a new, more comprehensive taxonomy. This then served as a conceptual framework throughout this doctoral thesis, and was progressively refined based on findings emerging from the studies presented in Chapters 4 through 6.
- **Chapter 4 – Outcome Measurement Review**: This chapter presents findings from a systematic review that aimed to establish what types of outcomes have

been reported in the recent treatment effectiveness literature for adolescent depression. These outcomes were mapped conceptually using the taxonomy developed in Chapter 3.

- **Chapter 5 – Post-Therapy Outcome Perspectives**: This chapter presents a study that aimed to identify the types of outcomes discussed by young people, parents, and clinicians following a course of psychotherapy for depression, using the taxonomy developed in Chapter 3 as an initial analytical framework. The salience of different outcomes in participants' narratives were compared with their salience in the recent treatment effectiveness literature, as well as between participant groups and treatment modalities.
- **Chapter 6 – Viewpoints on Outcome Priorities**: This study focussed on soliciting outcome priorities amongst young people and clinicians through use of a card-sorting exercise, and on identifying distinctive viewpoints within and between both groups through inverted factor analysis (i.e. Q-methodology).
- **Chapter 7 – Comparing Change Across Outcome Domains**: Quantitative analysis of change metrics across different outcome measures and domains was conducted using a large dataset of naturalistic, routinely collected outcome data. The study aimed to assess the extent to which measures designed to capture the same outcome yielded comparable findings for individual service users, and the extent to which levels of change in one domain translated to equivalent levels of change in another domain.
- **Chapter 8 – General Discussion:** This last chapter reviews the main aims of this doctoral thesis, its key findings, strengths and limitations, and implications for future research and practice.

This thesis deliberately refrains from approaching the question of 'good outcome' through a theoretical framework linked to any particular therapeutic modality. Instead, it aims to consider the widest possible range of outcomes, to devise a conceptual framework that can be applied across treatment modalities, and to explore stakeholder perspectives and priorities without excluding or prioritising views linked to any particular approach. As such, none of the studies constituting this thesis placed deliberate restrictions on the treatment types and contexts to consider.

Throughout this thesis, adolescents will be referred to as 'young people' or 'service users'. The term 'patient' will be avoided as it suggests a "a clear dividing line between the sick and the sane" within a medical framework of mental illness that is prone to creating stigma (Christmas & Sweeney, 2016, p. 11). Instead, young people accessing support will be defined by their use of services with the aim of positioning mental health difficulties as "part of the human condition" (Christmas & Sweeney, 2016, p. 12). For the purpose of consistency, ‘patient-centred care’ as an established term will therefore be replaced with the term ‘person-centred care’. For the sake of brevity, the term “parent” will henceforth be used to cover parents, as well as other primary caregivers. Where there is need to distinguish both groups, this will be made clear in the text. The term “clinician” will be used throughout to refer to mental health professionals who engage with service users in clinical practice, such as psychiatrists, psychologists, and nurses.

1.3 Outcome Concepts

1.3.1 Defining Outcome

In the context of psychotherapy research, outcomes are defined as changes that occur as a result of treatment, either towards the end of treatment, or in the longer term (Sperry et al., 1996). As discussed in the following sections, different types of change may be considered. Most treatment effectiveness research in child mental health has defined ‘good outcome’ in terms of symptom reduction and to a lesser extent as a reduction in functional impairment (Bear et al., 2020; Becker et al., 2011; Hoagwood et al., 2012; Weisz et al., 2005). However, other relevant changes may occur in other domains of young people’s lives, such as their relationships, personal growth, or physical health. Outcomes can also be intermediary steps in the process of realising higher-level benefits, in which case they are frequently determined with reference to the specific theoretical framework underpinning a given treatment approach (Cuijpers, 2019). Outcomes can also be negative when they take the form of side effects and adverse events. Outcomes may be considered at the level of individual service users and their families, or at the level of services and systems (Trauer, 1998). Person-centred outcomes correspond to the gains for a young person and their family, while service outcomes correspond to indicators of service performance, that is, metrics that reflect the quality of care. Such metrics include, for example, service use statistics; service user satisfaction; and the availability, resourcing, planning, and quality of services (Childs et al., 2013; Fonagy, 1997; Hoagwood et al., 1996). This thesis will focus on what constitutes a good outcome for service users and their families, thus taking a person-centred approach. While consideration of undesirable outcomes, and service or system-level outcomes are important aspects of considering outcome in psychotherapy research, it was beyond the scope of this project to cover these.

A concept that relates to personalised notions of ‘good outcome’ is that of treatment goals. Goal setting has been used in mental health for decades (e.g., J. D. Frank & Frank, 1993; Urwin, 2007) to identify a “desired endpoint of treatment”, as well as intermediate steps towards reaching this endpoint (Jacob et al., 2018, p. 111). Goals, typically defined at the outset, reflect the expectations and priorities that service users bring to treatment. They are thus conceptually related to the outcomes that these service users value. While outcomes usually describe the changes achieved ex post, goals describe notions of *desired outcome* ex ante. While the outcomes literature is closely linked to the assessment of treatment effectiveness, the goal literature focusses on goal-setting as a clinical tool, and on how it influences the treatment process and eventual outcomes (e.g., Cooper & Duncan, 2018). While the conceptual discussion presented in Section 1.3 will focus on the outcomes literature, the review of empirical studies of outcome perceptions presented in Sections 1.4 and 1.5 will also consider studies exploring goals. The body of empirical research on outcome perceptions in child mental health is scarce, and goal-focussed research provides useful additional evidence on the types of change that young people (as well as parents and clinicians) may value. A more detailed conceptual discussion of goal setting and goal-based outcome measurement is provided in Section 1.7.1.

1.3.2 Conceptual Issues and Debates

A Focus on Symptoms

As mentioned above, outcome measurement in child mental health – and in psychotherapy research more generally – has predominantly focussed on measuring change in symptom severity, and to a lesser extent, in functional impairment (Bear et al., in press; Becker et al., 2014; Becker et al., 2011; Hoagwood et al., 2012; Weisz et al., 2005). In the early 1990s, a set of consensus definitions were proposed for depression, which specified treatment outcomes in relation to the number and intensity of symptoms, and the timelines of symptomatic deterioration or improvement (E. Frank et al., 1991):

- ***Response*** denotes a clinically significant reduction in depression symptoms following the onset of treatment, even though moderate symptoms may remain.
- ***Remission*** is achieved when a service user no longer meets diagnostic criteria for depression, and experiences no more than minimal residual symptoms.
- ***Recovery*** denotes a sustained period of remission that lasts long enough to mark the end of the initial depressive episode, and during which diagnostic criteria for major depression are not met.
- ***Relapse*** denotes the return to clinically significant symptom levels warranting a diagnosis of major depression before recovery has been achieved.
- ***Recurrence*** refers to the onset of a new depressive episode meeting diagnostic criteria, following recovery from a previous episode.

These consensus definitions have since been widely adopted, with symptom remission often described as the ‘optimal outcome’ of acute treatment for depression in the adult literature (Ballenger, 1999; Ferrier, 1999; Nierenberg & Wright, 1999; Rush et al., 1998), as well as in child mental health (Birmaher, 2007; Lewandowski et al., 2013; Park & Goodyer, 2000; Tao et al., 2009). A ‘good outcome’ is thus equated to restoring a service user to a ‘healthy’ mental state (Keller, 2003).

Symptom-focussed approaches to outcome measurement have a number of strengths. First, they establish treatment success with reference to the same diagnostic criteria that are used for initial assessment, as per the Diagnostic and Statistical Manual of Mental Disorders (DSM; APA, 2013) or the International Classification of Diseases (ICD; WHO, 2018). These constitute familiar and widely respected frames of reference amongst clinicians (Cuijpers, 2019). Second, symptom-focussed definitions such as those proposed by Frank and colleagues (1991) have the benefit of conceptual clarity and can be operationalised with relative ease (Davidson & Roe, 2007). By reporting outcomes with reference to consensus definitions, results can – in theory – be compared across studies and service settings, although caveats have been raised about the equivalence of different measurement tools with regards to their sensitivity to change (Fried, 2017; Snaith, 1993; Chapter 7), and inconsistencies in the significance thresholds and norms applied (McGlinchey et al., 2008). Third, as a result of the long-standing focus on symptoms in efficacy trials, there is a proliferation of symptom measures, which have often been more widely and carefully validated than measures of other outcome concepts (Becker et al., 2011; Bickman & Rosof-Williams, 2000; McGlinchey et al., 2008). Fourth, remission and recovery have been consistently shown to predict longer-term outcomes and trajectories, including the risk of relapse and recurrence (Judd et al., 1998; Kennedy & Foy, 2005; Lin et al., 1998; Zimmerman et al., 2008), which suggests that they are clinically meaningful indicators.

The Challenge of Arbitrary Metrics

Despite the above-mentioned strengths, symptom-focussed outcome measurement faces a number of limitations and challenges. One challenge worth noting is that many commonly used symptom scales are not immediately interpretable with regards to how a score change translates into real-world changes in a service user’s life (Blanton & Jaccard, 2006; Kazdin, 1999b, 2006). While the clinical significance of a score change is typically assessed with reference to whether or not post-treatment scores fall into a normative range (see Section 1.7.2), this does not enable direct conclusions about the magnitude of change actually experienced by individuals in their daily lives. It has been argued that symptom scores should therefore be assessed and calibrated in relation to more concrete reference points, such as a young person returning to school, global measures of functioning, or service users’ perceptions of change (C. E. Hill et al., 2013; Kazdin, 1999, 2006; Sechrest et al., 1996). Where symptom change has been compared with change in other outcome domains, convergence has indeed been shown to be imperfect in relation to subjective perceptions of change (Karpenko & Owens, 2013), functioning (Becker et al., 2011; Brookman-Frazee, Haine, & Garland, 2006) educational outcomes (Becker et al., 2014), and family functioning and relationships (Brookman-Frazee et al., 2006; Kazdin & Wassell, 2000).

Considering Functional Impairment Alongside Symptom Change?

Functional impairment has been highlighted as a second outcome domain, that could provide valuable information if measured alongside symptom change. In child mental health, functioning has been described as “the ability of children to adapt to varying demands of home, school, peer group, or neighborhood”, representing a continuum from high levels of adaptation and competency on one end of the spectrum, to impairment on the other end (Hoagwood et al., 1996, p. 1060). A multitude of functioning measures exist that assess functioning either globally or in specific life domains (Canino et al., 1999). The *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) requires that clinically significant levels of symptom severity must be accompanied by significant functional impairment to warrant a DSM diagnosis (APA, 2013). A logical consequence might be to assess treatment success with reference to both outcome domains (Zimmerman et al., 2008). Moreover, assessment of functioning can help identify the life domains where the impact of symptoms is most strongly felt, and in tracking change in these specific domains, can help understand whether treatment is effective at alleviating relevant difficulties (McKnight & Kashdan, 2009). As such, measurement of functioning can help address concerns about the arbitrariness of symptom metrics discussed above (Becker et al., 2011; Blanton & Jaccard, 2006; Kazdin, 2006; Sechrest et al., 1996). Lastly, it has been shown that levels of functioning influence help-seeking, as well as providers’ decisions about whether or not a young person warrants clinical support, and the types of services they should receive (Hodges et al., 2000; Striley et al., 2003).

Existing evidence suggests that symptom change is an imperfect proximal indicator of improved functioning. On the one hand, many children with above-threshold symptoms are not significantly impaired, while on the other hand many children who do experience substantial impairment do not meet diagnostic criteria of symptom severity (Angold et al., 1999; Costello et al., 2003; Pickles et al., 2001; Simonoff et al., 1997). One study comparing change in symptoms, functioning, and family functioning in a sample of 112 adolescent outpatients (49% of whom had a mood disorder) found only minimal agreement in meaningful improvement ratings between these three outcome domains (Brookman-Frazee et al., 2006). A systematic review of randomised controlled trials of psychosocial and combined treatments for youth compared the strength of the evidence presented by each study in favour of a treatment effect on (a) symptoms and (b) functioning (Becker et al., 2011). The majority of studies did not report on functioning, but where symptoms and functioning were both measured, the evidence supporting changes in functioning was considerably weaker than the evidence for symptom reduction. The authors concluded that it was more difficult to provide even minimal empirical support for changes in functioning, compared to changes in symptoms. As functioning may improve independently from symptoms (and vice-versa), its separate assessment appears warranted (Keller, 2003).

Considering the Multidimensionality of Developmental Psychopathology

Another challenge to narrow, symptom-focussed approaches to outcome measurement is that the development of mental health difficulties is typically shaped by a multitude of biological, psychological and social factors. Mental health difficulties may disrupt different domains of life, such as daily functioning, relationships, and physical health. Depression in particular has been described as “multifactorial” (Weeks et al., 2016, p. 37) with regards to its risk factors (Clarke & DeBar, 2010; Kassis et al., 2017; Weeks et al., 2016) and its various adverse impacts on wellbeing and functioning (Clayborne et al., 2019; Fletcher, 2008, 2013; Holsen & Birkeland, 2017; Kessler et al., 2001; Klein et al., 2005). Psychotherapy is often equally multidimensional in attempting to address difficulties in different domains of life and their driving factors (Stulz & Lutz, 2007). If outcome measurement is to reflect the multifactorial nature of depression and of treatment mechanisms, a multidimensional approach to measurement may be required (Dirks et al., 2012).

In child mental health, such considerations have been taken further, influenced by insights from developmental psychopathology about the complex ways in which individual and environmental risk and protective factors interact to influence child development (e.g., Cicchetti & Toth, 1995), treatment effectiveness and the maintenance of treatment benefits (Kazdin & Wassell, 2000). Hoagwood and colleagues (2012, 1996) propose a “developmentally grounded, dynamic, and interactional” model of outcome for child mental health (1996, p. 1057). It recommends measuring outcomes across seven domains: *symptoms and diagnoses*, *functioning*, *consumer perspectives*, *environments*, *systems*, *parental symptoms* and *general health*. The authors specify that functioning should be assessed with attention to the different environments in which children interact (e.g. family, school, or community). A similar model suggesting five dimensions of outcome has been proposed by Fonagy (1997; see Appendix A, Table A.1 for a more detailed overview of both frameworks). These dimensions include (a) the *symptomatic level*, (b) *adaptation to the psychosocial environment* (i.e., functioning), (c) a *transactional level* focussing on the child’s interactions with their environment (e.g. family, community, school and wider society), and (d) *service utilisation* and quality of care. The fifth dimension are the cognitive and emotional *mechanisms* underpinning mental health difficulties (e.g., affect regulation, understanding emotions, self-representations), which treatment approaches will likely target. These have repeatedly been highlighted as a neglected topic in outcomes research (Cuijpers, 2019; Kazdin, 2009). Both models stress that in order to fully understand whether and how treatment is effective, outcomes must be considered across multiple dimensions and in terms of how they interact in shaping a child’s development. Importantly, these models further stress that children and families (and possibly their wider environment) influence one another reciprocally, where a reduction in child symptoms may have a positive effect on family functioning and parental wellbeing and vice-versa (Kazdin & Wassell, 2000; Kazdin & Weisz, 1998).

The Imperative to Consider Outcomes from the Service User’s Perspective

Within the framework of person-centred care, it appears imperative to measure outcomes that are meaningful to service users themselves, as well as to their families. In adult mental health, symptom-focussed outcome frameworks have been challenged by the service-user-led *recovery movement*, which emerged in the 1980s and advocated for more personalised notions of recovery (Bellack, 2006; Davidson & Roe, 2007; Deegan, 1988). In this vein, one influential definition describes recovery as:

> a deeply personal, unique process of changing one's attitudes, values, feelings, goals, skills and/or roles. It is a way of living a satisfying, hopeful, and contributing life even with limitations caused by the illness. Recovery involves the development of new meaning and purpose in one's life as one grows beyond the catastrophic effects of mental illness. (Anthony, 1993, p. 527)

Within the recovery literature, ‘good outcomes’ have frequently been described in relation to themes such as connectedness (e.g., relationships and peer support), hope and optimism about the future (e.g., having dreams and aspirations), identity (e.g., rebuilding a positive sense of self), meaning in life (e.g., quality of life), and empowerment (e.g., resuming control over one’s life) (Leamy et al., 2011). In child mental health, concepts of recovery have not generally been applied, partly due to concerns about the notion of restoring young people to a previous state, given that they are in constant development; and concerns that outcomes such as self-determination and responsibility may not apply to all age groups (Friesen, 2007). While the literature on experiences and perceptions of outcome amongst adolescent service users is scarce, a discussion of relevant studies is provided in the next section.

1.4 Service User Perceptions of Outcome

Two studies have suggested that adolescents and young adults experience recovery and outcomes in ways similar to those described in the adult recovery literature. Lavik and colleagues (2018) conducted semi-structured interviews and focus groups with 22 adolescents who were undergoing or had just completed treatment at CAMH services in Norway. They described good outcomes in terms of a journey towards “a stronger autonomy and safer identity” (Lavik et al., 2018, p. 4), and described outcomes in six more specific areas: (a) *developing a better understanding of and ability to cope with feelings and thoughts*; (b) *becoming the person they really are, independent from societal pressures*; (c) *opening up and feeling more connected with family and peers*; (d) *embracing new challenges*; and (e) *being able to cope with life’s challenges*. Symptom change was not identified as a core outcome theme. Bergmans and colleagues (2009) investigated recovery narratives and turning points amongst young adults who had completed a psychosocial intervention following recurrent suicide attempts. As in the Norwegian study, recovery was described as an individual journey, which involved learning to live with enduring symptoms, gaining back a sense of choice, connecting with others, restoring hope, being able to envisage the future, and learning to understand and cope with symptoms.

While demonstrating strong qualitative methodologies overall, both studies have some similar limitations, relating mainly to participant recruitment. Lavik and colleagues (2018) recruited participants with the help of clinicians who suggested “experienced service users” (i.e., young people who had been in treatment for at least six months), thus possibly introducing selection bias towards youth able to provide rich accounts at the expense of representing a broader range of experiences. Bergmans and colleagues (2009) recruited youth who had completed treatment between six months and six years prior, thus allowing for a high degree of heterogeneity in perspectives and experiences, which was, however, not explored. In both studies, dominant outcome themes were discussed in detail, but no attention was given to divergent views or heterogeneity in experiences.

Heterogenous experiences were, in turn, a core focus of a study conducted by Gibson and Cartwright (2014) with 22 adolescents who had accessed school-based counselling in New Zealand. While they did not analyse perceptions of outcome explicitly or systematically, they identified four distinct narratives about the counselling process which also touched upon outcomes: The "transformative" narrative described dramatic and profound changes, mainly in relation to the self; the "supportive" narrative described counselling as holding young people in place as they confronted challenges in their lives; the "pragmatic" narrative described counselling as helping young people resolve specific issues; and the "disappointed" narrative considered that no positive change had occurred (Gibson & Cartwright, 2014). Similarly to Lavik and colleagues (2018) and Bergmans and colleagues (2009), they concluded that young people's accounts did not support a linear, symptom-focussed understanding of outcome. Another notable finding was that having access to a therapeutic space constituted an important outcome in and of itself, which was consistent with findings from other studies focussing on helpful and unhelpful aspects of therapy (Binder et al., 2011; Dunne et al., 2000; Freake et al., 2007; Gibson et al., 2016).

Two studies completed in the UK examined therapy goals defined by children and young people at the start of treatment, using an existing goal taxonomy developed in adult mental health as an initial coding framework (i.e. the Bern Inventory; Grosse Holtforth & Grawe, 2002), which was then inductively adapted. Bradley and colleagues (2013) analysed a set of goals defined by 80 children and young people who had accessed one of six public CAMH services. The most frequently mentioned goal themes related to young people's *personal growth*, *functioning* and *coping with specific symptoms*. Due to the structure of the dataset, the authors were unable to specify key demographic characteristics relating to age and presenting problems, thus making it difficult to judge of whom these goals were representative. Rupani and colleagues (2014) presented a secondary analysis of goal data from two pilot randomised control trials (RCTs) of school-based counselling, which involved 73 adolescents with at least moderate psychological distress. The most frequent goal types were *increased self-confidence and self-acceptance*, *reducing anger*, *feeling happier and less upset*, *improved relationships with family and friends*, as well as *reduced anxiety*.

One qualitative study has examined outcome perceptions specifically amongst depressed adolescents (Cortés et al., 2018). A small sample of six adolescents (and their therapists, see Section 1.5) were interviewed following treatment at university-based or private mental health services. Young people frequently described outcomes in terms of *intrapsychic* *change*, which comprised improved well-being and calm, improved ability to manage mood and emotions, and acquiring coping skills and strategies, as well as increased initiative, motivation and ability to make decisions. Other outcomes mentioned included improved family interactions, strengthened parental support, and a decrease in depressive symptoms. While the study authors suggested that young people valued other changes more that symptom reduction, this conclusion appeared partly driven by their decision to split the symptom change category from the commonly endorsed domain of *intrapsychic change*. However, as examples of the coding frame are not provided, this is difficult to ascertain.

These studies showcase a range of treatment outcomes and changes valued by children and young people, which cut across different domains such as *symptom change*, *functioning*, *relationships* and *personal growth*. They suggest that there may indeed be a mismatch between symptom-focussed approaches to outcome measurement, and young people's perspectives and priorities. *Learning to cope with symptoms* was a predominant theme, as was *improvement in family functioning and relationships*. At the same time, this existing body of research is limited in a number of ways. First, only one study has focussed specifically on adolescent depression, and this was based on a very small sample and a coding framework that was less nuanced than those presented in other studies (Cortés et al., 2018). The majority of existing studies explored outcomes in a sample of youth with a mix of presenting problems, without exploring diverse perceptions and priorities (Bradley et al., 2013; Gibson & Cartwright, 2014; Lavik et al., 2018; Rupani et al., 2014). Second, none of these existing studies approached the data through a coding frame that was theoretically grounded in a developmental model of outcome such as those suggested by Hoagwood and colleagues (1996) or Fonagy (1997), and none discussed their findings with reference to this literature. Across most studies, initial preconceptions held by the researchers were not reflected on, and possible sources of bias to their analysis were thus not made transparent. The existence of an external frame of reference, such as in the form of an outcome taxonomy, could help avoid the introduction of such bias, and could also facilitate the comparison and integration of qualitative findings across studies.

1.5 Perspectives of Parents and Clinicians

Within a person-centred framework, young people may be considered the most important stakeholders in their treatment. However, as emphasised by developmentally informed models of outcome (e.g., Fonagy, 1997; Hoagwood et al., 1996), carers and family members also play an central role. It is often them who initiate treatment, rather than young people themselves, and they have an important role in supporting treatment (Kazdin & Weisz, 1998). Another crucial stakeholder group are clinicians. Specific training and professional experience may enable them to detect changes that service users themselves may not be aware of, or not consider important. Clinicians (as well as researchers) also typically lead the selection of outcome measures. Where perceptions diverge, this poses a dilemma: Whose views should be prioritised, which outcomes should be measured, and from which perspective? Any consensus on what constitutes a good and important outcome of treatment must consider all three of the above-mentioned perspectives, if not more.

In their tripartite model of outcome, Strupp and Hadley (1977) argue that outcome assessment is always subjective and driven by assumptions about desirable types of behaviour. They suggest that service users, family members (who they consider represent the view of the wider society), and clinicians judge outcomes from different vantage points, based on different values and priorities: Wellbeing and happiness may be most highly valued by service users; clinicians may judge 'good outcomes' in relation to their theoretical training; and family members and wider society may focus on manageable behaviour and conformity with social expectations. Strupp and Hadley (1977) suggest that outcome measurement should therefore consider all three vantage points, and anticipate a considerable level of discrepancy.

Indeed, young people, parents, and clinicians frequently provide discrepant reports on the severity of young people's presenting problems (Achenbach, 2006; Bird et al., 1992; Cantwell et al., 1997; De Los Reyes et al., 2015; De Los Reyes et al., 2019; De Los Reyes et al., 2011; De Los Reyes & Kazdin, 2004, 2005, 2008; Ferdinand et al., 2006; Salbach-Andrae et al., 2009), including in relation to depression and emotional problems (Angold et al., 1987; Bear et al., in press.; Castagna et al., 2019; Cole et al., 2000; Cuijpers et al., 2010; Makol et al., 2019; Makol & Polo, 2018). While such discrepancies complicate judgements about treatment effectiveness, they are now generally considered to be informative, as each informant is seen to represent a valuable and insightful perspective (De Los Reyes et al., 2015; De Los Reyes et al., 2013; De Los Reyes & Kazdin, 2005, 2008; Kraemer et al., 2003). This is based on the understanding that young peoples' behaviour and the manifestation of presenting problems often differ across the different contexts in which they interact. Disagreement may thus provide valuable information on how mental health problems manifest themselves, and how treatment affects these manifestation across different contexts (Achenbach et al., 1987; De Los Reyes et al., 2011; Dirks et al., 2012). As stressed by Weisz and colleagues (2017):

> Youth therapy outcome is always, to some extent, in the eye of the beholder, and […] different informants observe different samples of a youth's behavior, in different contexts, and bring different perspectives to what they observe. (p. 95)

While the above-mentioned research has focused on comparing multi-informant agreement on symptom ratings, the body of research investigating different perspectives on what constitutes a good outcome, or outcome perceptions beyond the symptom domain, is considerably smaller. Only one study has provided a qualitative comparison of outcomes described by young people and clinicians following treatment. Cortés and colleagues (2018) compared perceptions of outcome amongst six dyads of depressed adolescents and their clinicians and found that *intrapsychic changes* (i.e., changes in affect, attitude, and interpretation of events) were particularly emphasised by young people, while clinicians tended to describe changes in symptoms and behaviour. However, as mentioned above, some of the changes in wellbeing assigned to the category of *intrapsychic changes* could equally have been considered within the category of depressive symptoms and vice-versa. Young people and clinicians converged in describing improvements in relationships with family and friends, family functioning, and parenting. A systematic comparison of outcome perceptions (e.g., using indicators of salience in each group) was not provided.

A larger number of studies has focussed on comparing problem perceptions, treatment goals, or desired outcomes at the *start* of treatment. One of the earliest studies in this area explored agreement between youth aged 7-18 years and their parents on the principal problem that had brought them to therapy (Yeh & Weisz, 2001). Following intake, problem descriptions were collected from 381 dyads who sought treatment at community mental health services in the United States. They were coded into problem types and higher-level domains with reference to the Child Behavior Checklist (CBCL; Achenbach, 1991). Of all dyads, 63% did not agree on a single target problem and more than one third did not agree on a broad problem domain. Young people and parents agreed more frequently on behaviour problems than on emotional problems, and on aggressive behaviour in particular. In a subsequent study, Hawley and Weisz (2003) applied the same methodology to triads of youth, parents, and clinicians. Of 315 triads, 77% were unable to agree on a single target problem, and 44% were unable to agree on a broad problem area. Again, agreement about behaviour problems was found to be significantly higher than agreement on emotional problems, although remaining low overall. Across all three groups, the most reported target problem was aggressive behaviour (e.g., disobedience, temper tantrums). Youth were found to more frequently define target problems in relation to parenting and family functioning, while parents tended to focus on the child's behaviour and emotional problems. Clinicians were more likely to endorse emotional and behavioural problems than children, but also more likely than parents to define problems around family functioning and relationships.

Agreement on *desired outcomes* was assessed by Garland and colleagues (2004) in a sample of 170 triads of adolescents, parents, and clinicians at public outpatient services in the United States. They applied Hoagwood's (1996) conceptual model to categorise outcomes and found that across all three stakeholder groups, desired outcomes most frequently related to *symptoms* and *functioning*, with *reduced anger and aggression* as the most frequently endorsed individual outcome (Garland et al., 2004). The third most mentioned outcome domain was the *environment*, and a small share in all three groups further defined outcomes in the domains of *consumer perspectives* and *systems.* Only 38% of triads agreed on at least one outcome. Therapists and young people more often endorsed outcomes around family relationships, compared to parents, while parents were more likely to define outcomes in relation to obedience and improved self-esteem. Clinicians more frequently defined outcomes around communicating feelings and thoughts than either young people or parents.

Two studies have compared treatment goals at the start of treatment. Jacob and colleagues (2016) thematically analysed goals that had been defined by parents and youth (aged 4-17 years) as part of accessing routine mental health care in the UK. They found that youth tended to focus more on emotional problems and issues related to personal growth (e.g. managing negative thoughts or feelings, greater self-confidence, increased autonomy), while parents focussed primarily on behaviour management (e.g. better management of sleep, reduction of risky behaviours), family functioning, and their own wellbeing and parenting skills. However, the direct comparison of perspectives in terms of goal theme salience was hampered by the use of separate coding frameworks for young people and parents. Odhammar and Carlberg (2015) compared goals defined by 33 dyads including parents of 5-10-year olds, and clinicians, and found explicit agreement amongst half. They found that clinician goals appeared strongly influenced by their theoretical training. The most common goal category was the child's *intrapsychic development* (p. 284), such as their ability to manage impulse and emotions; understand and reduce negative affect, and express positive affect; increased self-esteem and self-confidence; empathy and reflexive ability. Parental goals, in turn, were focussed primarily on the child being "more secure and harmonious", better able to identify and seize possibilities in life, and on greater self-confidence and happiness (Odhammar & Carlberg, 2015, p. 287). Since the authors did not display their final coding frame, or indications of the salience of different outcome themes, it is difficult to draw conclusions about the salience of each outcome from both perspectives.

A number of findings emerge consistently from some or all of these studies. First, dyadic and triadic agreement tends to be relatively low, with youth, parents, and clinicians focussing on different types of problems or outcome. Second, agreement tends to be highest in relation to behaviour problems and aggression. This is in line with the literature around agreement on key symptoms, which equally shows higher agreement on behavioural as opposed to emotional difficulties (Hodges et al., 1990). Behavioural problems may be more observable in the form of delinquency, disruptive behaviour, and their effects on functioning at home, in school and in the community (Yeh & Weisz, 2001). Third, parents appear to focus more on managing and improving youth behaviour, while youth more frequently focus on emotional difficulties and self-esteem. Fourth, there is some evidence that clinicians focus particularly on *intermediate* outcomes that relate to their theoretical training and approach.

These studies reflect similar limitations as those discussed in the previous section in relation to youth perceptions. Only one study (Garland et al., 2004) referred to the conceptual framework proposed by Hoagwood and colleagues (1996) and used this to systematically assess outcomes across a range of domains, across different perspectives. Two studies used the CBCL as a guiding framework (Hawley & Weisz, 2003; Yeh & Weisz, 2001), and one used an existing goal taxonomy as a starting point (Jacob et al., 2016). The remaining studies neither approached the data through a conceptual framework, nor presented a detailed record of their coding frame. Conclusions about the salience of different outcome themes in different subgroups thus lack transparency. Another limitation is that most studies have used heterogenous samples, including both children and adolescents with a range of presenting problems, without disaggregating their findings for specific subgroups. It remains unclear what outcomes different stakeholder groups value and prioritise specifically for adolescent depression, and to what extent they agree. Given the high levels of agreement on behaviour problems, agreement could be expected to be even lower amongst those presenting with emotional problems. Most studies have assessed agreement at the start of treatment. It is possible that agreement might increase over the course of treatment, as stakeholders engage in more detailed discussions about primary problems and treatment goals (Garland et al., 2004).

A number of additional limitations to the existing evidence base are worth noting. While perspectives have been compared between youth, parents, and clinicians, little is known about diversity in views and perceptions along other lines, such as gender, socio-economic background, or ethnicity. In addition, all existing studies were conducted in high-income countries marked by Western cultural norms. As stressed by Hoagwood and colleagues (1996), change is in itself not meaningful unless it is assigned meaning within a given cultural context. Perceptions of good outcome are shaped by “the norms, values, and social structures of the culture in which the intervention is embedded” (Hoagwood et al., 1996, p. 1057). The existing evidence base is thus limited with regards to the experiences it reflects, both socio-demographically and culturally.

1.6 Influence of the Treatment Approach

While Sections 1.4 and 1.5 have focussed on different stakeholder perspectives on what constitutes a good outcome, another crucial factor is the treatment approach, and the “theory of therapeutic action” underpinning it (D. J. Cohen, 1995). Psychotherapeutic treatments are rooted in different theories about the mechanisms that drive the development of depression (and other mental health difficulties), and about the approaches best suited for deactivating or modifying these mechanisms, to promote adaptation and improvement (Fonagy, 1997). Different approaches may target different outcomes, and make different assumptions about the sequencing or trajectories of change. As suggested by Fonagy (2010), service user’s observations about their symptoms are of central interest in Cognitive Behavioural Therapy (CBT), which focusses on reducing symptoms via cognitive and behavioural strategies (A. T. Beck, 1976, 1993). In psychodynamic psychotherapy, therapists are trained to focus on intermediary outcomes, such as service users achieving transference in projecting feelings onto the therapist, or service users gaining insight into how past events may have influences their current thoughts, emotions, and behaviour (Fonagy, 2010; Odhammar & Carlberg, 2015). It is important to recognise the complexity of outcome measurement in practice, and the fact that different treatment approaches drive the outcomes they are likely to achieve. This is discussed in more detail in Section 3.4.

1.7 Measurement Issues

Beyond the selection of outcome domains and informants, methodological aspects of the measurement, analysis, and reporting of outcomes also have a bearing on how ‘good outcome’ is defined and assessed. Outcomes may be measured using nomothetic (i.e., standardised) measures, or they may be tracked through idiographic (i.e., personalised) tools. Once data has been collected, outcomes may be analysed and reported at the individual level or with reference to improvements in average group scores. Both of these aspects are discussed in turn below.

1.7.1 Nomothetic or Personalised Outcome Measurement?

Nomothetic outcome measures assess specific dimensions of symptoms, functioning, or other outcome concepts that researchers consider relevant to most service users. They locate the service user on these dimensions, and their location can then be compared to that of other service users and to population norms (Sales & Alves, 2016). In the case of symptom measures, this allows for the formulation of diagnoses and assessments of remission or recovery with reference to population norms and reference points that most clinicians are familiar with (Ashworth et al., 2007; Sales et al., 2018). Items of nomothetic measures are generally defined by researchers with varying degrees of input from service user representatives (Beresford & Branfield, 2006), and the validity and reliability of these items is then tested in relevant populations. Nomothetic scales are appreciated for their psychometric properties, and the possibility of comparison and benchmarking. Where normative data is available, they generally provide clinical cut-off points that can be used to establish clinically significant change (see a more detailed discussion of clinically significant change in Section 1.7.2). They further facilitate the comparisons of outcomes across groups of service users (Wolpert, Ford, et al., 2012). However, nomothetic measures have also been criticised for providing one-size-fits all solutions to clinical populations that are typically diverse in their presentations, needs, and priorities (Hurn et al., 2006; Jacob, 2019; Kiresuk & Sherman, 1968; Rockwood et al., 1997; Sales & Alves, 2016).

Idiographic outcome measures aim to remedy some of the shortcomings of nomothetic measures. Rather than imposing a predefined set of items, these measures are tailored to each service user’s primary concerns and priorities, containing only items that are relevant to them (Fitzpatrick et al., 1998). They take the individual, rather than the general population, as their central reference point (Ashworth et al., 2019). Progress with regards to these outcomes is then tracked, often using a standardised assessment scale. Idiographic outcome measures have been in use since the 1960s. They include instruments that track top problems or concerns, such as the Target Complaints (Battle et al., 1966); the Simplified Personal Questionnaire (PQ; Elliott et al., 2016); Youth Top Problems, which was developed specifically for young people (Weisz et al., 2011); and Psychological Outcome Profiles (PSYCHLOPS; Ashworth et al., 2004), which have recently been adapted for use with young people (i.e., PSYCHLOPS Kids; Godfrey et al., 2019). Other idiographic measures focus on goals that service users and their families would like to achieve with the help of treatment. Goal Attainment Scaling (GAS) is the oldest procedure for specifying individual goals (Kiresuk & Sherman, 1968). It involves (a) defining a set of personal goals, (b) specifying possible outcomes for each goal that can be mapped onto a standard attainment scale, and (c) assessing treatment effectiveness by tracking goal attainment using this scale (King et al., 1999). In the UK, a similar tool, the Goal Based Outcomes, has been developed for goal-based outcome measurement in child mental health (Law, 2006).

By assigning service users an active role in defining which success criteria should govern the evaluation of their treatment, idiographic goals promote shared decision making and person-centred care (Grenville & Lyne, 1995; King et al., 1999; Sales & Alves, 2012, 2016). Defining treatment goals and tracking progress toward them can facilitate conversations and collaboration between service users and clinicians (Jacob et al., 2018; Sales & Alves, 2016), and translate therapeutic reasoning into outcome measurement (Lachs, 1993). Idiographic outcome measures can also reduce respondent burden, as service users can be asked fewer questions based on what is relevant to them (Kiresuk & Sherman, 1968).

In clinical practice, the relevance and low burden of idiographic measures promotes their uptake (Sales & Alves, 2016). The practice of goal-setting has been shown to have beneficial effects on treatment process and experience: Idiographic outcome measurement had high levels of acceptability in a sample of youth accessing school-based mental health support who reported that personalised outcome measurement helped improve their self-awareness, problem-solving, and the achievement of behavioural goals (Duong et al., 2016). In a study examining treatment experiences, discharged adolescent inpatients reported that goal-setting had been one of the most meaningful aspects of their treatment experience (Grossoehme & Gerbetz, 2004). A recent retrospective chart review at two Australian *headspace* treatment centres for youth indicated that idiographic goal setting at assessment was associated with reduced risk of disengagement and more sessions attended (Cairns et al., 2019).

Idiographic outcome measures also allow covering a broader range of outcomes than nomothetic scales. For example, an exercise of mapping young people’s self-defined goals onto commonly used nomothetic measures found that outcomes around talking about feelings and thoughts, letting people know when help is needed, understanding anger, being more independent and responsible for oneself, feeling more confident, thinking about oneself and understanding one’s past were not covered by nomothetic measures frequently used in UK CAMH services (Jacob et al., 2017). This suggests that by combining nomothetic and idiographic measurement in clinical practice, the breadth of outcomes considered can be expanded, and assessment made more person-centred (Sales & Alves, 2016).

At the same time, idiographic measurement poses a number of challenges. There is limited evidence on the validity and reliability of idiographic measures, and how these relate to nomothetic measures (Jacob et al., 2018; Sales & Alves, 2016). Another challenge is that when used in clinical practice, clinicians may prioritise achievable outcomes, thereby raising their chances of demonstrating success, even though this may not reflect what matters most to the service user, or may not reflect a sufficiently ambitious goal to assess quality of care (Kiresuk & Sherman, 1968; Rockwood et al., 1997). A further disadvantage is that by nature of their idiosyncrasy, personalised outcome scores cannot easily be compared across service users or research studies, even when standardised scales are used to score change in personalised items. They may thus be particularly valuable tools for outcome measurement in clinical practice, with a focus on individual cases.

1.7.2 Measuring Outcome at Group or Individual Level?

Historically, clinical research has analysed outcome data in terms of differences in mean scores observed in a group of service users, before and after treatment, or when comparing an intervention and a control group (A. B. Hill, 1937). When converted into a standardised effect size (e.g. J. Cohen, 1988), the magnitude of this difference can be interpreted and compared across studies and groups. However, effect sizes cannot be readily interpreted in terms of how many service users have experienced change that is clinically meaningful, or have returned to previous levels of functioning. A change that is statistically significant is not necessarily clinically significant and vice versa (Wolpert, Görzig, et al., 2015). Large improvements for some individuals can provide a skewed impression of overall effectiveness (Edbrooke-Childs et al., 2018; S. A. Jensen & Corralejo, 2017). This makes it difficult to draw meaningful conclusions about the actual effectiveness of treatment, as the effect size may conceal substantial variation in treatment impact within the sample (S. A. Jensen & Corralejo, 2017; Westen & Bradley, 2005).

As argued by Lambert (2013), “a key element in psychotherapy quality management research is defining and operationalizing the concepts of positive and negative outcome for the *individual* [emphasis added] patient” (p.46). To establish whether change is not only statistically significant but also clinically meaningful, Jacobson and Truax (Jacobson, Follette, & Revenstorf, 1984; Jacobson & Truax, 1991) developed a methodology for classifying outcome based on two elements: a so-called Reliable Change Index (RCI), and normative data enabling the determination of a clinical cut-off (McGlinchey et al., 2008). The RCI serves to assert whether individuals demonstrate greater change in scores than could be attributed to measurement error alone. The *clinical significance* of this change is assessed based on whether or not the individual demonstrates reliable movement from the clinical scoring range to the normative scoring range, thus demonstrating normative levels of functioning or distress. Based on these two elements, an individual may be classified as follows (McGlinchey et al., 2008):

- **“recovered”:** the individual demonstrates reliable and clinically significant change;

- **“improved:** the individual demonstrates reliable change, but still scores within the clinical range;
- **“unchanged”:** the individual demonstrates neither reliable nor clinically significant change; or
- **“deteriorated”:** the individual demonstrates reliable change indicating a worsening rather than a movement towards normative levels.

This methodology serves to classify outcome across presenting problems and measurement scales (Evans et al., 1998). It has been recommended for use alongside the analysis of effect sizes in clinical research, as well as for the routine measurement of outcomes in clinical practice (Evans et al., 1998; Jacobson & Truax, 1991b; S. A. Jensen & Corralejo, 2017; Wolpert, 2017; Wolpert, Goerzig, et al., 2015). The consensus definitions of outcome for depression defined by Frank and colleagues (1991) – that is, *response*, *remission*, and *recovery* – build upon this notion of clinically meaningful change, by combining clinical significance and reliable change with consideration of the timelines over which these changes occur and are sustained (McGlinchey et al., 2008). The need for a clinical cut-off implies that clinically significant change can only be meaningfully specified within outcome domains for which population norms exist, such as symptoms and functioning. However, the reliable change metric can be applied across outcome domains.

1.8 Concluding Remarks

This chapter has provided a discussion of outcome concepts; a critical review of the existing empirical literature on outcome perceptions amongst youth, parents, and clinicians; and an introduction to aspects of outcome measurement that will be examined in more detail as part of this doctoral thesis. It has been shown that symptom-focussed approaches to measurement are often favoured for research purposes, as they are seen to provide clear and comparable indicators of treatment success. Such approaches have, however, been challenged for being too unidimensional, for providing ambiguous results that are not inherently interpretable, and for failing to reflect the broad spectrum of changes experienced by service users during the process of recovery. Indeed, the existing empirical literature suggests that young people, parents, and clinicians value a broad range of other changes, beyond symptoms, which are inadequately captured by purely symptom-focussed measurement. Similarly, symptom-focussed measurement may fail to generate evidence on intermediate outcomes and mechanisms through which treatment is expected to work.

Outcome measurement that is too narrow has real-world implications in clinical practice settings, for example where service users feel they have benefitted while this is not recorded. Service users may achieve personal treatment goals while not achieving clinically significant improvement on a symptom measure, which would possibly be classified as a treatment failure (Kazdin, 1999b). Indeed, a recent study of drop-out amongst adolescents participating in a treatment trial for depression found that some young people dropped out within three sessions after starting treatment, because they felt they had achieved their personal treatment goals. This was, however, not reflected by nomothetic outcome measures and would be flagged as treatment failure (O’Keeffe et al., 2019). Issues of convergence or divergence in outcomes measured across different outcome domains are the focus of the last study included in this thesis (Comparing Change Across Outcome Domains study, Chapter 7), which explores these in relation to functioning and idiographic goals, specifically for depressed adolescents.

The evidence base is limited by a frequent lack of theoretical grounding and by inconsistencies in how different outcomes have been categorised, which hampers comparisons between studies. In addition, most studies have been broad in scope, investigating goals or outcome perceptions amongst children and young people with a range of presenting problems, thus potentially masking more specific priorities held by depressed adolescents. Lastly, heterogeneity in perceptions and priorities *within* groups (i.e. amongst young people, parents, and clinicians) were rarely explored. This doctoral thesis aims to address these limitations by devising an initial taxonomy of treatment outcomes, which can serve as a consistent conceptual framework, and by empirically examining outcome perceptions amongst young people, parents, and clinicians specifically for adolescent depression, and with special attention to heterogeneity in viewpoints.

Chapter 2. Research Design and Methodology

2.1 Research Paradigm

2.1.1 Overall Research Design

This doctoral thesis represents a pragmatist approach to research (Howe, 1988; Johnson & Onwuegbuzie, 2004; Maxcy, 2003; Morgan, 2007; Tashakkori & Teddlie, 1998). It has an applied focus, as it aimed to generate knowledge that could help promote principles of value-based and person-centred mental health care for young people, in practice. As outlined in Figure 2.1, this thesis was structured around a set of complementary research questions, which are best answered using a mix of different qualitative and quantitative methods. The following sections briefly introduce pragmatism as a research paradigm, outline the mixed-methods approach employed, and provide a brief description of the research methods used.

A Brief Introduction to Pragmatism as a Research Paradigm

Research paradigms are systems of beliefs and practices that influence how researchers select the questions they study and the methods used to study them (Morgan, 2007). Research paradigms are characterised by the specific assumptions they make about the nature of reality (i.e., ontology) and of knowledge acquisition (i.e., epistemology). The most commonly used paradigms are positivism (or post-positivism) and constructivism. The positivist paradigm postulates that there is a single reality that can be studied objectively and directly, through measurement and rigorous hypothesis testing. The aim is to derive general conclusions that are valid beyond the specific research context, and unbiased by the role and perspective of the researcher. Post-positivism acknowledges that reality can never be observed directly, but only in an approximate fashion. In contrast, the constructivist paradigm postulates that there are multiple realities, which are constructed through social interactions. Knowledge is not obtained but created through the interaction of researchers and research participants, in a process that is conditioned by their respective values, perspectives, and environments (Howe, 1988; Todd et al., 2004). The role of the researcher in co-constructing knowledge is explicitly acknowledged. Knowledge is generated through interpretation rather than measurement, to gain a deeper understanding of human behaviour, attitudes, and experiences from the point of view of those experiencing them. Positivism and constructivism are commonly associated with quantitative and qualitative methodologies, respectively. As part of the so-called paradigm wars, proponents of each paradigm clashed over the respective merits of each approach, based on the sense that their underlying assumptions were incompatible (for discussions of the paradigm wars see Gage, 1989; Hammersley, 1992; Oakley, 1999).

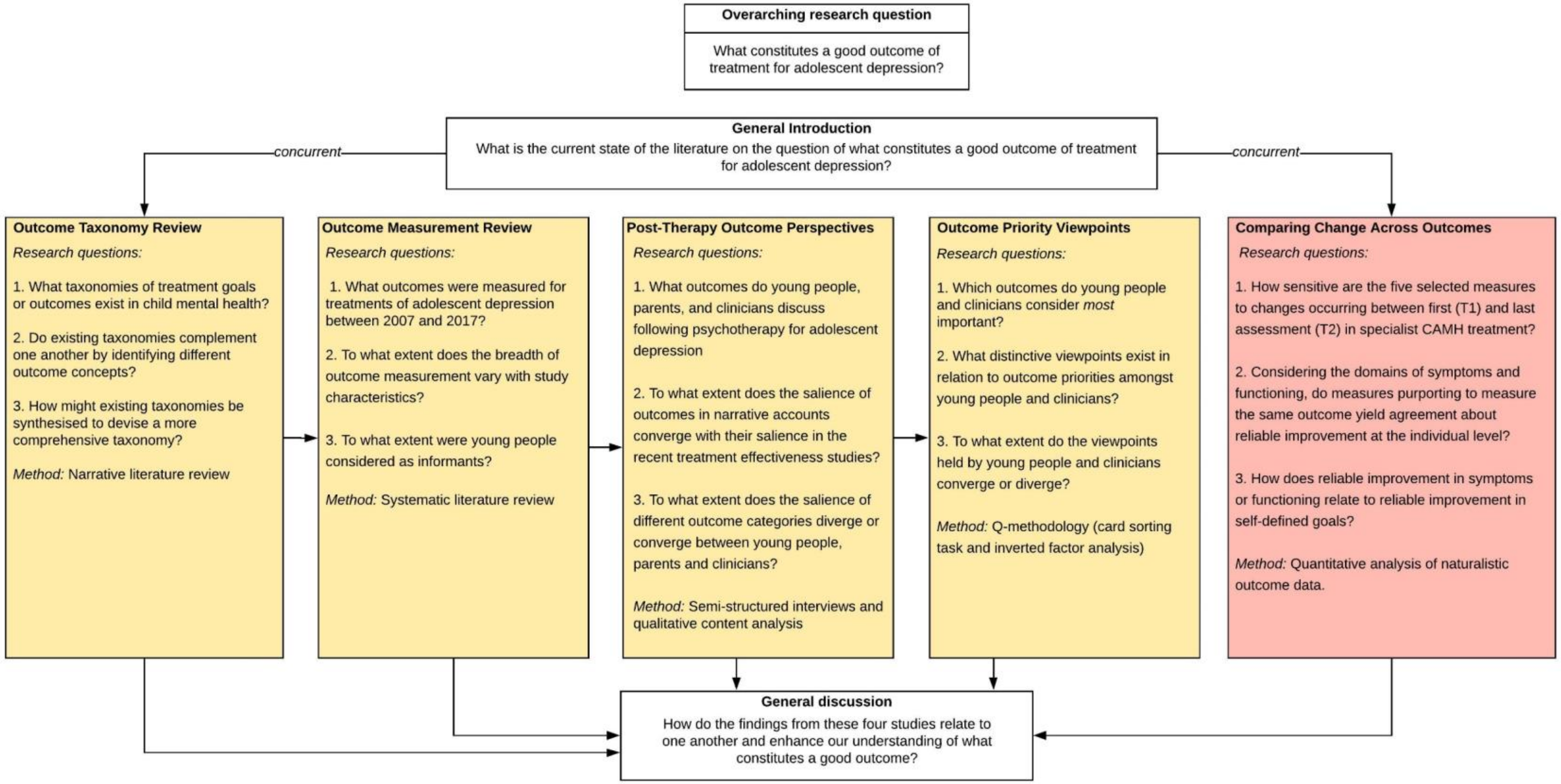
A third paradigm, that of pragmatism, refutes this notion of incompatibility by taking a philosophical and methodological “middle position” (Johnson & Onwuegbuzie, 2004, p. 17). Tashakkori and Teddlie (2003) define pragmatism as:

> a deconstructive paradigm that debunks concepts such as ‘truth’ and ‘reality’ and focuses instead on ‘what works’ as the truth regarding the research questions under investigation. Pragmatism rejects the either/or choices associated with paradigm wars, advocates for the use of mixed methods in research, and acknowledges that the values of the researcher play a large role in interpretation of results. (p. 713)

Within a pragmatist paradigm, this doctoral thesis combines constructivist and positivist elements. It is constructivist in that it considers the meaning of ‘good outcome’ to be contested and socially constructed. This premise motivated the exploration of outcome concepts and priorities from the perspective of different stakeholder groups, which forms the core of this doctoral thesis. At the same time, conclusions about predominant views were drawn with reference to the frequency at which specific outcomes were discussed or measured, following more positivist reasoning. While the thesis emphasises differences in perceptions and priorities, it ultimately aims to facilitate dialogue between different stakeholder groups, and to strengthen outcome measurement as a positivist practice.

This doctoral thesis also has elements of the transformative research paradigm, which posits that power relations are a central issue in research and must be considered and addressed at each stage of the process (Mertens, 2007, 2010). It reflects a transformative stance by recognising that young people accessing services often have less power of interpretation, compared with parents and clinicians, in defining ‘good outcome’. In addition, they often have limited influence on defining measurement approaches (Merry et al., 2004; Moran et al., 2012; Mulley et al., 2017). The transformative ambition is reflected in the choice of research methods for this thesis, and in placing young people at the centre of this inquiry: The Post-Therapy Outcome Perspectives study (Chapter 5) drew on semi-structured interview data to explore how young people describe and discuss outcomes in their own words; the study of Viewpoints on Outcome Priorities (Chapter 6) used Q-methodology, consisting of a card sorting task and inverted factor analysis, to elicit outcome priorities directly from the item configurations produced by young people. Q-methodology has been explicitly praised for its potential to empower participants and make marginalised voices heard (Brown, 2006; Donner, 2001). Finally, the Outcome Measurement Review (Chapter 4) gave special attention to the extent to which youth have been considered as informants by recent treatment efficacy and effectiveness studies.

Figure 2.1. Mixed Methods Research Design and Research Questions



Overarching research question

What constitutes a good outcome of treatment for adolescent depression?

General Introduction

What is the current state of the literature on the question of what constitutes a good outcome of treatment for adolescent depression?

Outcome Taxonomy Review

Research questions:

1. What taxonomies of treatment goals or outcomes exist in child mental health?

2. Do existing taxonomies complement one another by identifying different outcome concepts?

3. How might existing taxonomies be synthesised to devise a more comprehensive taxonomy?

Method: Narrative literature review

Outcome Measurement Review

Research questions:

1. What outcomes were measured for treatments of adolescent depression between 2007 and 2017?

2. To what extent does the breadth of outcome measurement vary with study characteristics?

3. To what extent were young people considered as informants?

Method: Systematic literature review

Post-Therapy Outcome Perspectives

Research questions:

1. What outcomes do young people, parents, and clinicians discuss following psychotherapy for adolescent depression

2. To what extent does the salience of outcomes in narrative accounts converge with their salience in the recent treatment effectiveness studies?

3. To what extent does the salience of different outcome categories diverge or converge between young people, parents and clinicians?

Method: Semi-structured interviews and qualitative content analysis

Outcome Priority Viewpoints

Research questions:

1. Which outcomes do young people and clinicians consider *most* important?

2. What distinctive viewpoints exist in relation to outcome priorities amongst young people and clinicians?

3. To what extent do the viewpoints held by young people and clinicians converge or diverge?

Method: Q-methodology (card sorting task and inverted factor analysis)

Comparing Change Across Outcomes

Research questions:

1. How sensitive are the five selected measures to changes occurring between first (T1) and last assessment (T2) in specialist CAMH treatment?

2. Considering the domains of symptoms and functioning, do measures purporting to measure the same outcome yield agreement about reliable improvement at the individual level?

3. How does reliable improvement in symptoms or functioning relate to reliable improvement in self-defined goals?

Method: Quantitative analysis of naturalistic outcome data.

General discussion

How do the findings from these four studies relate to one another and enhance our understanding of what constitutes a good outcome?

Note. Studies shaded in yellow were conducted sequentially, that is one method was used to elaborate on the findings from a preceding study that used a different method; the Comparing Change Across Outcome Domains study is shaded in orange because it was conducted concurrently, to complement the other four studies (Creswell, 2003).

2.2 Characteristics of the Mixed-Methods Approach Used

Within a pragmatist research paradigm, this thesis uses a mixed-methods research design. Mixed-methods research is characterised by the combination of both qualitative and quantitative approaches for data collection, analysis, and the extraction of findings and conclusions within a single study or research project (Tashakkori & Creswell, 2007b, p. 4). Pragmatism promotes “methodological eclecticism" (Hammersley, 1996) that understands qualitative and quantitative methods as opposite ends of a continuum along which researchers can move freely to answer different types of research questions in the best possible way (Newman et al., 2003). As such, it lends itself well to applied research that is designed to address practical rather than theoretical problems (Hammersley, 1996; Pope & Mays, 1995). Health outcomes research in particular has been identified as an area of application for mixed methods, with the goal of promoting the priorities of service users, as well as scientific rigour, with Pope and Mays (1995) suggesting: "We need a range of methods at our fingertips if we are to understand the complexities of modern health care (p.45).

This doctoral thesis consists of five studies, which combine quantitative and qualitative elements (see Section 2.3 for an outline of the methods used in each study). These include narrative literature review, systematic literature review, qualitative content analysis of semi-structured interviews, Q-methodology, and quantitative analysis of a large naturalistic outcomes data set. These methods complement one another by illuminating different aspect of the overarching research question. The predominantly qualitative approaches employed by the studies of Post-Therapy Outcome Perspectives (Chapter 5) and Viewpoints on Outcome Priorities (Chapter 6) shed light on how outcomes are constructed, experienced, and prioritised by young people, parents, and clinicians (McLeod, 2001, 2011). Q-methodology in particular is tailored to the analysis of subjectivity (S. R. Brown, 1993), with a focus on identifying and interpreting subtle differences in perceptions and priorities that may be missed by large-scale quantitative studies. In turn, the quantitative approach employed in the Comparing Change Across Outcome Domains study (Chapter 7) enables general inferences about the levels of convergence between commonly used change metrics, and thus helps advance knowledge about important methodological issues inherent in outcome measurement.

In a typology of mixed-methods research designs, Creswell (1994, 2003) distinguishes sequential and concurrent mixed-methods designs: In sequential designs each new method helps elaborate on the findings generated through previously used methods; in concurrent designs, different methods and types of data are used in parallel, and combined to achieve more comprehensive findings. (Franz 2013). The present doctoral thesis used a hybrid mixed-methods design, which combined sequential and concurrent strands of research. The Outcome Taxonomy Review, Outcome Measurement Review, Post-Therapy Outcome Perspectives, and Viewpoints on Outcome Priorities were conducted sequentially, with the initial taxonomy developed through the Outcome Taxonomy Review serving as a consistent conceptual framework that was iteratively strengthened based on the insights from each subsequent study. The Comparing Change Across Outcome Domains study formed a separate, concurrent strand of research that was conducted in parallel. The general discussion provided in Chapter 8 integrates the findings from all five studies to derive overarching conclusions (Tashakkori & Creswell, 2007). Figure 2.1 (above) provides an illustration of this mixed-methods research design.

Quantitative and qualitative methods are also integrated within studies, at the level of research questions, sampling designs, data collection, and analysis (Hammersley, 1996). For example, the study of Post-Therapy Outcome Perspectives (Chapter 5) draws on rich qualitative data that was collected through semi-structured interviews. These were conducted as part of a randomised control trial, and all youth recruited into the trial within a specific geographic area were invited to be interviewed. The sampling approach was thus not guided by considerations typical for qualitative research, such as the purposive representation of different perspectives or the principle of saturation[1] (e.g., Dworkin, 2012; Mason, 2010). The analytical method chosen by the doctoral candidate, qualitative content analysis, further combined elements of qualitative and quantitative analysis, by fusing aspects of thematic analysis with the quantification that characterises content analysis. The Viewpoints on Outcome Priorities study (Chapter 6) provides another example of mixing methods within a study, as Q-methodology combines quantitative analysis using factor analytical techniques with in-depth qualitative interpretation of data collected from small, purposively selected samples (Baker et al., 2006; Ellingsen et al., 2010).

2.3 Specific Methods Used

2.3.1 Outcome Taxonomy Review

An integrative narrative review (Torraco, 2005; Whittemore & Knafl, 2005) was conducted to identify existing outcome taxonomies in child mental health; critically appraise them; and synthesise their content into a new, more comprehensive taxonomy. While systematic reviews focus on producing replicable findings by following explicit and formalised methodological protocols (e.g., PRISMA guidelines for defining preferred reporting items for systematic reviews and meta-analyses; Moher et al., 2009), narrative reviews typically use less formalised approaches. Instead they focus on providing in-depth, critical, and reflective appraisals of the literature – often for the purpose of theory development (Greenhalgh, Thorne, & Malterud, 2018). Through the narrative review conducted in this study, a new taxonomy was obtained, which was subsequently revised and reviewed in light of the evidence emerging from other studies conducted as part of this doctoral project.

---

[1] Saturation describes the point at which the data collection process ceases to produce new insights or perspectives (Dworkin, 2012).

2.3.2 Outcome Measurement Review

A systematic review was conducted (Chapter 4) to explore what outcomes were reported in the recent treatment efficacy and effectiveness literature for adolescent depression. Academic databases were systematically searched for peer-reviewed studies published between 2007 and 2017 that evaluated treatment outcomes either quantitatively or qualitatively. Outcome concepts and measurement tools were systematically extracted and categorised using the initial taxonomy, derived through the Outcome Taxonomy Review. Descriptive statistical analysis and hypothesis tests were performed on studies presenting quantitative outcome measures to explore differences in the breadth of outcome measurement in relation to study characteristics. The Outcome Measurement Review represented the second of four sequentially conducted studies exploring the types of outcomes valued by different stakeholder groups and in different contexts, focussing specifically on the outcome concepts that currently predominate clinical research.

2.3.3 Post-Therapy Outcome Perspectives

The Post-Therapy Outcome Perspectives study (Chapter 5) drew on semi-structured interviews conducted with triads of young people, parents, and clinicians following their participation in a treatment trial for adolescent depression (Midgley et al., 2014). Qualitative content analysis (Elo & Kyngäs, 2008; Hsieh & Shannon, 2005) was used to explore what outcomes young people, parents, and clinicians described; how outcome reports differed between (a) participant groups, and (b) the three treatment modalities covered by the trial arms; and how they corresponded to the outcomes reported in the treatment effectiveness literature (as per the Outcome Measurement Review). As part of the qualitative content analysis, narrative segments describing change were categorised, using the taxonomy developed through the Outcome Taxonomy Review as an a priori coding framework. Changes to the initial taxonomy were then made based on outcome themes emerging from the interview data.

2.3.4 Viewpoints on Outcome Priorities

The Viewpoints on Outcome Priorities study (Chapter 6) built upon the above-mentioned studies by using the refined outcome taxonomy as a starting point. It explored which outcomes young people and clinicians' value the most and why, whether there was diversity in priorities within both groups, and how priorities differed between these two groups. The study used Q-methodology to answer these research questions. In Q-methodology, participants sort a set of stimuli representing different perspectives on a given phenomenon, thus articulating their own point of view. Inverted factor analysis is then used to identify a smaller set of distinctive viewpoints from the individual item configurations produced by all participants. As such, Q-methodology combines the in-depth examination of subjectivity provided by qualitative research methods with the added transparency and structure provided by quantitative analysis even in small samples (Baker et al., 2006).

2.3.5 Comparing Change Across Outcome Domains

The Comparing Change Across Outcome Domains study (Chapter 7) involved quantitative analysis of a large dataset of routinely collected outcome data from CAMHS in England. The study assessed the convergence of reliable change ratings across five commonly used measures of symptoms, functioning, and progress towards self-defined goals. Agreement between different measures and different outcome domains in signalling reliable change was tested using McNemar’s test of correlated proportions (McNemar, 1947), and Cohen’s kappa (κ) of chance-corrected agreement (J. Cohen, 1960).

2.3.6 Methodological Limitations

A number of methodological limitations should be noted. First, while the Outcome Taxonomy Review and the Outcome Measurement Review considered literature from around the world (as long as it was published in English, German, French, Spanish or Italian), the studies focussing on Post-Therapy Outcome Perspectives, Viewpoints on Outcome Priorities, and Comparing Change Across Outcome Domains used data collected exclusively in England. Since outcome perceptions and priorities are likely to differ across cultural contexts (Binder et al., 2010; Fonagy, 1997; Hoagwood et al., 1996), inferences can therefore not be made for other cultural contexts. This is an important area for future research.

Second, the studies of Post-Therapy Outcome Perspectives and Viewpoints on Outcome Priorities used predominantly qualitative research methodologies. The reliability and validity of qualitative findings cannot be assessed with the same criteria used for evaluating quantitative studies (Yardley, 2000). The interpretative component of qualitative analysis implies that findings are inherently subjective, as the proposed interpretation represents only one possible view on the phenomenon under study. Such views are likely to be influenced by the researcher’s training, pre-existing assumptions, and expectations. Reflecting on possible sources of bias, and making these transparent is a characteristic of good qualitative research (Yardley, 2000). Therefore, Chapter 5 and Chapter 6 include dedicated reflexivity sections.

Lastly, while not exactly a methodological limitation, this doctoral thesis approached the topic of outcomes through a deliberately narrow conceptual lens. It focused on outcomes in relation to depression, reflecting calls for exploring the question of ‘good outcome’ within a problem-specific framework (Porter, 2010). At the same time, evidence of a general psychopathology factor in adolescence suggests that there may be common underlying factors driving both emotional and behavioural problems (Patalay et al., 2015). This might imply that outcome measurement should be tailored to capturing changes in these common factors, beyond specific presenting problems. The aim of this doctoral thesis was to identify outcomes that are of particular relevance to young people with depression, as there is currently a lack of clarity in this area, coupled with growing attention to measuring outcomes for this specific population. The taxonomy developed through this thesis can inform broader cross-diagnostic studies, by ensuring that outcomes that are important to youth with depression are considered.

Chapter 3. Outcome Taxonomy Review

3.1 Introduction

3.1.1 Background

As the discussion of the existing empirical literature on outcome perceptions in child mental health has demonstrated, ambiguity and variation in how outcomes are named and categorised across different studies complicate the synthesis of findings. Several studies have failed to disclose the coding frame through which they approached qualitative data analysis, thus providing limited conceptual transparency. A taxonomy of treatment outcomes in child mental health could help address such issues. A taxonomy is the hierarchical classification of phenomena into categories and sub-categories, for the purpose of establishing a standardised terminology (American Society for Indexing, 2019). In health research, referential taxonomies include the International Classification of Functioning, Disability and Health (ICF), which provides a framework for the description of health states (WHO, 2008); and standard diagnostic systems like the International Classification of Diseases for Mortality and Morbidity Statistics, 11th Revision (ICD-11; WHO, 2018); and the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5; APA, 2013).

Taxonomies can promote comparability by providing a common frame of reference and by “canvassing the full range” of concepts that may be relevant to the examination of a specific phenomenon (Michie et al., 2011, p. 2). A taxonomy of outcomes would enable the systematic synthesis of findings from existing qualitative and quantitative studies. It would also enable greater conceptual transparency and terminological consistency in future qualitative studies. Even where such studies develop categories inductively based on themes emerging from the raw data rather than by applying a pre-existing framework, the categories identified could still be referred back to a more universal taxonomy, thus promoting the integration of the qualitative evidence base. In the longer term, a taxonomy of outcomes in child mental health could also facilitate systematic literature searches focussed on specific outcomes of interest, and provide conceptual guidance to those interested in the development of core outcome sets (Dodd et al., 2018)

A taxonomy of treatment outcomes for the whole of health research has been proposed by Dodd and colleagues (2018) with the aim of reducing inconsistencies in how outcomes are described across clinical trials, systematic reviews, and clinical registries. Their proposed taxonomy has a broad coverage of outcomes relating to various physical health conditions, but identifies only one overarching domain of ‘psychiatric outcomes’ without further disaggregation. This lack of granularity, as well as the lack of a developmental perspective, limits its utility for use in child mental health. Partly recognising this limitations, Dodd and colleagues have encouraged mental health researchers to advance further disaggregation of the ‘psychiatric outcomes’ domain, by using a standard diagnostic system as a guiding framework. This would, however, seem to focus the psychiatric outcome domain on types of disorder, rather than types of outcome. Finally, the taxonomy proposed by Dodd and colleagues (2018) spans not only positive individual-level outcomes, but also negative effects, as well as service- and system-level outcomes. These were not the focus of the present study, which aimed to identify person-centred concepts of ‘good outcome’. There is thus value in considering whether other taxonomies exist that are more closely tailored to child mental health, and in appraising their capacity to serve as a comprehensive conceptual framework for this doctoral thesis.

3.1.2 The Present Study

This narrative literature review aimed to identify existing taxonomies of treatment outcome or treatment goals in child mental health, to compare their content and relative merits, and to synthesise them into a more refined taxonomy as necessary. This initial taxonomy would then be revised iteratively, as appropriate, based on the findings emerging from the remaining four studies conducted as part of this doctoral thesis. This narrative review was guided by the following research questions:

1. What taxonomies of treatment outcomes or goals exist in child mental health?
2. How might existing taxonomies be synthesised into a new, more comprehensive taxonomy?

3.2 Method

3.2.1 A Note on Integrative Narrative Review

This study used the method of integrative narrative review to identify, critically appraise, and synthesise existing outcome taxonomies in child mental health. Integrative literature reviews are “a form of research that reviews, critiques, and synthesizes representative literature on a topic in an integrated way such that new frameworks and perspectives on the topic are generated” (Torraco, 2005, p. 356). Alongside research agendas and metatheories, taxonomies are a typical research output of integrative reviews (Torraco, 2005).

Narrative reviews differ from systematic reviews in a number of ways. They are generally more inclusive of different study designs, and more focussed on identifying and discussing variations in how a phenomenon has been approached in the literature (Whittemore & Knafl, 2005). For this purpose, evidence is often selected purposively rather than through an extensive and exhaustive screening process. While not offering the same level of replicability and methodological rigour as systematic reviews, narrative reviews can provide more in-depth, critical, and reflective appraisals of the literature, and thus fulfil a complementary purpose (Greenhalgh et al., 2018).

3.2.2 Review, Appraisal, and Synthesis of the Literature

For the present study, relevant literature was purposively and gradually identified over a three-year period, from October 2016 to October 2018. The following channels were used:

- A highly specific systematic literature search was conducted in the PsychInfo database to identify studies of outcome or goal taxonomies in child mental health. The search syntax used for this search is displayed in Table 3.1.
- Over 7500 studies relating to outcome research in child mental health were screened as part of the systematic Outcome Measurement Review presented in Chapter 4 (see Table B.1 in Appendix B for the full search syntax). Any relevant literature identified in the process was earmarked for consideration in this narrative review.
- Reference lists of relevant studies were routinely searched.
- Relevant studies were equally shared by collaborators, who are experts in the field.

Studies were included in this narrative review if they explicitly developed an outcome taxonomy or framework for child mental health, based on theoretical considerations or based on the analysis of empirical data. They had to be written in English, German, French, Spanish, or Italian to be considered. Because of the nature of the research questions and the integrative narrative review approach, the PRISMA guidelines (Moher et al., 2009) defining preferred reporting items for systematic reviews and meta-analyses could not be applied to this review.

For the purpose of synthesis and comparison, each of the outcome categories included in the identified taxonomies were extracted and tabulated (Michie et al., 2011). Each outcome category was appraised for its relevance to a person-focussed examination of outcomes for adolescent depression. If relevant, they were synthesised into a new, more comprehensive taxonomy. The conceptual models suggested by Hoagwood and colleagues (1996) and Fonagy (1997), while not providing nuanced taxonomic structures, were used to guide the theoretically informed classification of outcomes into higher-level outcome domains. The resulting synthesis is displayed in Table 3.2. See Table A.1 (Appendix A) for an overview of each individual framework and its original terminology.

Table 3.1. Automated Search Syntax for the Outcome Taxonomy Review

Line	Syntax
1	(outcome* or goal*).ti.
2	(taxonomy or framework or model).ti, ab.
3	(adolescen* or youth or child* or young people).ti.
4	(mental health or depress* or low mood or mood disorder or emotional problem* or emotional difficult* or internali#ing problem* or internali#ing difficult*).ti.
5	(therap* or psychiatr* or counselling or treatment).ti.
6	1 and 2 and 3
7	4 or 5
8	6 and 7

3.3 Findings

3.3.1 What Outcome Taxonomies Already Exist in Child Mental Health?

The integrative narrative review identified five existing goal or outcome frameworks for child mental health. This included the two conceptual models proposed by Hoagwood and colleagues (1996) and Fonagy (1997), which were considered for the identification of higher-level outcome domains. It further identified one goal taxonomy for adult mental health (Grosse Holtforth & Grawe, 2002) that was considered relevant because it formed the conceptual reference point for two of the child-focused taxonomies.

A taxonomy of adult treatment goals was developed by Grosse Holtforth and Grawe (2002) in Switzerland. The so-called Bern Inventory of Treatment Goals was derived from a database of 1031 goals defined by 298 adult service users at a psychotherapy outpatient clinic over a 20-year period. The Bern Inventory comprises five high-level goal domains: *coping with specific problems or symptoms*; *interpersonal goals*; *wellbeing and functioning*; *existential issues*; and *personal growth*. It is further disaggregated into 26 goal categories and 52 sub-categories. The Bern Inventory has demonstrated good inter-rater reliability, exhaustivity, and partial association with diagnostic status in adult outpatients (Grosse Holtforth & Grawe, 2002) and inpatients (Grosse Holtforth et al., 2004).

Two studies conducted in the UK have used the Bern Inventory as a starting point to categorise treatment goals defined by children and young people (Bradley et al., 2013; Rupani et al., 2014; see Chapter 1, Section 1.2.2 for a more detailed discussion of the study methodologies). Bradley and colleagues (2013) thematically analysed a set of treatment goals defined by 80 children and adolescents accessing CAMH services and identified three high-level outcome domains: (a) *coping with specific problems and symptoms*, (b) *relationship/interpersonal*, and (c) *personal growth & functioning*. These were further disaggregated into 25 goal categories. Rupani and colleagues (2014) analysed goals defined by 73 secondary-school students who had participated in school-based counselling (McArthur et al., 2013; Pybis et al., 2015). They identified four high-level outcome domains: (a) *emotional goals*, (b) *interpersonal goals*, (c) *goals targeting specific issues*, and (d) *personal growth goals*. These were further disaggregated into 16 specific goal categories.

Table 3.2. Synthesis of Existing Outome Taxonomies in Child[1] Mental Health

Outcome domains and categories	Grosse Holtforth & Grawe (2002)[1]	Childs et al. (2012)	Bradley et al. (2013)	Rupani et al. (2014)
	Adults	Children and young people		
Symptoms				
Depressive symptoms	✓	*	✓	✓
Suicidality	✓	*	—	✓
Self-harm	✓	*	✓	✓
Anxiety	✓	*	✓	✓
Behaviour and anger	✓	*	✓	✓
Sleep	✓	*	✓	✓
Eating	✓	—	—	—
Somatic problems	✓	—	—	—
Obsessions and compulsions	✓	—	—	—
Sexuality	✓	—	—	—
Substance use	✓	—	—	—
Self-management				
Coping with mood, thoughts, and feelings	✓	—	✓	✓
Understanding feelings and thoughts	—	—	✓	—
Functioning				
Academic functioning	✓	—	✓	✓
Executive functioning	✓	—	—	—
Global functioning	✓	✓	—	—
Social functioning	✓	—	—	—
Personal growth				
Autonomy and responsibility	✓	—	✓	—
Assertiveness	✓	—	—	—
Empowerment and self-efficacy	—	✓	—	—
Meaning and purpose	✓	—	—	—
Processing past & present	✓	—	✓	✓
Recognising ad fulfilling desires & wishes	✓	—	—	—
Self-concept and identity	✓	—	—	—
Self-confidence, esteem, and acceptance	✓	—	✓	✓
Relationships				
Being able to talk about feelings and thoughts	—	—	✓	✓
Family functioning and relations	*	*	✓	✓
Friendships	—	*	✓	✓
Peer relationships (incl. bullying)	*	*	✓	✓
Romantic relationships	✓	—	—	—
Connectedness & intimacy	✓	—	—	—

Outcome domains and categories	Grosse Holtforth & Grawe (2002) [1]	Childs et al. (2012)	Bradley et al. (2013)	Rupani et al. (2014)
	Adults	Children and young people		
Wellbeing				
Exercise and activity	✓	—	—	—
Enjoying life	—	—	✓	—
Loneliness & grief	✓	—	—	—
Relaxation and composure	✓	—	—	—
Wellbeing	✓	—	—	—
Physical health	—	—	—	✓
Parental support and wellbeing				
Parental wellbeing	—	✓	—	—
Parental support / parenting	✓	✓	—	—

Note. ✓ indicates that an outcome or goal concept is explicitly identified in the relevant taxonomy. ∗ indicates that an outcome or goal concept is implicitly covered by the relevant taxonomy, for example as part of a higher-level category. — indicates that an outcome or goal has not been included by the relevant taxonomy.

[1] The Bern Inventory developed by Grosse Holtforth and Grawe (2002) was developed based on goals defined by adults. However, it has been adapted by two child-focussed studies, thus demonstrating considerable relevance for child mental health. It has therefore been included alongside child-focussed frameworks.

A different approach was taken by Childs and colleagues (2013) who led a Delphi-consultation (Turoff, 1971) with adolescent service users, clinicians, researchers and commissioners to explore notions of ‘good outcome’. Outcomes were discussed in relation to different age groups and scenarios (e.g., where improvement was not likely), and a taxonomy was devised that comprises three higher-level domains: the level of the individual child, the interpersonal level, and the service/society level (which was not been considered for this review). These were further disaggregated into six sub-domains, which cover many of the outcomes suggested by Hoagwood and colleagues (1996) and Fonagy (1997) in their conceptual models: *symptom change*, *functioning*, *interpersonal relationships*, *service user experience* (not considered here), *parental wellbeing,* and *system-level outcomes*. The consultation group stressed that for application, the proposed sub-dimensions would need to be tailored to the relevant presenting problem, developmental status, and other potentially relevant characteristics (e.g., ethnicity or socio-cultural background) of the target population.

3.3.2 How Might Existing Taxonomies Be Synthesised?

The four taxonomies had a number of similarities. First, all distinguished an individual outcome dimension from a relational dimension, in line with conceptual outcome models (Fonagy, 1997; Hoagwood et al., 1996). Second, all identified symptom change as a key individual outcome, although with varying degrees of nuance in the disaggregation of specific symptoms. Third, all identified outcomes related to functioning, although these were classified and specified differently: Bradley and colleagues (2013) identified a category of academic functioning within the domain of personal growth; Childs and colleagues (2013) specified functioning as a higher-level domain at the individual level; and Grosse Holtforth and Grawe (2002) did not use the term functioning, but specified *difficulties in specific life domains* (e.g., issues with housing, work or education, and time management) within the symptom domain.

While the three goal taxonomies identified outcomes related to *coping and self-management*, and Childs and colleagues (2013) emphasised the importance of *empowerment and self-efficacy*. Hoagwood's (1996) conceptual model includes no explicit reference to coping and self-management, and does not comment on whether these may be subsumed within the domains of *functioning* or *symptoms*. Fonagy (1997) suggests a separate domain to distinguish *mechanisms,* that is, "the cognitive and emotional capacities that probably underpin both symptomatology and adaptation" (p. 586). Change mechanisms have repeatedly been highlighted as an important outcome domain where evidence is still lacking (Cuijpers, 2019; Kazdin, 1999b, 2009). As discussed in Chapter 1, improved coping and self-management skills have also been frequently identified as a salient outcome by young people (e.g., Bergmans et al., 2009; Lavik et al., 2018). For these reasons, the new proposed taxonomy includes coping as an explicit higher-level outcome domain.

The concept of personal growth features most prominently in the Bern Inventory (2013), and to lesser extents in the child-focused goal taxonomies (Bradley et al., 2013; Rupani et al., 2014). Like coping, personal growth was not explicitly identified as an outcome domain by either Hoagwood (1996) or Fonagy (1997). The common mentioning of increased autonomy, improved self-esteem or strengthened identity by adolescents in the studies reviewed in Chapter 1 suggests that personal growth may gain importance as young people transition from childhood into adulthood. Therefore, personal growth is included as a higher-level outcome domain in the revised taxonomy.

Four further points are of note. First, while specified as a higher-level outcome domain by Hoagwood and colleagues (2012), parental symptoms and wellbeing were covered only by Child and colleagues (2013). This may reflect the broader consultation approach employed by that study, although parents were not consulted. Third, while the Bern Inventory (Grosse Holtforth & Grawe, 2002) proposed a number of outcomes related to the domain of wellbeing, such outcomes were generally not included in the child taxonomies. Nevertheless, wellbeing has been tentatively included as a domain in the revised taxonomy, in an effort to not discard its relevance prematurely. This decision will be reassessed based on the evidence from the remaining four studies constituting this doctoral thesis. Finally, the specific symptom domain of sexuality included in the Bern Inventory may not be relevant to adolescents. However, like wellbeing, it was tentatively included at this stage.

3.4 Discussion and Conclusions

This integrative narrative review aimed to identify existing taxonomies and frameworks of outcome in child mental health, and to synthesise these into a new, more comprehensive taxonomy that can serve as a conceptual framework throughout this thesis. Two conceptual models and four taxonomies were identified. The conceptual models were theoretically grounded and identified higher-level outcome domains, while the four taxonomies were empirically derived and identified higher-level domains as well as specific outcome categories.

None of the existing taxonomies was exhaustive in its coverage of possible treatment outcomes in child mental health, but rather complemented one another. The above discussion showcased the overlap as well as inconsistencies between these existing taxonomies, and the potential gains from synthesis. A new taxonomy was devised that consists of eight domains and 40 specific outcome categories. Definitions of each high-level domain are provided in Table 3.3 (below). This taxonomy served as an initial conceptual framework for the Outcome Measurement Review, the study of Post-Therapy Outcome Perspectives, and the study of Viewpoints on Outcome Priorities, presented in Chapters 4 through 6. It was iteratively reviewed and revised based on the outcome categories emerging from these studies.

The outcomes identified by this narrative review are framed around areas for positive change, rather than in relation to initial presenting problems. Taxonomies of outcome direct attention away from problems to visions of what improvement might look like. The latter cannot always immediately be inferred from the former, as the outcomes anticipated may vary, depending on the treatment approach and mechanisms chosen to address a given problem. Nevertheless, there is an association between problem categories and outcome categories. Weisz and colleagues began devising a system for categorising presenting problems in the early 1990s (Weisz & Weiss, 1991). This has since been applied and refined, including as part of the development of the idiographic *Youth Top Problems* outcome measure (Hawley & Weisz, 2003; Weisz et al., 2011; Yeh & Weisz, 2001). Many of the target problems identified by Weisz and colleagues map onto outcome categories proposed in this chapter: symptoms of *depression and anxiety*, *aggressive or delinquent behaviour*, or *somatic complaints* map onto the symptom domain; problems related to *daily living skills* link into the domain of functioning; *family and life stress* link into the category of relationships; *social problems* can be assigned to the domain of functioning or relationships; and *identity problems* link into the category of personal growth (Hawley & Weisz, 2003).

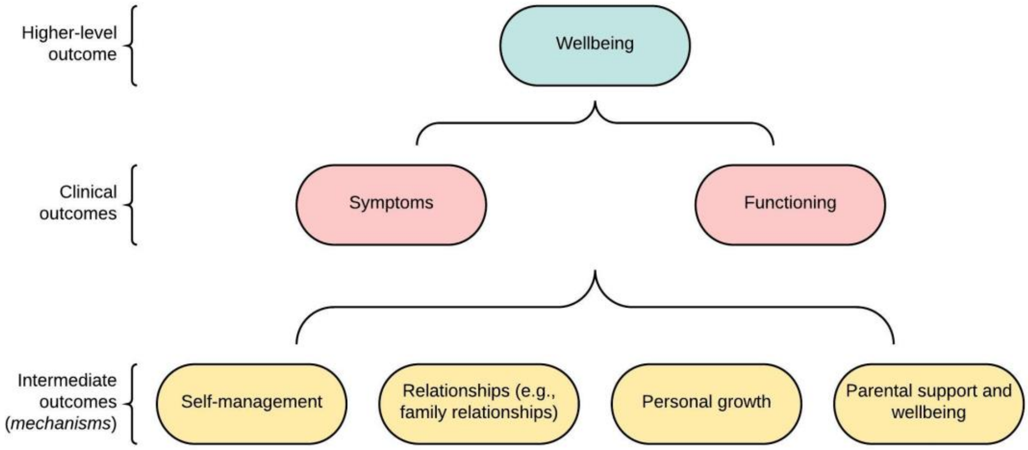
3.4.1 A Note on Mechanisms

The taxonomies considered as part of this narrative review all focussed on grouping outcomes thematically. None organised outcomes in terms of whether they related to higher-level (or ultimate) treatment aims, or to intermediate changes that might also be described as treatment mechanisms. The latter have generally been neglected in psychotherapy research (Cuijpers, Stringaris, & Wolpert, 2020; Kazdin, 1999a, 2009)

The taxonomy domain of *self-management* identifies a number of mechanisms by which young people may be able to help themselves, and cope more effectively. However, self-management is only one example for a group of changes described as *mechanisms* by Fonagy (1997). They include mechanisms of treatment, as well as mechanisms of disease, and are factors that drive both depressive symptomatology and improvement. For example, cognitive theories postulate that negative cognitive bias in the interpretation of available information, or attention bias towards negative as opposed to positive information represent cognitive styles that commonly underpin depression (A. T. Beck, 1976). CBT targets such cognitive distortions through treatment mechanisms such as thought identification or cognitive restructuring (Micco, Henin, & Hirshfeld-Becker, 2014). Other mechanisms commonly used in CBT include, for example, problem solving, emotion regulation, or behavioural activation (Kennard, Clarke, et al., 2009). Changes in these capacities constitute *intermediate outcomes* that may be instrumental in bringing about higher-level change in symptoms and functioning. As discussed in Section 1.6, mechanisms more typical for psychodynamic psychotherapy concern changes within the personal growth domain. Mentalisation-based psychoanalytic treatments may focus on attachment processes (Fonagy & Bateman, 2006), while systemic family therapy may focus on the influence of family functioning and relationships (Cottrell, 2002).

While the taxonomy devised in this chapter organises outcomes thematically, it is important to consider that these may occur at different times, be interrelated, and have different chances of occurring based on the treatment approach used. It may be helpful to conceive of a change hierarchy or sequential impact chain, as tentatively illustrated in Figure 3.1. Any psychotherapeutic treatment typically aims to alter intermediate outcomes (or *mechanisms*) in a first instance. This should facilitate symptom reduction and improved daily functioning, which constitute critical clinical outcomes, based on which diagnostic status is established. Epidemiologists or health economists may turn to even higher-level concepts of wellbeing or health-related quality of life are to judge the effectiveness of the clinical response at a population level (e.g., Jia, Zack, Thompson, Crosby, & Gottesman, 2015).

Figure 3.1. Illustration of Possible Outcome Levels



While developmental psychopathology has produced theory and phenomenological evidence on the mechanisms that underpin the development of mental health difficulties, no reliable biological marker of recovery from depression has yet been identified (Keller, 2003). As suggested by the National Institute of Mental Health (NIMH), biomarkers are biological indicators of disease processes, and may take the form of “a genetic variant, an abnormal amount of a certain protein, a distinct neuroimaging pattern from a brain scan, a certain response during a cognitive test, or any number of indicators from blood, sweat, or other biological fluids” (NIMH, 2008, p. 7). While a number of biomarkers for response to antidepressant treatment have been suggested, none currently has a solid evidence base (Cleare et al., 2015). However, if a relevant biomarker was identified, then this would constitute a critical addition to any outcome taxonomy for adolescent depression.

3.4.2 Limitations and Future Research

The present study focussed on considering and refining existing outcome taxonomies through an integrative narrative review. Since taxonomies aim to establish terminological reference points that can attract wide uptake within the relevant research community, their relevance and acceptability are important to consider. Future research might examine the acceptability of this taxonomy through a consultation process involving a reference group. This could be done using the Delphi technique, where several sequential rounds of structured consultation, feedback, and voting are used to generate consensus within an expert group (Linstone & Turoff, 1975; Powell, 2003). Similarly, it was beyond the scope of this study to examine the inter-rater reliability of the taxonomy when applied to empirical data. This may form another area for future research (e.g., Dodd et al., 2018; Michie et al., 2013).

Table 3.3. Conceptual Definitions of High-Level Outcome Domains

Outcome Domain	Definition
Symptom change	Changes in the severity or frequency of symptoms of depression or comorbid mental health difficulties
Self-management	Changes in young peoples' ability to manage their symptoms, feelings and thoughts, to cope with life's ups and downs, and in their self-efficacy and empowerment to take back control.
Functioning	Changes in young peoples' ability to meet role expectations at home, at school, and with peers and the wider community (Hoagwood et al., 1996).
Personal Growth	Changes related to young people developing themselves personally, which include becoming more mature and independent, becoming more assertive and self-confident, having greater self-regard and a stronger sense of identity, and being able to reflect on their personal history and its effect on the present.
Relationships	Changes in young people's relationships with family, friends, peers, and romantic partners.
Wellbeing	Wider improvements in young people's wellbeing, outlook into life, and future orientation.
Physical health	Improvements in biological markers (e.g., cortisol levels), general physical condition, or weight.
Parental support and wellbeing	Changes in the capacity of parents to understand their children's difficulties and provide adequate support and parenting; changes in parental wellbeing.

Chapter 4. Outcome Measurement Review[2]

---

[2] Parts of this chapter have been published as Krause, Bear, Edbrooke-Childs, & Wolpert, 2019.

4.1 Introduction

4.1.1 Background

A number of systematic reviews have taken stock of the types of outcomes that have typically been reported in treatment effectiveness studies constituting the evidence base in child mental health. A systematic review of 236 studies published between 1962 and 2002 found that all of the reviewed studies included at least one measure of symptoms, 28% reported on changes in functioning, 8% on consumer satisfaction, and 5% on environmental factors such as parenting (Weisz et al., 2005). A second review found that of 238 eligible treatment efficacy and service effectiveness studies published between 1980 and 2011, 95% assessed change in symptoms; 51% in functioning; 34% in services and systems; 29% in relationships and the family environment; 23% in consumer-oriented outcomes, 9% in parental symptoms; and 3% in physical health (Hoagwood et al., 2012; P.S. Jensen et al., 1996). A third systematic review looking specifically at treatment effectiveness studies measuring mental health and educational outcomes found that of 602 studies screened, only 15% reported on at least one educational outcome measure, even though 20% related to school-based interventions (Becker et al., 2014). Taken together, these reviews suggest that the evidence base for interventions in child mental health is primarily built upon the measurement of symptom change as a primary outcome, while other outcome domains are rarely covered.

As mental health services move towards more person-centred care, there is a call for greater participation of young people and families in the process of outcome measurement through the use of so-called patient-reported outcome measures (Black, 2013; Coulter, 2017; De Los Reyes et al., 2015; Department of Health, 2011, 2015). By assessing change not only from the clinician’s perspective but also from that of young people and parents, outcome measurement can inform and enhance communication and promote shared decision-making, as it may help establish a common understanding of the progress achieved (Valderas et al., 2008). There is growing awareness that high integrity health care systems cannot simply rely on clinicians’ views of what is most important, or on symptom change alone, but must consider service user perspectives to ensure that treatment meets their needs, priorities, and preferences (Mulley et al., 2017). Given the growing attention to person-centred care and shared decision-making, one might expect that more recent studies would show a broader approach to measurement, reflecting shifting mind-sets that increasingly recognise the importance of measuring outcomes that genuinely matter to service users, and of doing so by including them as informants.

The above-mentioned existing reviews have a number of limitations. They were broad in scope, covering childhood and adolescence, as well as a range of presenting problems. None examined approaches to outcome measurement specifically for adolescent depression. All commented on the frequency with which high-level outcome domains were reported, but none disaggregated these frequencies to a more granular taxonomic level. This limits conclusions about the frequency with which frequently endorsed outcomes identified in qualitative studies with young people (e.g., coping skills, autonomy, identity or family functioning) have been covered in research studies. Finally, all of these existing studies focused exclusively on controlled studies, in which the need to ensure the highest possible internal validity may have favoured symptom-focussed measurement approaches. Little is known about the extent to which a similar focus dominates studies using other designs.

4.1.2 The Present Study

The present study aimed to expand on the above-mentioned reviews and to address their gaps. This systematic review focussed on the most recently published treatment efficacy and effectiveness studies in the field of adolescent depression (considering studies published between 2007 and 2017). It broadened the inclusion criteria used by previous studies by explicitly searching for non-controlled studies (including studies that assessed outcome purely qualitatively). In addition, it specifically examined associations between study characteristics and the breadth of outcome measurement, based on the hypotheses that outcome measurement was likely to be more multidimensional (a) in more recently published studies, and (b) in studies using non-controlled designs. While previous reviews identified overarching outcome domains, this review undertook a nuanced mapping of outcome domains and outcome categories, using the initial taxonomy devised in the Outcome Taxonomy Review (Chapter 3). In light of the growing emphasis on person-centred care, additional attention was accorded to the extent to which young people were consulted as informants during outcome measurement. The three research questions guiding this review were:

1. What outcomes were measured for treatments of adolescent depression between 2007 and 2017?
2. To what extent does the breadth of outcome measurement vary with study characteristics?
3. To what extent were young people considered as informants?

For the sake of brevity, the term “treatment effectiveness study” will be used to refer to both treatment efficacy and effectiveness studies, unless the explicit distinction of both types is warranted.

4.2 Method

4.2.1 Search Strategy

Given this review’s broad scope with regards to eligible study designs, a search strategy was devised that prioritised sensitivity over specificity, in order to identify a broad range of outcome studies focussing on treatments for adolescent depression. Boolean operators were used to identify pre-specified search terms in publication titles and abstracts. The search syntax comprised dedicated sets of search terms to identify (a) randomised trials (e.g. “trial adj2 clinical) and (b) non-randomised studies (e.g., “routine adj3 care”, “service adj3 data”, or program* adj3 evaluat*), as well as terms specifically devised to narrow down the search to intervention studies (e.g., "intervention adj4 effect*) rather than epidemiological studies. In addition, search terms specified the study population (e.g., "child", "adolescent*", "depress*"), and a number of exclusion terms that were applied to study titles only (e.g., "toddler*", "preschool*"). The full search syntax can be consulted in Appendix B, Table B.1.

Three academic databases were systematically searched: PsycINFO, Medline and Embase. In order to be included, studies had to have been published in peer-reviewed journals between January 2007 and July 2017, and measure treatment outcomes either quantitatively or qualitatively. To complement the automated search, reference lists of key articles were reviewed, and relevant studies missed by the automated search were added manually. While no explicit restrictions were placed on study design or data quality, the search excluded papers that had not passed peer-review as a quality control mechanism, such as gray literature and doctoral dissertations. Language of publication was restricted to English, French, German, Italian or Spanish, as these are the languages spoken by the doctoral candidate.

4.2.2 Study Eligibility Criteria

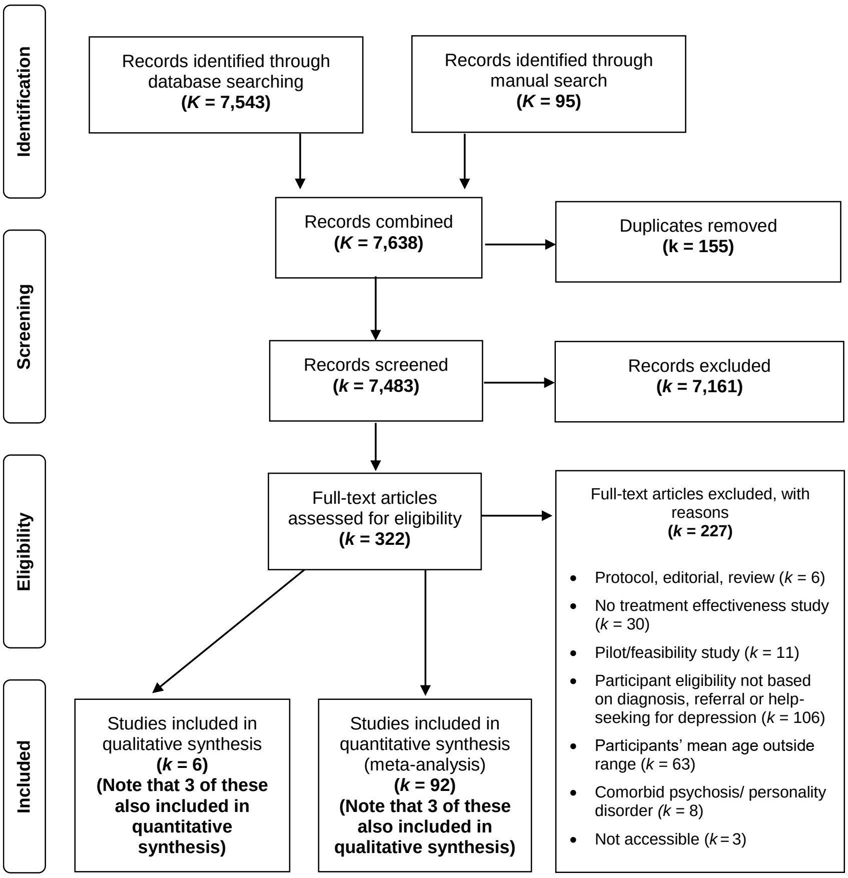
To be included, studies had to assess treatment efficacy or service effectiveness, *or* report on treatment outcome as part of a secondary analysis of moderators or predictors of treatment response. Any research design was eligible, including purely qualitative studies. Studies were included if the mean age of the treatment sample was between 12 and 19 years, representing mid to late adolescence as defined by the WHO (2017b). Study participants had to be diagnosed with a depressive disorder, help-seeking, or referred for depressive symptoms. Studies testing treatments in non-clinical populations (e.g., interventions universally targeted at pupils) were excluded. Studies focusing on young people with an underlying physical illness, developmental disorder, comorbid psychosis, or personality disorders were also excluded, in order to focus the assessment on outcomes related to depression as the primary presenting problem (note that this was reflected in some of the exclusion terms applied to study titles, e.g., "autistic" or "diabetes"). No restrictions were placed on the study setting or type of mental health treatment. Pilot and feasibility studies were excluded, as were studies focusing exclusively on prevention, maintenance, safety, treatment adherence, or engagement.

4.2.3 Study Selection and Data Extraction

Overall, the automated and manual searches yielded 7638 titles. After removing duplicates, 7483 titles were included in the screening of titles and abstracts (see Figure 4.1 for a flowchart of the screening process). Due to the inclusive search strategy, a large number of studies initially identified were subsequently excluded as they did not focus specifically on adolescent depression, or did not test treatment effectiveness or efficacy. Following the title and abstract screening, 322 studies were retained for full-text screening. An independent second reviewer (Holly Bear) replicated the full-text screening for 10% of the retained studies, yielding a kappa coefficient of 0.68, indicating substantial inter-rater agreement between both reviewers (Landis & Koch, 1977). The final sample included 95 studies. Of these, 92 were considered for quantitative synthesis, and six for qualitative synthesis, with three studies considered for both. Information was systematically extracted about study characteristics (e.g., study design, participant characteristics, treatment approach, publication date), the specific outcome measures used, and informants consulted for outcome measurement (e.g., clinician, parent, young person). An overview of all included studies can be found in Appendix B, Table B.4 and Table B.5).

The term “study” will be used in the remainder of this chapter to refer to each of the 95 publications considered in this review. However, it must be caveated that several of these publications presented secondary analyses of data collected through larger anchor studies, and so do not represent primary research studies in themselves.

Figure 4.1. Flowchart of the Screening Process



Note. Adapted from Moher et al. (2009)

4.2.4 Assessment of Data Quality and Risk of Bias

The methodological quality of the included studies was assessed using an abbreviated version of the Downs and Black (1998) checklist for data quality in randomised and non-randomised treatment studies. Four of the originally 27 items in this checklist were removed because most studies did not report the required information to assess these criteria. The abbreviated checklist included 23 equally weighted evaluation criteria, and the total attainable data quality score was 23. Purely qualitative studies ($n = 3$) and a single case study were excluded from this data quality assessment, as the quality criteria could not be meaningfully applied to these studies. The adapted checklist (Table B.2) and data quality scores for each study (Table B.3) can be consulted in Appendix B.

4.2.5 Data Synthesis and Statistical Analysis

The types of outcomes and outcome domains covered in the studies were identified by mapping each reported outcome measure to the principal concept they were designed to measure. This was determined with reference to the taxonomy devised through the Outcome Taxonomy Review (see Chapter 3), handbooks of psychometric measures (Rush, First, & Blacker, 2008), and the original studies reporting on the relevant measure's development or validation. At this stage, 121 distinct measures were identified. To ensure the reliability and validity of the conceptual mapping, a second reviewer (Holly Bear) independently coded 25% of these 121 measures. Diverging codes were then discussed and a final coding agreed. The coding framework thus obtained was also used for a brief thematic analysis (Braun & Clarke, 2006) of outcomes described in the six studies measuring outcomes qualitatively.

Descriptive bivariate analysis was performed on the 92 included studies that presented quantitative outcome measures to explore differences in outcome domains measured and informants consulted over time, and in relation to study characteristics. A series of Kruskal–Wallis tests were conducted to assess whether there were any differences based on study characteristics with respect to the number of outcome domains covered (Kruskal & Wallis, 1952). Where the Kruskal–Wallis test showed a significant group difference ($p < .05$), post hoc pairwise comparisons were carried out using Dunn's test with Benjamini-Hochberg adjustment for false discovery rates in multiple comparisons, allowing for positively dependent test statistics (Benjamini & Hochberg, 1995; Dunn, 1961). Spearman's correlation coefficient for ordinal and non-parametric data (r_s) was computed to test whether there were associations between the number of outcome domains covered and the data quality assessment score or publication year (Spearman, 1987). In addition, a Mann-Whitney U test (Mann & Whitney, 1947) was carried out to test whether there were any significant differences in the breadth of outcome measurement based on whether studies had been published in the first or second half of the review period.

4.3 Results

4.3.1 Study and Participant Characteristics

Table 4.1 (below) displays the characteristics of the 95 studies under review. The majority (67%) originated in the United States or Canada; nine studies (10%) had each been implemented in the United Kingdom or in Australia/New Zealand; and the remainder were implemented in Continental Europe, the Middle East and Asia, or across multiple regions.

Table 4.1. Study Characteristics

Study characteristics	*k* (%)	Average number of domains covered	Average number of measured used
Publication year			
2007 - 2012	68 (71.6)	1.9	3.2
2013 - 2017	27 (28.4)	2.3	4.7
Country			
United States and Canada	62 (67.4)	1.9	3.5
United Kingdom	9 (9.5)	2.6	5.3
Continental Europe	7 (7.4)	1.9	3.1
Australia and New Zealand	9 (9.5)	1.9	3.9
Middle East and Asia	6 (6.3)	2.0	4.2
Multiple regions [a]	2 (2.1)	2.0	2.0
Study type			
Randomized controlled trial	69 (72.6)	2.0	3.6
Observational study	11 (11.6)	2.0	4.3
Case study or qualitative study		3.3	5.3
Clinical follow-up	5 (5.3)	1.4	2.4
Retrospective chart review	3 (3.2)	1.3	2.0
Study objective			
Assess treatment effectiveness	49 (51.6)	2.3	4.5
Study predictors / mediation factors	33 (34.8)	1.6	2.4
Examine long-term outcomes	6 (6.3)	1.5	3.7
Cost-effectiveness	4 (4.2)	2.0	2.0
Other objective [b]	3 (3.2)	2.5	6.0
Sample size			
1–10	7 (7.4)	2.8	4.4
11–50	14 (14.7)	2.4	5.5
51–100	15 (15.8)	1.9	4.3
101–300	23 (24.2)	1.8	3.1
301–500	36 (37.9)	1.8	2.9
Study intervention			
Combined treatment	41 (43.2)	1.8	3.1
Psychotherapy	31 (32.6)	2.4	4.9
Medication	17 (17.9)	1.9	3.3
Routine care	6 (3.2)	1.4	1.8

[a] Includes one study conducted in the United States, Canada, Mexico, and Argentina; and one study conducted in the United States, Eastern Europe, South Africa, and Western Europe. [b] Other objectives include assessing the transportability of an intervention, reflecting on lessons learned, and assessing treatment experience qualitatively.

More than two thirds of the reviewed studies were RCTs, and 12% were observational studies, with the remainder being case studies, retrospective chart reviews and purely qualitative studies. Half of the studies assessed the efficacy or effectiveness of treatments or service delivery initiatives, while the other half assessed predictors or moderators of treatment response, cost-effectiveness, or long-term treatment outcomes using a follow-up design. These latter studies often presented secondary analyses of data from large primary studies such as the Treatment for Adolescents with Depression Study (Treatment for Adolescents With Depression Study [TADS] Team, 2004) or the Treatment of SSRI-resistant Depression in Adolescents (TORDIA) study (Brent et al., 2008).

The majority of studies (71%) were conducted in outpatient settings. The most common type of intervention were combined treatments involving antidepressant medication and psychotherapy (45%), followed by psychotherapeutic approaches (33%), medication (18%), and routine service provision or quality improvement initiatives. The mean data quality score for the 92 quantitative studies based on the modified Downs and Black (1998) checklist was 17 score points (*SD* = 2.8) out of 23 possible score points, with individual scores ranging from 10 to 22. Non-randomised studies (*n* = 20) had a mean quality score of 13.7 points (*SD* = 2.0), whereas randomised studies had an average quality score of 18.1 points (*SD* = 2.2).

The studies involved a total sample of 7394 young people aged 12 to 19 years. The mean age of participants across studies was 15.1 years, 65% were female, and the majority (82%) had been recruited based on meeting diagnostic criteria for a depressive disorder, whereas the remainder had been recruited based on a clinical referral or self-referral.

4.3.2 Outcome Measurement Across the Sample

Across the 92 quantitative studies, 121 different outcome measures were identified. These mapped onto the nine outcome domains defined in the Outcome Taxonomy Review: *symptoms*, *self-management*, *functioning*, *relationships*, *personal growth*, *service use and satisfaction*, *wellbeing*, *physical health*, and *parental wellbeing*. They further mapped to 33 more specific outcome concepts (see Appendix B, Table B.5 and Table B.6 for further detail on what measures were identified and how these were mapped conceptually).

On average, each study assessed treatment effectiveness or efficacy across 2.0 outcome domains (*SD* = 0.9), using four different measures. While 15 studies used only a single measure, one study used as many as 14 different measures (G. N. Clarke et al., 2016). The domain most frequently assessed (by 94% of studies) was that of *symptoms* (see Table 4.2, below). This domain included specific outcome categories relating to core depressive symptoms, or other symptoms related to comorbid presenting problems, such as self-harm, substance use, behaviour problems, and issues with eating and sleep. Overall, 58 studies reported a change within the symptom domain as their primary outcome, which constituted 85% of all studies that identified a primary outcome measure.

Table 4.2. Outcome Domains and Subdomains by Frequency of Measurement

	Number (and %) of studies covering outcome:		Covered in qualitative studies (*k* = 6)	Identified by the Outcome Taxonomy Review
Outcome	in general (*k* = 92)	as primary outcome (*k* = 68)		
Symptoms	**86 (93.5)**	**58 (85.3)**	✓	✓
Depressive symptoms	86 (93.5)	57 (83.8)	✓	✓
Suicidality	15 (16.3)	3 (4.4)	✓	✓
Self-harm	1 (1.1)	—	—	✓
Anxiety	7 (7.6)	1 (1.5)	—	✓
Substance use	6 (6.5)	1 (1.5)	✓	✓
Behaviour and anger	4 (4.3)	1 (1.5)	—	✓
Disordered eating	1 (1.1)	—	—	✓
Sleep	2 (2.2)	—	—	✓
Other comorbidities	3 (3.3)	—	✓	✓
General psychopathology	6 (6.5)	—	—	—
Sexuality	—	—	—	✓
Self-management	**14 (15.2)**	**2 (2.9)**	✓	✓
Cognitive processes	9 (9.8)	—	✓	—
Behavioural and physical activation	4 (4.3)	2 (2.9)	✓	✓
Coping with mood, thoughts & feelings	2 (2.2)	—	✓	✓
Understanding feelings and thoughts	—	—	—	✓
Functioning	**51 (55.4)**	**27 (39.7)**	✓	✓
Academic functioning	—	—	✓	✓
Executive functioning	2 (2.2)	—	✓	✓
Global functioning	48 (52.2)	27 (39.7)	✓	✓
Social functioning	3 (3.3)	—	✓	✓
Relationships	**4 (4.3)**	—	✓	✓
Attachment style	1 (1.1)	—	—	—
Family functioning and relations	4 (4.3)	—	✓	✓
Friendships / peer relationships	1 (1.1)	—	✓	✓
Able to talk about feelings & thoughts	—	—	—	✓
Connectedness & intimacy	—	—	—	✓
Romantic relationships	—	—	—	✓
Personal growth	**7 (7.6)**	**1 (1.5)**	✓	✓
Self-concept and identity	6 (6.5)	1 (1.5)	—	✓
Assertiveness	1 (1.1)	—	✓	✓
Self-esteem, confidence, acceptance	1 (1.1)	—	✓	✓
Autonomy and responsibility	1 (1.1)	—	✓	✓
Empowerment and self-efficacy	—	—	✓	✓
Fulfilling desires and wishes	—	—	✓	✓
Processing past and present	—	—	—	✓
Meaning and purpose	—	—	—	✓

Outcome	Number (and %) of studies covering outcome:		Covered in qualitative studies (k = 6)	Identified by the Outcome Taxonomy Review
	in general (k = 92)	as primary outcome (k = 68)		
Wellbeing	**7 7.6)**	—	✓	✓
Health-related quality of life	7 (7.6)	—	—	✓
Relaxation and composure	—	—	✓	✓
Physical health	**3 (3.3)**	**2 (2.9)**	✓	✓
Biomarkers	3 (3.3)	2 (2.9)	—	—
Parental symptoms	**2 (2.2)**	—	✓	✓
Parental symptoms and coping	2 (2.2)	—	✓	✓
Parental support / parenting	—	—	✓	✓
Idiographic outcome measures [a]	2 (2.2)	—	✓	✓
Service-level outcomes	**8 (8.7)**	—	✓	✓
Client satisfaction	3 (3.3)	—	—	✓
Use of other services	4 (4.3)	—	—	—
Therapeutic alliance	2 (2.2)	—	—	—
Treatment retention	1 (1.1)	—	—	✓
Retention of therapeutic content	1 (1.1)	—	—	—

Note. ✓ indicates that an outcome category was identified by at least one of the six reviewed studies that measured outcomes qualitatively, or by the Outcome Taxonomy Review. — indicates that an outcome category was not identified.

[a] The item content of idiographic outcome measures is determined by individual service users and can cover any of the thematic outcome domains or categories identified.

The second most frequently assessed domain was that of *functioning*, measured by 51 studies (55.4%), and identified as a primary outcome by 27. Most of these studies (k = 48) included a measure of global functioning, which in two cases included the achievement of developmental milestones such as graduating from high school or entering employment (Melvin et al., 2013; Peters et al., 2016). Three studies assessing the efficacy of Interpersonal Psychotherapy for Adolescents specifically measured social functioning (Mufson et al., 2014; Rosselló et al., 2008; Spence et al., 2016), in line with the approach's focus on interpersonal difficulties (Klerman & Weissman, 1994). In addition, two studies measured executive functioning (Ariga et al., 2010; Bloomquist et al., 2016). While some global functioning measures like the Impairment Rating Scale (Fabiano et al., 2006) and the Behavioral Assessment System for Children-2 (BASC-2; C. R. Reynolds, 2010) include items related to functioning in school, none of the studies explicitly reported on outcomes related to school attendance or attainment.

The outcome domain of self-management was covered by 14 studies (15%). Within this domain, nine studies measured changes in *cognitive or behavioural patterns* associated with depression, such as ruminative thinking or negative attributional styles. In addition, four measured outcomes related to behavioural and physical activation (Carter et al., 2015; Clarke et al., 2016; Parker et al., 2016; Pass et al., 2015), of which two studies specifically tested the effectiveness of physical exercise in combatting depression, and one was a behavioural activation intervention. Two studies measured changes in young peoples' ability to cope with emotions and thoughts (Eskin et al., 2008; Lusk & Melnyk, 2011).

Changes related to other outcome domains were less frequently measured. The domains of *interpersonal relationships*, *personal growth*, *wellbeing, parental symptoms*, and *physical health* were each covered by less than one in ten studies, and only three studies measured change in any of these as a primary outcome. The domain of *personal growth* (k = 7) covered aspects such as self-concept, self-esteem, assertiveness, and autonomy. The domain of *wellbeing* (k = 7) captured a broader concept of well-being and health-related quality of life. Six studies assessed change with regards to *interpersonal relationships*, covering family relationships and functioning, and peer relationships. Only three and two studies, respectively, measured changes in *physical health,* or *parental well-being*. Eight studies covered service-level outcomes within the domain of *service quality and satisfaction* (k = 8), including aspects such as consumer satisfaction, therapeutic alliance, and treatment adherence.

Of the six studies that included qualitative analysis of treatment response, all but one measured change in the domain of personal growth, four each covered symptoms, self-management, and interpersonal relationships, three covered functioning, and one each covered service satisfaction and physical health. The qualitative data touched upon additional subdomains not measured by the quantitative studies, such as school performance and attendance (k = 3), increased motivation and aspirations (k = 2; allocated to the category of 'fulfilling desires and wishes') and greater peace of mind (k = 1). They also highlighted a range of additional parent-related outcomes, including parental ability to cope, acceptance of the child's strengths and weaknesses, and parenting practices.

Commonly Used Outcome Measures

As mentioned above, as many as 121 different measurement tools were used across the 95 reviewed studies. The fifteen measures most commonly reported all focussed on measuring outcomes in the domains of symptoms or functioning (see Table 4.3, below). Of these, eight were clinician-rated measures and seven were self-report measures that were to be filled in by young people and/or their parents.

The single most commonly used instrument was the Children's Depression Rating Scale Revised (CDRS-R; Poznanski & Mokros, 1996), a 17-item clinician-led interview that assesses depression severity and impairment. The CDRS-R was used in close to half of all quantitative studies (n = 45) and served as a primary outcome measure in 37.

Table 4.3. The Fifteen Most Commonly Reported Outcome Measures

Rank	Outcome measure	Principal outcome domain	Informant(s)	Studies using this measure *k* (%) (*K* = 92)	Studies using this as primary outcome measure *k* (%) (*K* = 68)
1	Children's Depression Rating Scale (Poznanski & Mokros, 1996)	Symptoms	CL	45 (48.4)	37 (54.4)
2	Clinical Global Impressions Scale Improvement (Busner & Targum, 2007)	Functioning	CL	32 (34.4)	21 (30.9)
3	Beck Depression Inventory [a] (A. T. Beck, Steer, & Carbin, 1988)	Symptoms	YP	16 (17.2)	5 (7.4)
4	Clinical Global Impressions Scale Severity (Busner & Targum, 2007)	Functioning	CL	16 (17.2)	3 (4.4)
5	Kiddie-SADS-Present and Lifetime Version (Kaufman et al., 1997)	Symptoms	CL	14 (15.1)	3 (4.4)
6	Children's Global Assessment Scale (Dyrborg et al., 2000)	Functioning	CL	13 (14.0)	3 (4.4)
7	Suicidal Ideation Questionnaire Junior (W. M. Reynolds, 1988)	Symptoms	YP	10 (10.8)	3 (4.4)
8	Children's Depression Inventory [b] (Kovacs, 1992)	Symptoms	YP or PA	10 (10.8)	3 (4.4)
9	Hamilton Rating Scale for Depression (Hamilton, 1986)	Symptoms	CL	10 (10.8)	5 (7.4)
10	Mood and Feelings Questionnaire (short/long version)(Angold et al., 1995)	Symptoms	YP	6 (6.5)	2 (2.9)
11	Health of the Nation Outcome Scale (Gowers et al., 1999)	Functioning	CL	6 (6.5)	1 (1.5)
12	Reynolds Adolescent Depression Scale (W. M. Reynolds, 1987)	Symptoms	YP	6 (6.5)	2 (2.9)
13	Adolescent Longitudinal Interval Follow-up Evaluation (Keller et al., 1987)	Symptoms	CL	5 (5.4)	4 (5.9)
14	Child Behaviour Checklist (Achenbach, 1991)	Symptoms	YP or PA	4 (4.3)	0 (0.0)
15	Centre for Epidemiologic Studies Depression Scale (Radloff, 1977)	Symptoms	YP	4 (4.3)	1 (1.5)

Note. CL = clinician; PA = parent; YP = young person.

[a] Includes the Beck Depression Inventory first and second version. [b] Includes the Children's Depression Inventory first and second version.

The second most commonly used scale was the clinician-rated Clinical Global Impression Scale Improvement (CGI-I; Busner & Targum, 2007). The CGI-I is a one-item assessment of the clinician's sense of change in service user's symptom burden (Busner & Targum, 2007). It was used by 32 studies and served as a primary outcome measure in 21. The third most commonly used scale was the Beck Depression Inventory (BDI; A. T. Beck et al., 1988), a youth self-report measure of depressive symptoms used by 16 studies and serving as a primary outcome measure in five. As many as 81 of the 121 measures identified were used in only one study.

The majority of measures used across the 95 studies were nomothetic instruments. Idiographic measures that captured young people's self-defined treatment goals or presenting problems were used by only two studies. One study assessed progress towards achieving self-defined goals (Pass et al., 2015), and another (G. N. Clarke et al., 2016) assessed improvement in initial presenting problems using the Target Complaints Checklist (TCC; Elkin et al., 1985). The TCC lets young people and parents identify up to 3 target complaints at the start of treatment (e. g., "Having no friends"), and rate the degree of improvement for all targets at every follow-up meeting. Idiographic measures provide a means of tracking change across a wide range of outcome domains without the need for dedicated nomothetic instruments.

4.3.3 Multidimensional Outcome Measurement by Study Characteristics

A number of Kruskal-Wallis H tests and Spearman correlation coefficients were computed to explore differences in the breadth of outcome domains measured by study characteristics. Kruskal–Wallis tests indicated no significant difference in the number of domains covered by country of implementation, and Spearman correlation coefficients showed no correlation between the data quality score and the approach to outcome measurement. There were, however, significant differences in outcome measurement between different study designs, sample sizes, and the type of treatment tested.

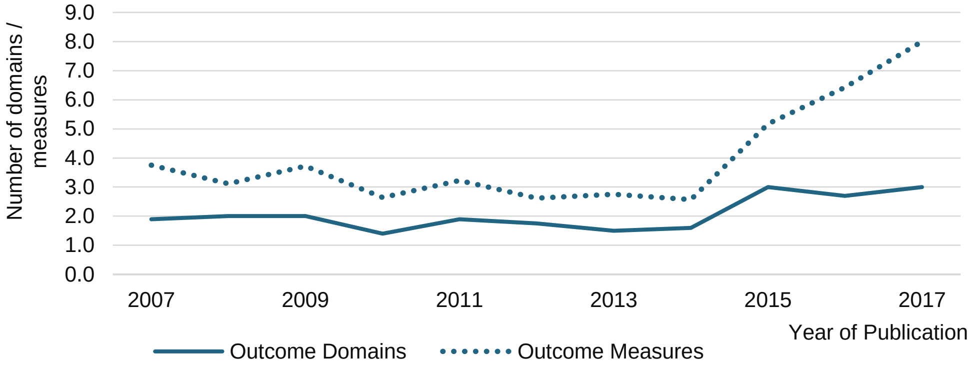
There was a statistically significant difference in the number of outcome domains covered (Kruskal Wallis χ^2 (3) = 10.80, p = .013) according to the treatment approach tested. Dunn's pairwise post hoc comparisons with Benjamini and Hochberg adjustment for multiple comparisons (Benjamini & Hochberg, 1995) indicated that studies testing purely psychotherapeutic treatments tended to take a broader approach to outcome measurement than studies testing the efficacy of combined treatments (p = .007) or the effectiveness of routine care (p = .017). On average, studies of psychotherapeutic treatments covered 2.4 outcome domains, compared with 1.8 in the studies testing combined treatments and 1.4 domains in studies of routine care interventions. There was no evidence of a difference between the other pairs.

There further was a statistically significant difference in the breadth of outcome measurement according to study type (Kruskal-Wallis χ^2 (4) = 10.071, p = .039). Dunn's pairwise post hoc tests showed a significant difference between case studies and qualitative studies on the one hand, and clinical follow-ups (*p* = .024) and retrospective chart reviews (p = .023) on the other hand. The mean number of domains covered by the former was 3.3, compared with 1.4 and 1.3 in the latter groups. The number of domains covered appeared to decrease with study size, from 2.8 (*SD* = 1.6) outcome domains measured in studies with up to 10 participants, to 1.8 (*SD* = 0.7) domains in studies with more than 300 participants. However, a Kruskal-Wallis test showed no significant difference (Kruskal Wallis χ^2 (4) = 6.505, *p* = .165).

In terms of the country of implementation, the broadest approach to outcome measurement was apparent in studies from the UK, which covered 2.6 domains (*SD* = 1.1) on average, while other countries and regions used only 2.0 measures on average (see Table 4.1, above). UK-based studies also used the largest average number of outcome measures, namely 5.3 (*SD* = 4.1), compared with 4.2 (*SD* = 2.5) in the Middle East and East Asia, 3.9 (*SD* = 2.6) in Australia and New Zealand, 3.5 (*SD* = 2.4) in North America and 3.1 (*SD* = 2.0) in Continental Europe. However, a Kruskal-Wallis test showed no significant difference by country or region of implementation (Kruskal Wallis χ^2 (5) = 4.072, *p* = .539).

Spearman's correlation coefficient showed no significant association between year of publication and the number of outcome domains or measures reported. However, a Mann-Whitney U test indicated a significant difference in the number of outcome domains covered between studies published from 2007 to 2012 on the one hand, and studies published between 2013 and 2017 on the other hand (*z* = -2.058, p= 0.040). While studies published in the first half of the review period covered 1.9 domains on average, it was 2.3 domains in the second group. This recent upward trend over the 10-year reference period is also evident from Figure 4.2 (below). Whilst an average of 1.9 (*SD* = 0.4) domains were covered by studies published in 2007, this had increased to 3.0 (*SD* = 0.9) by 2015. In parallel, the number of outcome measures used increased even more steeply, from an average of 3.8 (*SD* = 1.5) in 2007, to 5.2 across the six studies published in 2015, to 6.4 (*SD* = 4.9) in the seven studies published in 2016, and 8.0 (*SD* = 1.4) in the two studies published in 2017.

Figure 4.2. Mean Number of Outcome Domains and Measures Reported, 2007– 2017



Note. The figure shows the average number of outcome domains and outcome measures reported in the reviewed studies by publication year.

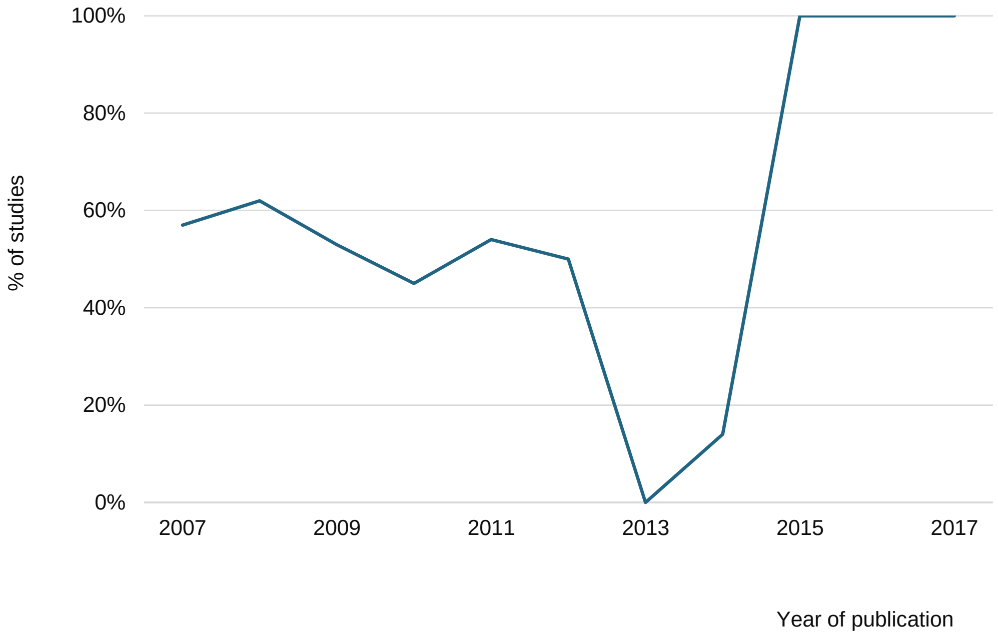
4.3.4 Informants Consulted Across Outcome Domains

Thirty-nine studies (42%) assessed change in outcomes using exclusively clinician-rated measures, while 30 studies (33%) used both clinician and youth-reported measures, ten (11%) used youth and parent self-report, and eight (9%) relied entirely on youth report. Overall, youth report was considered by more than half of the studies (53%). However, primary outcome measures, as identified by the relevant study authors, predominantly relied on clinician-report (75%).

As shown in Figure 4.3, there was an inconsistent trend in the use of youth self-report over time. Of the seven studies published in 2007, 57% used at least one youth self-report measure and the percentage remained relatively stable until 2012, then dropping to zero in 2013 (*n* = 4) and remaining low at 14% for the seven studies published in 2014. In contrast, all fourteen studies published since 2015 have included a youth self-report measure.

In some outcome domains, youth self-report served as the primary source of information, whereas other domains were mainly assessed using clinician report. All relevant studies consulted young people to measure change in cognition and behavior, and quality of life, with youth being the only relevant informant in 80% of these studies. Likewise, youth self-report was the main source of information concerning the domains of personal growth, service satisfaction, and interpersonal relationships. Clinician-report dominated the measurement of depressive symptoms and functioning. While youth report on depressive symptoms was gathered by half of all studies, only six percent included youth-self report for functioning.

Figure 4.3. Percentage of Studies Including Youth Self-Report, 2007-2017



Note. The figure shows the percentage of reviewed studies published in a given year that included at least one self-reported outcom measure, filled in by the young person receiving treatment.

4.4 Discussion

4.4.1 Summary of Findings

This systematic review aimed to establish the types of outcomes that were measured in the recent treatment effectiveness literature for adolescent depression, to explore associations between study characteristics and measurement approaches, and to assess the inclusion of young people as informants. The review found that the measurement of treatment efficacy and effectiveness for adolescent depression over the past decade focused primarily on symptoms and functioning. Other domains have occasionally been covered as secondary outcomes, using a plethora of different measures (see Appendix B, Table B.6), generally without providing a detailed rationale for their selection. Case studies and qualitative studies tended to employ broader approaches to measurement, while retrospective case reviews and clinical-follow-up tended to cover the narrowest range of outcome domains. On average, evaluations of psychotherapeutic treatments covered a broader range of domains than evaluations of combined treatments or routine care. Clinician-report was the predominant source of information for primary outcomes. Although over half of the reviewed studies included youth self-report, this was mainly to assess secondary outcomes. A trend towards more multidimensional measurement and the inclusion of youth self-report was observed in the most recently published studies.

The focus on measuring outcomes in the domains of symptoms and functioning was consistent with findings from earlier reviews of treatment outcomes in child mental health across age groups and presenting problems (Becker et al., 2011; Hoagwood et al., 2012; P. S. Jensen et al., 1996; Weisz et al., 2005). This suggests a high degree of continuity in the predominance of symptom-focussed outcome measurement over the past decades, with change only becoming apparent in the past five years. This recent increase may be reflective of the rising policy interest in outcome measurement and a parallel growth in research about self-reported and idiographic outcomes (Coulter, 2017; Edbrooke-Childs et al., 2016; Jacob et al., 2017; Wolpert, Ford, et al., 2012). With one exception (Weisz et al., 2009), calls for multidimensional and developmentally informed approaches to measuring outcomes in child mental health (Fonagy, 1997; Hoagwood et al., 1996; Kazdin, 1999b) were not explicitly acknowledged or discussed by any of the reviewed studies. Rationales for selecting specific outcome domains and discarding others were not generally provided.

The predominance of clinician-report in the measurement of primary outcomes may reflect a belief that clinicians provide a global view of clinical improvement, as their assessments frequently draw on information provided by young people and their parents (De Los Reyes et al., 2011). The evidence, however, suggests that clinicians attach greater importance to parent report than to youth self-report (Hawley & Weisz, 2003; Kramer et al., 2004; Loeber et al., 1990) and that different informants often provide inconsistent but equally valid accounts (De Los Reyes & Kazdin, 2005; Hunsley & Mash, 2007). Discrepancies between these accounts may suggest that treatment has altered a young person’s behavior in one context but not another, which provides information that can be both clinically useful and insightful with regards to specifying treatment mechanisms (Hunsley & Mash, 2007). The reviewed studies did, however, not explore such differences.

The recent rise in studies that include youth self-report may reflect a move towards more person-centred and youth-guided care as guiding principles for providers and funders of mental health care (Department of Health, 2010; National Quality Forum, 2017). For instance, the US Substance Abuse and Mental Health Services Administration (SAMHSA) mandates that young people be systematically consulted on the design, implementation and evaluation of community-based systems of care funded through its Children’s Mental Health Initiative (CMHI; Davis-Brown et al., 2012). Nevertheless, the reviewed studies rarely elaborated on their rationale for including youth self-report and did not refer to the youth advocacy movement spearheaded by SAMHSA in the United States, or comparable initiatives.

4.4.2 Review of the Initial Outcome Taxonomy

The initial outcomes taxonomy developed through the Outcome Taxonomy Review served as a reference framework for the extraction and systematic classification of outcomes reported by the reviewed studies. The systematic Outcome Measurement Review thus presented the first application of this taxonomy, and a first attempt at proof of concept.

A majority of outcome categories proved applicable to the data. The review covered the full range of symptoms specified in the taxonomy, with the exception of symptoms related to sexuality. Given the focus on adolescence as a phase when sexual activity cannot be assumed for all study participants, this may not be surprising. All four functioning outcomes listed in the taxonomy were also identified in the review. Within the domain of self-management, two changes were made to the initial taxonomy. First, the outcome category *exercise and activity*, which was previously included within the wellbeing domain, was renamed *behavioural activation and physical activity,* and moved to the *self-management* domain. This change was made because behavioural activation constituted a deliberate intervention mechanism in a number of the reviewed studies, rather than just an aspect of wellbeing. Second, the outcome category *cognitive processes* was newly added to the taxonomy, as several studies reported on changes in the cognitive processes and biases thought to underpin depression (e.g., negative interpretation bias).

In the domain of *personal growth*, most outcomes included in the initial taxonomy were also reported by at least one of the reviewed studies, with the exception of *meaning and* purpose, *recognising and fulfilling desires*, and *processing past and present experiences*. In the *relationship* domain, *being able to talk about feeling and thoughts*, greater *connectedness and intim*acy and *romantic relationships* were outcomes covered in the taxonomy but not identified in any of the reviewed studies. While this doctoral thesis focusses on person-centred rather than service-level outcomes, the systematic review identified a process-related outcome concept that may be relevant at an individual level, namely *therapeutic alliance*. This will be considered as an addition to the draft taxonomy in subsequent studies (Chapters 5-6).

The *wellbeing* domain displayed a certain amount of conceptual ambiguity. The original taxonomy included a number of loosely related concepts (i.e., *exercise and activity*, *enjoying life*, *loneliness and grief*, *relaxation and composure*, *wellbeing*), many of which were extracted from the adult-focussed Bern Inventory (Grosse Holtforth & Grawe, 2002). None of the conceptual outcome models considered for conceptual guidance discussed wellbeing as an outcome domain (Fonagy, 1997; Hoagwood et al., 1996). Studies reviewed as part of the Outcome Measurement Review reported on *quality of life* rather than wellbeing, thus adding another concept for consideration. As mentioned above, *Exercise and activity* was moved to the self-management domain. The outcome categories of *Enjoying life* and *Loneliness and grief* were not explicitly reported. *Enjoying life* can be described as the inverse of anhedonia (a typical depressive symptom) and be subsumed within the category of depressive symptoms, as might *Loneliness and grief.*

4.4.3 Implications

Using multiple outcome measures and informants is likely to be challenging in practice. Young people are most willing to engage if the changes measured are meaningful to them (Childs et al., 2013; Moran et al., 2012; Stasiak et al., 2013). Clinicians require measures that deliver actionable results such as alerting them to warning signs for individual service users, whereas commissioners may prefer data that can support benchmarking (Boswell et al., 2015). There may be inherent tension between these needs. A mitigation strategy may be to develop flexible core sets of outcome measures to streamline approaches to measurement and facilitate benchmarking (Szatmari et al., 2019). Such a core outcome set could include a small number of outcomes that different stakeholder groups can agree on, and recommend additional measures within other domains that could be integrated, based on shared decision-making with parents and young people. Such a multi-dimensional core battery would also benefit clinical research, enhancing the comparability of findings across studies and addressing the current fragmentation of measures (Szatmari et al., 2019).

Indeed, policy makers have begun to encourage or mandate the use of common sets of outcome measures, such as the Core Set of Children's Health Care Quality Measures for Medicaid and the Children's Health Insurance Program (CHIP) in the United States (Measure Applications Partnership, 2017) or the Australian Government's National Outcomes and Case mix Collection (Department of Health and Ageing, 2003). In the summer of 2018, the United States National Committee for Quality Assurance (NCQA) launched a National Collaborative for Innovation in Quality Measurement (NCINQ), which will convene a national panel of advisors including youth and family representatives to support health plans and states in researching possible improvements to depression care for adolescents using three core quality measures included in the NCQA's Healthcare Effectiveness Data and Information Set

(HEDIS; National Committee for Quality Assurance (NCQA), 2018). In light of such significant national attention to quality and outcomes measurement for adolescent depression, clarifying what is important to measure appears both imperative and urgent.

4.4.4 Limitations

Several limitations of this study should be noted. First, while the language of publication was not restricted to English, no foreign-language study met the inclusion criteria, and around half of the included studies were conducted in the United States. As a result, regional differences in outcomes measurement could not be explored in detail. Second, the systematic literature search aimed to identify a broad range of studies, including uncontrolled designs, but only considered studies published in peer-reviewed journals to apply a minimum level of quality control. It may have missed reports of programme or service evaluations published in the grey literature. Studies conducted in routine care settings may therefore be under-represented relative to more rigorously designed clinical trials that passed peer review. Additional research is needed to explore what outcome domains are measured in real-world CAMH settings, and the extent to which the picture converges with that found by this review.

Third, as part of this review, measures were mapped to the outcome constructs that they were *primarily* designed to operationalise, based on the original development and validation studies or relevant handbooks, and based on the concept that the relevant study authors described when reporting the relevant scores. However, some scales, especially within the functioning domain, are themselves multidimensional, covering aspects of social functioning, relationships at home or with friends and peers, or academic functioning (Canino et al., 1999). It is thus possible that some of the reviewed studies did cover a broader range of outcomes at the item-level. In conducting a similar exercise for overall health outcomes, Dodd and colleagues (2018) mapped so-called composite outcome indicators to *all* the taxonomy domains they pertained to. A similar approach could have added value to this study, in providing a more fine-grained mapping of outcomes covered at the item or sub-scale level of multidimensional measures. However, given the high degree of fragmentation, mapping the item-level content of the 121 measures identified would have constituted a separate exercise, which was beyond the scope of this review, but could illuminate further inquiries. A systematic mapping of functioning dimensions covered by different functioning measures, for example, would appear particularly useful, as would the mapping of specific symptoms covered by depression measures. This would complement similar mappings that have recently been undertaken for adult depression measures (Fried, 2017).

Fourth, it is worth highlighting that several of the reviewed studies linked to larger clinical trials. For instance, 17 studies used data from the TADS (TADS Team, 2004) and 14 studies used data from the TORDIA (Brent et al., 2008) study. Even though each of these studies answered different research questions and used different combinations of measures, all used the same primary outcome measures, which means that the relevant studies in the review sample are not entirely independent from one another. The rationale for proceeding in this way was that these trials had often administered a range of secondary measures that were not reported in the primary publications, and often only analysed and reported as part of secondary studies looking at moderators or predictors of treatment response. The approach taken allowed to consider these secondary studies and the outcomes they covered. In addition, the review aimed to take stock of the outcomes prioritised and reported across the recent literature, and the decisions of study authors to include or exclude secondary outcomes from their secondary analyses of trial data were considered informative in this context.

4.5 Conclusions

As this review demonstrates, the recent treatment effectiveness literature for adolescent depression focuses mainly on change in symptoms and functioning as reported by clinicians, at the expense of exploring change across a broader range of domains and informants. The current momentum around person-centred and youth-guided care provides a historical chance to accelerate multi-dimensional and multi-informant measurement and to reconsider what is important to measure for different stakeholders, including young people themselves. At a time when there is more scrutiny on treatment outcomes for adolescent depression than ever before, making sure that outcome measurement reflects what matters most to service users is paramount. The extent to which the focus on symptoms and functioning revealed by this review matches the outcome priorities of youth, parents, and clinicians will be the focus of the next two studies, presented in Chapters 5 and 6.

Chapter 5. Post-Therapy Outcome Perspectives

5.1 Introduction

5.1.1 Background

The Outcome Measurement Review presented in Chapter 4 suggested that over the past decade, outcome measurement in clinical research for adolescent depression has largely reflected a clinical understanding of outcome, focussed on symptom reduction and functional impairment, consistent with previous systematic reviews that covered a broader range of developmental stages and presenting problems (Becker et al., 2011; Hoagwood et al., 2012; P. S. Jensen et al., 1996; Weisz et al., 2005). As discussed in Chapter 1 (Section 1.4), existing qualitative studies exploring outcome priorities amongst young people, parents, and clinicians suggest that such symptom-focussed measurement does not fully reflect what matters most to these stakeholder groups. There is evidence that youth, parents, and clinicians value a broader range of outcomes (Bergmans et al., 2009; Binder et al., 2011; Cortés et al., 2018; Dunne et al., 2000; Freake et al., 2007; Gibson & Cartwright, 2014; Gibson et al., 2016; Lavik et al., 2018; Troupp, 2013), although they frequently disagree on which outcomes these are (Garland et al., 2004; Hawley & Weisz, 2003; Jacob et al., 2016; Yeh & Weisz, 2001).

This existing body of research is, however, limited and heterogenous, as described in more detail in Chapter 1. Only one study specifically examined outcome perceptions amongst depressed adolescents (Cortés et al., 2018), but used a small sample and provided limited account of the analytical approach used. Lavik and colleagues (2018), in contrast, provided an emic and thick description (Geertz, 1973; Morrow, 2005) of outcome themes based on semi-structured interviews and focus groups. However, the coding frame was not made available; the study authors did not discuss their own subjectivity and pre-existing assumptions; and negative or diverging cases were not examined. It is thus not clear whether the same attention has been given to all outcome themes, or whether preconceived assumptions may have biased analysis towards expected themes, such as those that were familiar from the recovery literature (see Section 1.3.2). The same strengths and caveats apply to Bergmans and colleague's (2009) grounded theory of recovery trajectories amongst young adults with a history of repeated suicide attempts. These existing studies have consistently suggested that symptom change does not constitute an outcome of primary importance in the accounts of young people.

Other studies have taken a more transparent and deductive approach to coding by applying (and often adapting) existing conceptual frameworks of outcome, treatment goals, or presenting problems (Bradley et al., 2013; Garland et al., 2004; Hawley & Weisz, 2003; Rupani et al., 2014; Yeh & Weisz, 2001). In contrast with the above-mentioned studies, these drew on 'thinner' data in the form of brief descriptions of treatment goals, desired outcomes or key presenting problems at the *start* of treatment, rather than detailed accounts of outcomes that were actually observed (Jacob et al., 2016). Their findings contradict those of the first set of studies, by suggesting that symptom change is, in fact, an important outcome category for young people. The evidence on this matter is thus inconclusive. To date, no research has joined the strengths of both approaches outlined above.

5.1.2 The Present Study

This study expanded on the existing literature by taking a systematic and conceptually grounded approach to investigating outcomes specifically for adolescent depression, based on detailed accounts of change processes provided by young people, their parents, and clinicians following the completion of a treatment trial for youth depression. The focus of this study was on providing a comprehensive mapping of outcomes described by these three groups, examining the salience of these outcomes across and between the three groups and three treatment arms, and comparing it to the salience of the same outcomes in the recent treatment effectiveness literature (as indicated by the Outcome Measurement Review).

The post-treatment perspective has been described as the most informative for investigating outcome perceptions, compared to studies conducted at the start or during treatment (Connolly & Strupp, 1996). To reduce the risk of omitting outcomes that may only be relevant to a minority, and to provide a comprehensive mapping that can be cross-referenced with the mapping undertaking in the Outcome Measurement Review, the outcome taxonomy developed in Chapter 1 was used as an initial coding framework. The study aimed to answer five research questions:

1. What outcomes do young people, parents, and clinicians discuss following psychotherapy for adolescent depression?
2. To what extent does the salience of outcomes in narrative accounts converge with their salience in the recent treatment effectiveness literature?
3. To what extent does the salience of different outcome categories diverge or converge between young people, parents, and clinicians?
4. To what extent do triads of young people, parents, and clinicians agree on the outcomes experienced in relation to individual cases?
5. To what extent does the salience of different outcome categories diverge or converge between the three treatment arms?

The existing studies that most resemble the present inquiry with regards to the study population, and type of data scrutinised suggest that symptom change may not constitute one of the most salient outcome themes, in line with the adult recovery literature (Bergmans et al., 2009; Cortés et al., 2018; Gibson & Cartwright, 2014; Lavik et al., 2018). It can further be hypothesised that levels of agreement between young people, parents, and clinicians, may be higher than those observed in previous studies, as data was collected following treatment rather than ex-ante, allowing triads more time to communicate and foster a common understanding of the changes achieved.

5.2 Method

5.2.1 Overall Design

The analysis presented in this chapter drew on data that was collected through the IMPACT-My Experience (IMPACT-ME) study (see Midgley et al., 2014 for full details of the study design and rationale), a qualitative study embedded in a pragmatic effectiveness superiority trial of psychotherapeutic treatments for adolescent depression.

The Improving Mood with Psychoanalytic and Cognitive Therapies (IMPACT) trial involved 15 statutory child and adolescent mental health services across England, and randomised 467 clinically depressed adolescents to one of three manualised treatments: a Brief Psychosocial Intervention (BPI), Cognitive Behavioural Therapy (CBT), or Short-term Psychoanalytical Psychotherapy (STPP; Goodyer et al., 2011). BPI involved psychosocial management over the course of 20 weeks, with up to eight individual sessions and up two four parent or family sessions. The focus was on psychoeducation, behavioural activation, problem solving, risk management and physical and mental hygiene (Goodyer et al., 2017; Kelvin et al., 2010). CBT involved 20 individual sessions over 30 weeks, plus up to four family or parental sessions, and focussed on identifying and challenging negative automatic thoughts and their linkages with behaviour, and developing more adaptive cognitive and behavioural techniques (Goodyer et al., 2017; Impact Study CBT Sub-Group, 2010). STPP consisted of a planned course of 28 sessions over 30 weeks with the option of parents or carers accessing additional sessions with a parent worker. Using a psychodynamic approach, the clinician encouraged young people to express difficult feelings and experiences, and guided them in making sense of these through a non-judgmental process (Cregeen et al., 2016; Goodyer et al., 2017, 2011; Midgley et al., 2017).

The nested IMPACT-ME study aimed to complement quantitative outcome assessment using nomothetic measures with the qualitative longitudinal exploration of change through semi-structured interviews. Participants were recruited from the five London-based trial centres and completed individual semi-structured interviews at the start of treatment, at the end, and at one-year follow-up. The present secondary analysis drew on interviews conducted at the end of treatment.

5.2.2 Participants

Young people qualified for participation in the IMPACT trial if they were aged 11 through 17 years at assessment, and had a current DSM-IV diagnosis of unipolar Major Depressive Disorder with moderate to severe functional impairment (Goodyer et al., 2011). Youth were excluded if they presented with generalised learning difficulties or a pervasive developmental disorder; a substance use disorder; or a primary diagnosis of bipolar disorder, schizophrenia, or eating disorder. Other exclusion criteria included pregnancy, the use of medication that could interfere with pharmacotherapy for depression, and having completed one of the study treatments in the past (Goodyer et al., 2017).

Of all youth (*N* = 80) who participated in the trial in the greater London area, two did not complete the qualitative study because of time constraints and one withdrew from the study, leading to a final qualitative sample of 77 (Midgley et al., 2015). Interviews with all three members of a triad (i.e., young person, parent, and clinician) at the end of treatment were available for 40 cases. In five of these, the young person did not remain in therapy for more than three sessions and the interviews provided limited information on outcomes achieved during this period. In one other case, the young person was referred to inpatient care for several months and most outcomes discussed referred to this experience rather than outpatient treatment as part of IMPACT. These six interviews were excluded from analysis. The final sample consisted of 34 triads, involving 102 individual interviews. Young people were aged 12-19 years at the time of the interview with a mean age of 16.2 (*SD* = 1.5) years, and 21 (61%) were female. Nine young people had been treated in the BPI arm, nine in the CBT arm, and 16 in the STPP arm.

5.2.3 Data Collection Method

By using qualitative methodology, the IMPACT-ME study encouraged participants to provide in-depth accounts of their experiences of therapy in their own words (McLeod, 2011). Semi-structured interviews were conducted individually with young people, one or both parents, and the clinician after therapy had ended, using the Experience of Therapy Interview guide (Midgley et al., 2011). The guides were tailored to each respondent group with regards to question wording, but were similar in content. Participants were asked about their experience of therapy, any changes they had observed since the start of treatment, any helpful or unhelpful aspects of therapy, and any outstanding moments or turning points. Interviews were conducted by post-doctoral psychologists, recorded, and transcribed verbatim for analysis.

5.2.4 Data Analysis

The present study aimed to describe and comprehensively map outcomes discussed by young people, parents, and clinicians following treatment, and to compare the salience of outcome themes between the groups, and in relation to their measurement in the recent treatment effectiveness literature (as explored in Chapter 4). The goal was not to undertake an in-depth, thick examination of individual change narratives, which would have required reducing the analytical sample, at the risk of identifying a narrower set of outcomes.

To meet these objectives, data was analysed using qualitative content analysis (Elo & Kyngäs, 2008; Hsieh & Shannon, 2005), which typically involves “the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns” (Hsieh & Shannon, 2005, p. 1278). In contrast with regular content analysis, qualitative content analysis moves beyond the counting of specific words or expressions and involves the careful examination and interpretation of language, and narratives. These are then coded thematically and systematically based on patterns of shared meaning (Weber, 1990). Meanings can be explicit or inferred, and categories may be derived inductively from the data, or from a pre-existing framework. As such, the initial coding resembles thematic analysis (Braun & Clarke, 2006). However, qualitative content analysis ends by coding the data into categories and quantifying their occurrence, whereas thematic analysis proceeds to aggregating these categories into higher-level themes, by examining relationships and patterns of meaning amongst them (Crowe et al., 2015). Qualitative content analysis is particularly suited for systematically condensing a phenomenon of interest into a conceptual system (Elo & Kyngäs, 2008), and as such, was the method of choice for this study.

In this study, the initial thematic coding was conceptually guided by the taxonomy developed in Chapter 3, to ensure equal attention to the full range of possible outcome concepts, and to promote consistency in terminology and categorisation with the Outcome Measurement Review presented in Chapter 4. However, this preliminary taxonomy presented only a starting point and was iteratively and inductively revised to reflect themes and patterns emerging from the data. Categories were created, modified, merged, or removed as necessary until a point of saturation was reached. The final coding frame included eight higher-level outcome domains: *symptoms*, *self-management*, *functioning*, *personal growth*, *relationships*, *therapy process*, *wellbeing*, and *parental support and wellbeing*. Within these domains, 31 outcome categories were identified. The final coding frame is displayed in Table C.1 (Appendix C) and displays definitions of each code and illustrative quotes. The specific changes made to the taxonomy in the course of this study are discussed in Section 5.4.1.

The number of interviews mentioning each outcome domain and category were counted. Frequencies reported in the findings section relate to the number of participants mentioning an outcome at least once, relative to the full number of participants in the reference group (i.e., this may either be the full sample, or the relevant subgroup). Agreement between dyads and triads of young people, parents, and clinicians was also examined. Beyond the quantitative reporting of outcome frequencies, narrative interpretations of the most commonly mentioned outcome categories are also provided to generate a richer understanding of what these meant to study participants.

To assess the significance of the difference in outcome frequencies between the Post-Therapy Outcome Perspectives study (PTOP) and the Outcome Measurement Review (OMR), a chi square statistic was computed that compared the frequencies at which outcomes were mentioned in participant narratives (defined here as the *observed values*), with the frequency of their measurement in the recent treatment efficacy and effectiveness literature as per the OMR (defined here as the *expected values*). For each row (i.e., outcome domain or category), the expected frequencies of the OMR were subtracted from the observed frequencies of the PTOP; the difference was squared; and divided by the expected frequency ((Observed – Expected)^2 / Expected). This was repeated for each row, and the results summed across all outcome domain rows, and all outcome category rows, respectively. The sum provided the chi square statistic, which was compared to the critical value for a 0.001 probability, at 7 degrees of freedom for the outcome domains, and 23 degrees of freedom for the outcome categories[3]. Outcomes that had no observation in the OMR were excluded. The resulting critical values were 24.32 (*df* = 7, *p* = 0.001) for the comparison of outcome domains, and 49.73 for the comparison of outome categories (*df* = 23, *p* = 0.001).

5.2.5 Ethical Considerations and Approval

The original study protocol for the IMPACT trial was approved by Cambridgeshire 2 Research Ethics Committee, Addenbrookes Hospital Cambridge, UK (REC Ref: 09/H0308/137). All participants above the age of 16 provided informed written consent. Parental consent and youth assent were obtained for younger adolescents. In order to ensure confidentiality, the interview data were anonymised, and any identifying details removed. Young people's names have been replaced with pseudonyms.

5.2.6 A Note on Reflexivity

Even though the doctoral candidate had no direct involvement in conducting the semi-structured interviews underpinning this analysis, her training and expectations may have influenced her approach to data coding and interpretation (Morrow, 2005; Yardley, 2000).

This study formed part of a doctoral thesis in evidence-based child mental health research. Even though the candidate is now based at a Psychology department, her previous degrees were in sociology. This social science background may express itself in a less clinical perspective on the data than a fully trained clinical psychologist might have applied. For example, a trained psychologist may have assigned a broader range of outcome concepts to the higher-level domains of *symptoms* and *functioning*, while the candidate favoured assigning them to the domains of *self-management*, *relationships,* and *wellbeing*, thereby constructing a less clinically focussed coding frame. In addition, the candidate's interest in multidimensional outcome measurement and ambition to identify a wide range of possibly relevant outcomes may have led to a focus on distinguishing distinctive outcome concepts rather than seeking commonalities for the purpose of aggregation.

A further point worth noting is that the doctoral candidate did not approach the data through a theoretical framework linked to a particular therapeutic approach. The resulting analysis had more descriptive than interpretive features, as the candidate attempted to remain faithful to the meanings conveyed by the participant narratives, and to represent these in the participants' own words (although code names were not entirely constructed in vivo, as they were conceptually informed by the outcome taxonomy developed through the Outcome Taxonomy Review). A researcher approaching the data through a more theoretical lense,

---

[3] Degrees of freedom were calculated as follows: (number of columns minus one) x (number of rows minus one). With two columns each, and 24 outcome categories that had at least one observation for each the PTOP and OMS, and 8 outcome domains, this resulted in 23 and 7 degrees of freedom, respectively.

linked to a specific therapeutic approach, may have yielded a greater depth of interpretation and theory-led conclusions (Timulak & Creaner, 2010).

Ideally, the credibility of this analysis would have been strengthened by having a co-analyst replicate the coding independently, and by validating the emerging categories with the original study participants (Morrow, 2005). Unfortunately, the doctoral candidate gained access to the data several years following the end of the study, and had no ethical clearance to re-contact participants for the purpose of validation. Co-analysis by a second rater was made difficult by the large volume of data (which involved over 100 interviews of roughly 30-60 minute-length) and the iterative process of devising the final coding frame. Co-analysis would have required substantial inputs from a second researcher, which would have made the candidate's contribution to the study less clearly distinguishable. However, emerging codes were discussed regularly with both PhD supervisors, and amended based on their feedback.

5.3 Findings

5.3.1 What Outcomes did Participants Describe at the End of Treatment?

The change narratives provided by young people, parents, and clinicians following their participation in the IMPACT trial touched upon a wide range of different outcomes across all eight higher-level outcome domains. This section first presents the most common themes across participant groups, and subsequently discusses differences in views and priorities.

On average, each study participant discussed outcomes in relation to four outcome domains, mentioning six outcome categories. As shown in Table 5.1 (below), the outcome domain most commonly mentioned across participant groups was *symptom change* (79%). However more than half of the participants across the sample also described changes in *functioning* (66%), *relationships* (62%), *self-management* (61%), *personal growth* (59%), and benefits related to the *therapeutic process* (53%). The only two domains that were less commonly mentioned were *wellbeing* (38%) and *parental support and wellbeing* (24%). The top five most mentioned specific outcome categories were *mood and affect* (65%), improved *coping and resilience* (50%), changes in *family functioning and relationships* (49%), the value of the *therapeutic space* (47%) and improved *academic and vocational functioning* (45%). Outcomes discussed less frequently, but still by more than one in five participants related to *social functioning* (35%), *confidence and self-esteem* (33%), *friendships* (27%), *parental wellbeing* (22%).

Within the outcome category of *mood and affect*, participants described young people feeling less low, being less prone to mood swings, less withdrawn, and more cheerful. Clinicians in particular often described such changes using clinical terms such as 'depression', 'symptomatology', 'cure' or 'recovery', whereas young people and parents tended to use more colloquial terms, such as 'being happier' and 'enjoying life more'. Some participants described that over the course of therapy, young people went back to being the person they used to be, or that they had become an entirely different person from their former depressed self. For example, one young person explained: “If I hadn’t gone there, I think I’d be a completely different person like I’d still be really low” (Natalie, 15 years, STPP; see Table C.1 in Appendix C). Others described that low mood and negative affect were still present, but more fleeting, and less overwhelming, which was often linked to young people leaning to cope more effectively.

Within the theme of *coping and resilience*, participants described improvements in young peoples’ ability to cope, which involved applying techniques learned during therapy (e.g., breathing or counting exercises, or keeping thought diaries), or identifying personal strategies for managing feeling and thoughts (e.g., allowing themselves to cry when they felt sad, rather than letting feelings build up). Participants further described that bringing these skills into their daily lives strengthened young peoples’ self-efficacy and sense of control, and made them more resilient and able to face life’s ups and downs. Another aspect of coping was young people having a better understanding of their feeling and thoughts, and how these might link to their behaviour. This enabled them to identify triggers, anticipate challenging situations, and manage them more effectively.

> It did wake me up to how my-, sort of how it all works and like how my brain works […] the fact that if you can understand something you can fix something that’s my motto. So, if I can understand like in a computer game if I can understand why it’s not working, I can fix the problem. (Dylan, 16 years, STPP)

Narratives about change in *family relationships and functioning* were multifaceted. Some young people reviewed and adjusted their role within the family system, for example by learning to impose boundaries between their own needs and those of their parents and siblings. In some families, parents, and young people grew closer as they learned to speak more openly about their difficulties. Some families experienced a decrease in conflict as young people learned to cope more effectively and family members grew more understanding. In contrast, others felt that therapy had taught them to tolerate a ‘healthy’ amount of conflict. Some young people were able to clarify a fraught relationship with a specific family member, by processing resentment during therapy and learning to interact with them differently:

> I know it sounds weird, but I can hold a good conversation with [stepfather] now. And, even...I had so many issues and I blamed them all on him […] ...and I kind of realise now that it wasn’t his fault and it’s never really been an issue with him just the fact that out of all the things that were going wrong, he was the one thing which was...I could blame everything on. And it’s realising that and it’s knowing that it’s, not his fault that have made it like seem easy to talk to him now and I have a really good relationship with him now and it makes everything so much easier. (Ella, 15 years, BPI)

Within the theme of *academic and vocational functioning*, participants discussed changes in young peoples’ attendance, commitment, and performance at school or college. Improved attendance involved young people missing fewer hours or days of school, or returning after a sustained period of leave. Frequently, participants also described young people being better able to motivate themselves and commit to their schoolwork, which often resulted in superior grades and exam results.

As part of their IMPACT-ME interview, young people and parents were asked which aspects of therapy they found particularly useful, and whether they would recommend therapy to friends. Many young people explained that therapy was a space where they felt safe, listened to, and able to open up, which was also frequently endorsed by parents. Young people valued the opportunity to speak about their feelings and thoughts without being judged. They appreciated the confidentiality of the space, as well as the clinician's position as an "outsider", which helped them "think about things differently."

Table 5.1. Prevalence of Outcome Categories in the PTOP and OMR

Outcome domain and category	PTOP participant narratives (*n* = 102)			OMR included studies (*k* = 92)		
	n	%	Salience [a]	*k*	%	Salience [a]
A. Symptom change	**81**	**79%**	**typical**	**86**	**93%**	**general**
A.1. Mood & Affect	66	65%	typical	86	93	general
A.2. Anger and aggression	17	17%	rare	4	4%	rare
A.3. Eating and weight	13	13%	rare	1	1%	rare
A.4. Sleeping and energy	20	20%	rare	2	2%	rare
A.5. Self-harm	12	12%	rare	1	1%	rare
A.6. Suicidality	14	14%	rare	15	16%	rare
A.7. Anxiety	12	12%	rare	7	8	rare
A.8. Other comorbid issues	4	4%	rare	9	10%	rare
B. Self-management	**62**	**61%**	**typical**	**14**	**15%**	**rare**
B1. Behavioural activation	20	20%	rare	4	4%	rare
B2. Coping and resilience	51	50%	typical	2	2%	rare
B3. Cognition and behaviour	19	19%	rare	9	10%	rare
C. Functioning	**67**	**66%**	**typical**	**51**	**55%**	**typical**
C1. Global functioning	9	9%	rare	48	52%	typical
C2. Executive functioning	20	20%	rare	2	2%	rare
C3. Academic / vocational funct.	46	45%	variant	0	0%	rare
C4. Social functioning	36	35%	variant	3	3%	rare
D. Personal Growth	**60**	**59%**	**typical**	**7**	**8%**	**rare**
D1. Assertiveness	13	13%	rare	1	1%	rare
D2. Autonomy and responsibility	16	16%	rare	1	1%	rare
D3. Identity	14	14%	rare	6	7%	rare
D4. Processing past & present	18	18%	rare	0	0%	rare
D5. Confidence and self-esteem	34	33%	variant	1	1%	rare
E. Relationships	**63**	**62%**	**typical**	**4**	**4%**	**rare**
E1. Ability to talk	13	13%	rare	—	—	—
E2. Family funct. & relationships	50	49%	variant	4	4%	rare
E3. Friendships	28	27%	variant	1	1%	rare
E4. Other peer relationships	7	7%	rare	1	1%	rare
F. Therapeutic process	**54**	**53%**	**typical**	**9**	**9%**	**rare**
F1. Therapeutic space	48	47%	variant	2	2%	rare
F2. Practical or systemic progress	9	9%	rare	—	—	—
G. Wellbeing	**39**	**38%**	**variant**	**7**	**8%**	**rare**
G1. Peace of mind	14	14%	rare	—	—	—
G2. Optimism and hope	12	12%	rare	—	—	—
G3. Future orientation	19	19%	rare	—	—	—

H. Parental support & wellbeing	**24**	**24%**	**variant**	**3**	**9%**	**rare**
H1. Parental support	7	7%	rare	—	—	—
H2. Parental wellbeing	22	22%	variant	2	2%	rare

Note. Funct. Is short for functioning. PTOP stands for Post-Therapy Outcome Perspectives study. OMR stands for Outcome Measurement Review.

[a] The salience of each outcome was defined as follows: Outcomes were demarcated as 'general' if mentioned by at least 90% of participants in the reference group; as 'typical' if mentioned by at least 50% but less than 90% of participants; as 'variant' if mentioned by at least 20% but less than 50% of participants; and as 'rare' if mentioned by less than 20% of participants (De Smet et al., 2019; Knox et al., 2006). — indicated that an outcome was not identified.

Due to these qualities, young people (and parents) felt that they could "get things off their chest" as opposed to "bottling it all up." For some young people, therapy was the first experience of being at the centre of somebody else's attention, and of feeling truly listened to:

> Back then I felt like nobody cared about me and I don't I think it made me feel good within myself because it was just it's kinda what I needed like to feel like someone … (breathes out) does care and that like they are there for me and like after that like I don't know it just it made me realise that like everything I thought was just in my head, like nobody cares about me and stuff like that, like it was kinda like she proved to me that she does care about me now. (Natalie, 15 years, STPP)

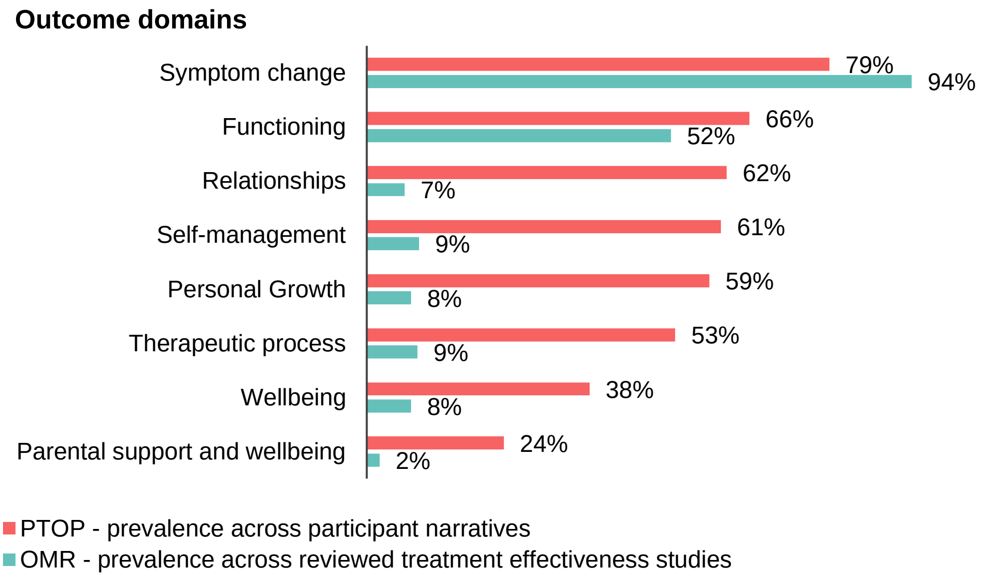
While these experiences constitute procedural aspects and facilitators of change, rather than changes in the young person or their lives, they were described as so transformative by several young people and parents, that being able to have these experiences might be considered an outcome in its own right. For this reason, two outcome categories related to the therapeutic process were included here alongside categories designating treatment outcomes.

The nine outcomes commonly mentioned across the full participant sample (i.e., by more than one in five participants) mapped across seven out of eight outcome domains, which illustrates the multidimensional nature of participants' change narratives.

5.3.2 Comparing Outcome Salience in the PTOP and OMR

The high salience of the *symptom* and *functioning* domains in participants' change narratives matched their frequent reporting in recently published treatment efficacy and effectiveness studies (see Chapter 4, and Figure 5.1, below). In this study, 79% of participants discussed changes within the *symptom* domain and 66% discussed changes in *functioning*. In comparison, across the 92 studies included in the Outcome Measurement Review, 94% measured change in the *symptom* domain and 53% measured change in *functioning.* In addition, however, study participants typically also discussed change in other domains such as *self-management*, *personal growth*, *relationships*, and *therapeutic process*. In the reviewed treatment efficacy and effectiveness studies, these outcomes were rarely measured. When comparing the frequencies observed in the PTOP with the frequencies expected based on the OMR, a chi square of 1959.70 was obtained, which indicated a highly significant difference between the two distributions ($df = 7$, $p = 0.001$).

Figure 5.1. Salience of Common Outcome Domains in the PTOP and OMR

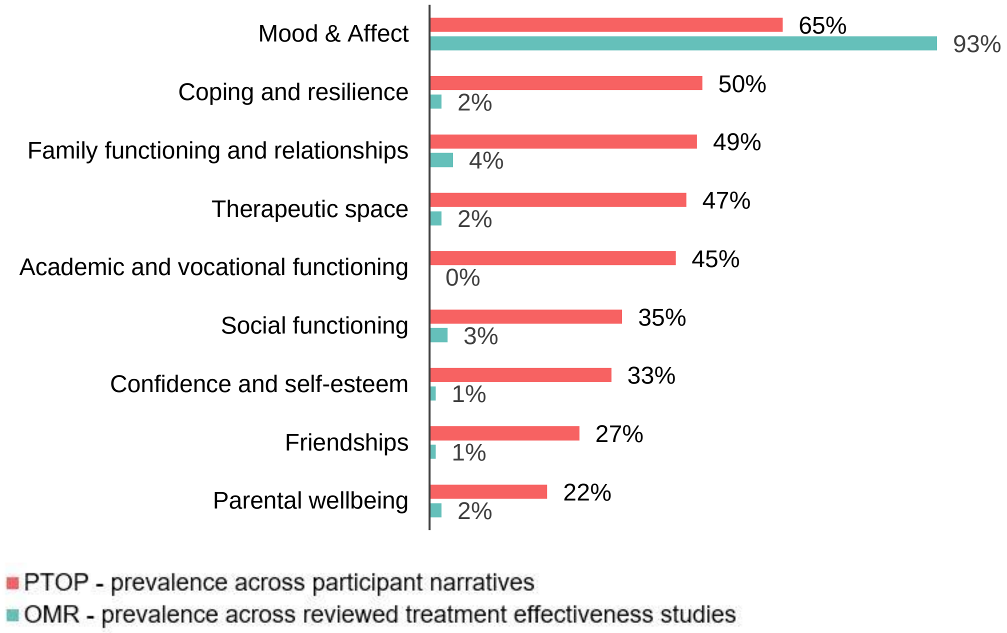


Note. PTOP stands for Post-Therapy Outcome Perspectives study. OMR stands for Outcome Measurement Review. Red bars indicate the percentage of study participants (i.e., young people, parents, and clinicians) in the PTOP study mentioning the respective outcome domain in their change narratives. Turquoise bars indicate the percentage of treatment effectiveness studies reviewed as part of the OMR that reported an outcome in the respective domain.

Within the domains of *symptoms* and *functioning*, there were differences in focus. In this study, between ten and twenty percent of participants discussed changes in secondary symptoms such as *anger and aggression*, *eating and weight*, *sleeping and energy*, or *self-harm*. In contrast, these were explicitly reported by less than five percent of the reviewed studies, although relevant items may have been included in depression symptom measures. Many treatment effectiveness studies reported on changes in *global functioning*, that is functioning across different life domains. This was frequently assessed through the Clinical Global Impression Scale (CGI; Busner & Targum, 2007) or the Children's Global Assessment Scale (Dyrborg et al., 2000), both of which are single-item tools gauging the clinician's overall impression of the young person's functioning. In IMPACT-ME , participants rarely discussed functioning in global terms, but tended to discuss specific changes in relation to academic, social, or executive functioning.

Of the nine specific outcome categories mentioned at least by one in five participants in this study, only *mood and affect* was also generally reported in recent treatment efficacy and effectiveness research. All other outcome categories were reported by less than five percent of the studies included in the Outcome Measurement Review (see Figure 5.2, below). For example, while half of all participants mentioned changes in coping skills and resilience, these were measured by only two of the 92 reviewed studies. Academic and vocational functioning were mentioned by 45% of study participants, but not reported in any of the reviewed studies. When comparing the frequencies observed in the PTOP with the frequencies expected based on the OMR, a chi square of 8358.28 was obtained, which indicated a highly significant difference between the two distributions (*df* = 23, *p* = 0.001).

Figure 5.2. Salience of Commonly Mentioned Outcome Categories in the PTOP & OMR



Note. PTOP stands for Post-Therapy Outcome Perspectives study. OMR stands for Outcome Measurement Review. Red bars indicate the percentage of study participants across the full sample who mentioned the respective outcome in their change narratives. Turquoise bars indicate the percentage of treatment effectiveness studies for adolescent depression included in the Outcome Measurement Review that reported the respective outcome.

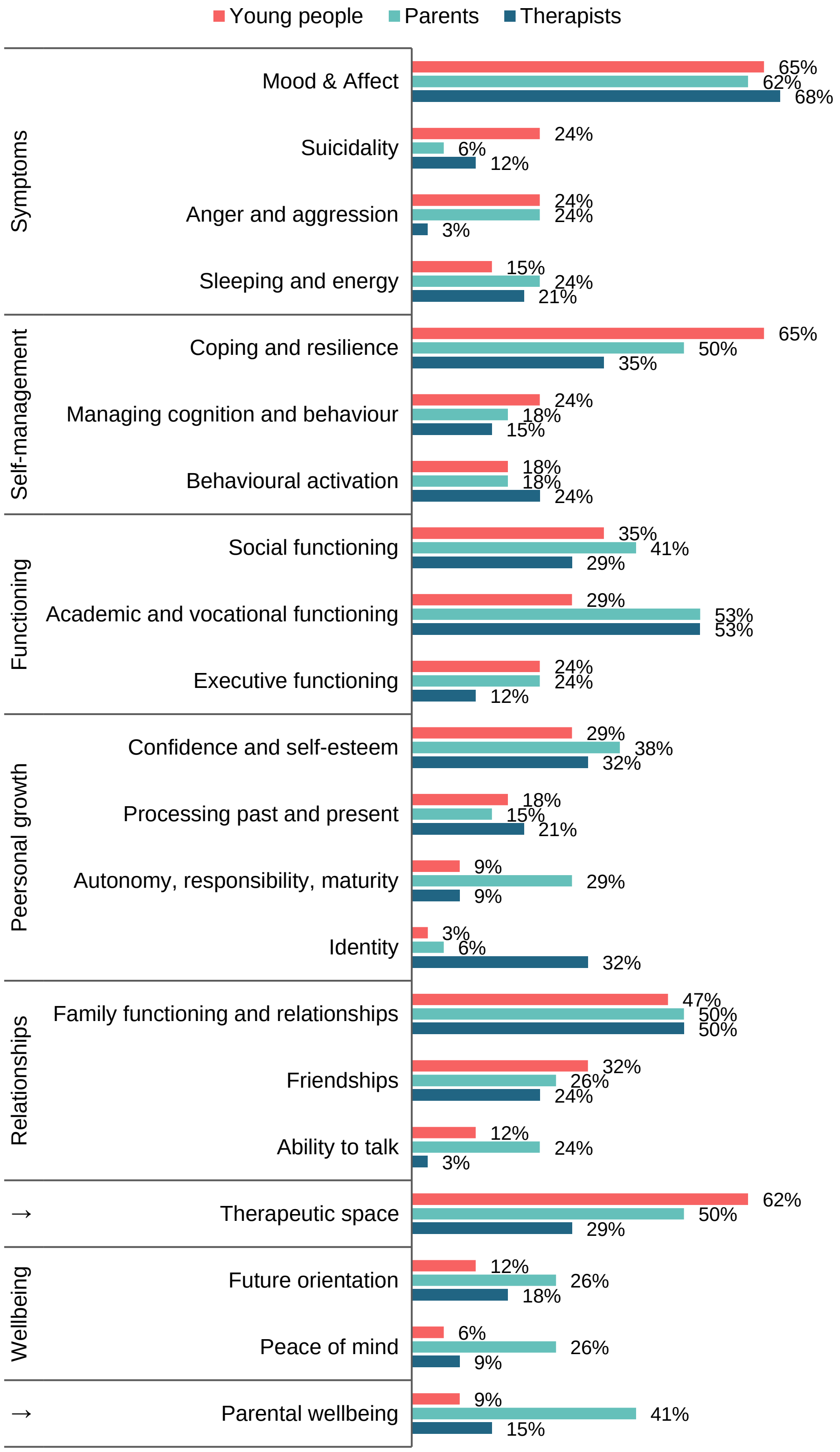
[a] Outcomes were classified as “commonly mentioned” if discussed by more than 20% of participants in the full sample.

5.3.3 Convergence of Outcome Perspectives Between Groups

Figure 5.3 (below) shows the outcome categories that were mentioned by at least one in five participants, in at least one of the three participant groups, and displays their salience in each group. Amongst young people, the most typically discussed outcome categories were *mood and affect* (65%), *coping and resilience* (65%), and *therapeutic space* (62%), each discussed by close to two thirds. A number of additional categories were discussed by at least one in five young people, including *family functioning and relationships* (47%), *social functioning* (35%), *friendships* (32%), *academic and vocational functioning* (29%), *confidence and self-esteem* (29%), and *suicidality*, *anger*, *managing cognition and behaviour*, and *executive functioning* (24% each).

Outcomes related to *mood and affect* were also the most discussed category amongst parents and clinicians, mentioned by around two thirds of participants in both groups. Similarly, improvements in *family functioning* were a common outcome theme amongst all three groups, with around half the participants in each group discussing change in this category.

Figure 5.3. Common Outcomes ^{a} by Participant Group



Note. Percentages are based on the number of participants in each respective subgroup mentiniong an outcome in their change narratives.

[a] Outcomes were defined as “commonly mentioned” if mentioned by more than 20% of participants in the relevant subgroup.

Social functioning, and *confidence and self-esteem* were consistently discussed by roughly 30% to 40% of participants in each of the three groups, although parents mentioned both outcomes more frequently than young people or clinicians. Outcomes related to *anger and aggression* and *executive functioning* were discussed by around one in four young people and parents, but only rarely by clinicians (3% mentioned anger, and 12% mentioned executive functioning).

While *coping and resilience* was the most commonly discussed outcome category (alongside *mood and affect*) amongst young people (65%), it was discussed less frequently by parents (50%) and clinicians (35%). Similarly, while 62% of young people described benefits of the *therapeutic space* in and of itself, these were also mentioned by 50% of parents, but by only 29% of clinicians. *Suicidality* was a third outcome category discussed considerably more frequently by young people (24%) than by parents (6%) or clinicians (12%). In contrast, outcomes related to *academic and vocational functioning* were discussed by more than half of parents and clinicians (53%), but by less than a third of young people (29%).

Outcomes discussed specifically by parents included young people's *autonomy, responsibility and maturity* (29%), *young people's ability to talk about their feelings and thoughts* (24%), as well as young people's *future orientations* and *peace of mind* (26%). Parents were considerably more likely than either young people (9%) or clinicians (15%) to discuss change in relation to *parental wellbeing* (41%). The one outcome discussed substantially more often by clinicians than by either parents or young people were changes related to the young person's *identity*. Clinicians often described that young people had developed a more balanced sense of self that could accommodate strengths and weaknesses, positive and negative feelings. While such changes were discussed by one in three clinicians, they were mentioned by just one young person and two parents (an overview of frequencies by participant group for all 31 outcome categories is provided in Appendix C, Table C.2).

Table 5.2. Dyadic and Triadic Convergence in Outcome Themes

Outcome domain	YP & PA		YP & CL		PA & CL		YP, PA, CL	
	%	*n* [a]	%	*n* [a]	%	*n* [a]	%	*n* [a]
A. Symptom change	64%	33	67%	33	71%	31	53%	34
B. Self-management	47%	30	45%	29	46%	26	28%	32
C. Functioning	50%	30	46%	28	60%	30	34%	32
D. Personal Growth	50%	26	50%	26	45%	29	30%	30
E. Relationships	65%	26	46%	28	40%	30	30%	30
F. Therapeutic process	46%	26	48%	25	57%	21	31%	26
G. Wellbeing	29%	21	11%	18	7%	27	3.5%	29
H. Parental support and wellbeing	12%	17	14%	7	11%	19	0.0%	19
Any domain	94%	34	85%	34	91%	34	77%	34

Note. YP = young person. PA = parent. CL = clinician. The table shows the percentage of dyads and and triads agreeing on at least one outcome within each outcome domain (considering only dyads and triads in which at least one participant mentioned an outcome within the relevant domain).

[a] The number of observations indicate the (denominator). It was defined as the number of dyads or triads in which at least one participant mentioned an outcome within the respective domain.

Of the 34 triads of youth, parents, and clinicians, 77% converged in discussing at least one common outcome following treatment. Convergence was higher amongst dyads, with 94% of youth-parent dyads, 85% of youth-clinician dyads, and 91% of parent-clinician dyads converging in discussing at least one common outcome domain (see Table 5.2). Across triads and dyads, convergence was strongest in the symptom domain, and lowest in the domains of parental support and wellbeing, and youth wellbeing.

5.3.1 Comparing Salience Across Treatment Arms

Across all three treatment arms, the most frequently discussed outcome was an improvement in *mood and affect*. Improved *coping skills and resilience* and *academic and vocational functioning* were the third and fourth most salient outcomes. *Family functioning and relationships* were the second most salient theme in CBT and BPI, but only the fifth-most salient in STPP. Instead, the second most discussed outcome theme in STPP was the benefit of the *therapeutic process*, which was the fifth-most salient in CBT and did not rank amongst the five most discussed outcomes in BPI. Here, improved *self-confidence and self-esteem* was the fifth most salient outcome.

Outcomes discussed more often in CBT than in the other two treatment arms included changes in *mood and affect*; *coping skills and resilience*; *managing cognition and behaviour*; *executive, academic, and social functioning; processing past and present experiences; friendships; therapeutic alliance; greater peace of mind*; and *feeling more optimistic and hopeful*. Outcomes discussed more often in STPP than in other arms included a reduction in *anger and aggression*, outcomes related to *personal growth* (e.g., assertiveness, autonomy, and identity), and improved *parental wellbeing*. Outcomes discussed more frequently in the BPI arm than in other arms included changes in *eating and weight*, *sleeping and energy*, *suicidality*, *anxiety*, and other comorbid symptoms, as well as *behavioural activation*, *global functioning*, *peer relationships*, *practical or systemic progress*, *future orientations and goals*, and *parental support* (see frequencies in Table C.2, Appendix C).

5.4 Discussion

5.4.1 Summary of Findings

This study had three aims: (a) to describe and comprehensively map the outcomes discussed by young people, parents, and clinicians following their participation in a treatment trial for adolescent depression; (b) to compare this with the frequency of their measurement in the treatment effectiveness literature; and (c) to assess the divergence or convergence in the salience of different outcome categories between young people, parents, and clinicians.

Young people, parents, and clinicians described a broad range of treatment outcomes across the domains of *symptom change*, *self-management*, *functioning*, *personal growth*, *relationships*, *therapeutic process*, *wellbeing*, and *parental support and wellbeing*. The most commonly discussed outcome categories across the three groups were *mood and affect*, *coping and resilience*, *family functioning and relationships*, *academic and vocational functioning*, and *therapeutic process*. Of these, only *mood and affect* was consistently measured in recent treatment efficacy and effectiveness studies (see Outcome Measurement Review, Chapter 4). In line with the initial hypothesis, levels of dyadic and triadic agreement were higher than those observed by studies comparing desired goals or outcomes at the start of treatment.

While a number of core outcome categories were salient across all three IMPACT treatment arms, the extent to which each of these was discussed, and the salience of other outcomes reflected different perspectives amongst young people, parents, and therapists, and varying priorities across trial arms, such as a focus on cognition and behaviour in the CBT arm, a focus on developing a healthy life-style (e.g., behavioural activation, restoring healthy sleeping and eating habits) in the BPI arm, and a focus on changes within the sense of self in the STPP arm.

Outcomes Frequently Described by Young People

Symptom change was the most prominent theme in young peoples' change narratives, at par with changes in coping and resilience. This contradicted findings of comparable studies about experiences of change and outcome conducted with adolescents following treatment (Bergmans et al., 2009; Cortés et al., 2018; Lavik et al., 2018). These studies suggested that youth did not primarily define outcome in relation to symptom change, but rather in the tradition of the adult recovery model (e.g., Leamy et al., 2011), by emphasising themes such as identity, hope, connectedness, and empowerment. This finding is, however, consistent with studies analysing goal or outcome data that was collected at the start of treatment using pre-existing conceptual frameworks to guide their analysis (Bradley et al., 2013; Garland et al., 2004; Jacob et al., 2016; Rupani et al., 2014). These found that outcomes related to symptoms, such as *managing mood, negative thoughts, and feelings;* or *reduced anger and aggression* were amongst the most desired outcomes or goals.

The precise reason for observing greater emphasis on symptoms than some previous studies is difficult to ascertain. This finding may point to the benefits of approaching similar inquiries with reference to a pre-defined conceptual framework. The latter can articulate the assumptions and terminology with which analysis has been approached, promote attention to all possible outcome categories, and help avoid bias from giving greater attention to outcomes frequently identified in the literature or valued within a certain theoretical framework. Another possible explanation is that these differences stem from the different treatment approaches employed across these studies (Connolly & Strupp, 1996). However, the specific approaches are rarely described in detail, which makes it difficult to explore this explanation further. Instead, they have been described as eclectic (Bradley et al., 2013; Jacob et al., 2016; Lavik et al., 2018), or in broad terms as 'psychotherapy' or 'counselling' (Cortés et al., 2018; Rupani et al., 2014). In this study, a strong focus on symptom change was consistently observed across all three treatment arms. A third explanation for the strong focus on symptom resolution in this study may be that this is particularly salient for youth with depression, because of the distressing nature of these symptoms. Youth with behaviour problems, for example, may experience less acute emotional suffering and therefore be less focussed on symptom change.

Other outcomes frequently discussed by young people were also identified by previous studies, including *learning to cope with challenges and emotions* (Bergmans et al., 2009; Lavik et al., 2018), *improved family functioning and family relationships* (Aarons et al., 2010; Garland et al., 2004; Hawley & Weisz, 2003), and the value of the therapeutic space and alliance, which included *being listened to and cared for without being judged* (Dunne et al., 2000; Freake et al., 2007; Gibson & Cartwright, 2014; Gibson et al., 2016). While Lavik and colleagues (2018) found that youth were generally describing 'good outcome' as the strengthening of autonomy and identity, these outcomes were rarely discussed by young people in this study. Autonomy was primarily discussed by parents, in line with findings by Garland and colleagues (Garland et al., 2004), while change in identity was primarily discussed by clinicians.

Outcomes Frequently Described by Parents and Clinicians

The study suggests that young people, parents, and therapists provide complementary accounts of change, focussing on different outcomes, to varying extents. This is in line with a large existing body of research about discrepancies in symptom and outcome ratings between different informants, as discussed in Chapter 1 (Section 1.5). In this study, parents and therapists demonstrated a greater interest in functioning than young people (especially in relation to academic and vocational functioning). In addition, parents were particularly concerned with young people's mental hygiene and lifestyle at home, as well as young people recovering peace of mind, becoming autonomous, developing aspirations for the future, and pursuing them. Similar themes have previously been identified by studies of parental goals, defined at the start of treatment (dosReis, Camelo Castillo, Ross, N'Dri, & Butler, 2018; Garland et al., 2004; Odhammar & Carlberg, 2015). While existing studies involving cross-diagnostic samples of children and adolescents have found that parents were concerned with managing youth behaviour (Hawley & Weisz, 2003; Jacob et al., 2016), this was not a dominant theme in this study, possibly due to its specific focus on adolescent depression.

Therapists demonstrated a particular concern with managing self-harm, but also frequently discussed changes in young people's self-image and self-confidence, which may reflect that close to half of the cases 34 were treated in the STPP treatment arm. Desired outcomes related to young people's sense of self were also frequently defined by therapists prior to psychodynamic psychotherapy as part of another study (Odhammar & Carlberg, 2015). This is in line with suggestions that therapists tend to pay particular attention to the intermediary outcomes or presumed treatment mechanisms linked to their specific theoretical treatment approach (C. E. Hill et al., 2013; Odhammar & Carlberg, 2015).

Overall agreement amongst young people, parents, and clinicians was higher than in previous studies, which had participants define desired goals or outcomes at the start of treatment (Garland et al., 2004; Hawley & Weisz, 2003; Jacob et al., 2016; Odhammar & Carlberg, 2015). In this study, looking back at the treatment experience, more than three quarters of triads and dyads described at least one common domain. Dyadic agreement was highest amongst youth and parents, and lowest amongst youth and clinicians.

5.4.2 Implications for Clinical Research and Practice

The observed mismatch between the multidimensionality of participant narratives on the one hand, and the predominance of unidimensional, symptom-focussed measurement in recently published treatment efficacy and effectiveness studies, calls for an urgent need to review measurement approaches in clinical research, as well as in clinical practice. As demonstrated by the Outcome Measurement Review in Chapter 4, primary outcome measures reported in clinical trials in particular tend to also be unidimensional with regards to the informants considered, as they typically draw on clinician report. They are frequently interpreted as global indicators of efficacy, although they do not necessarily reflect the perspectives of young people and parents (De Los Reyes et al., 2011). The Post-Therapy Outcome Perspectives study has shown that, young people and parents provide valid and complementary alternative accounts of change. To enable deeper and more nuanced insights into treatment efficacy and effectiveness, and the conditions under which change is observed by different informants in different outcome domains, conventions of defining primary outcomes around clinician-rated change in symptoms should be replaced with multidimensional and multi-informant approaches (The Lancet Psychiatry, 2020).

Two initiatives are currently under way, which aim to develop new standards for reporting outcomes across multiple domains and informants. So-called core-outcome sets (COS) recommend a battery of outcomes that should be tracked by all trials for a given disorder, or by all those providing relevant care in practice settings, as a *minimum* (M. Clarke & Williamson, 2015). Often, they also recommend a suite of relevant measurement tools. This study identified a range of outcomes that such efforts should consider, and demonstrates the importance of including young people, parents, and clinicians in the process. The fact that different outcomes were discussed to varying extents, across the three treatment arms underscores the importance of considering treatment mechanisms, when designing core outcome sets – for example, by including an explicit recommendation for researchers or services to add additional outcomes that are of theoretical importance (Fonagy, 1997; Kazdin, 1999a, 2009).

In person-centred clinical care, outcome measurement should mirror the concerns and perspectives of young people and families, and also cater to the information needs of clinicians and service managers. This requires an element of tailoring in accordance with individual or local needs, beyond the minimum suggested by core outcome sets. Services must balance the burden from administering additional questionnaires, with the need for tailoring. To help with this, questionnaires may be distributed between young people, parents, and clinicians, so that each party reports on the outcomes most relevant to them (both in terms of priorities, and in terms of chances for observing the relevant changes). For example, parents appear well placed to report on young people's eating and sleep hygiene at home, while young people may be better able to report on their use of coping skills and strategies, or their sense of self and wellbeing. Another important tool for tailoring measurement beyond a core set of outcomes while managing respondent burden, are idiographic measures that allow young people and parents to track change in relation to personalised indicators, such as target complaints or treatment goals (see a more detailed discussion of these measures in Chapter 1, Section 1.7.1).

5.4.3 Limitations

This study has a number of limitations. First, it aimed to provide a comprehensive mapping of outcome concepts and their salience across eight outcome domains and three participant groups to avoid the omittance of outcomes that may only be relevant to one group, or to a minority overall. For this purpose, a maximum number of eligible post-treatment interviews from the IMPACT-ME study were considered, and an analytic technique was chosen that is suitable for the systematic analysis of a large body of qualitative data. The deliberate prioritisation of breadth over depth comes at the expense of thicker and more emic inquiry into how individuals construct and understand outcomes, into the phasing and temporality of change trajectories, or into nuanced differences in the experiences and accounts of young people, parents, and clinicians. Future research could follow such lines of inquiry using a smaller subset of this large qualitative dataset, and an analytical approach more conducive to detailed narrative analysis, such as Interpretative Phenomenological Analysis to name just one example (Smith & Osborn, 2008).

Second, the data analysed were collected at the end of treatment and represented a snapshot in time. It is possible that participants' outcome narratives subsequently changed or evolved with time. Future analysis could draw on data from IMPACT-ME interviews conducted at the start of treatment, and at the one-year follow up post completion, to compare whether and how outcome themes evolved over time, and how actual changes discussed may relate to initial outcome expectations.

Third, it was not always possible to disentangle outcomes that had been brought about by therapy from changes that were caused by external factors. Any changes mentioned by participants as part of reflecting on their experience of therapy and evolving mental health were therefore coded, based on the understanding that the relevant *type of change* appeared to be meaningful regardless of whether or not therapy had succeeded in bringing it about.

Fourth, as qualitative research relies on narratives provided by participants, it risks favouring the voices of individuals who are particularly articulate and confident in sharing their thoughts and experiences, and risks missing voices of those less articulate or less comfortable with speaking to strangers. While even very short descriptions of change were coded, the study may still have failed to grasp the experiences of youth who provided very short responses. It is further worth noting that the participant sample was limited to young people, parents, and clinicians who lived in the Greater London area, and who may not be representative of young people in other regions or rural areas of the UK, or indeed, in other countries or regions of the world.

Fifth, a possible limitation of exploring outcomes following the end of treatment is that young people and parents may have been influenced by the values, terminology, or priorities conveyed by the clinician over the course of treatment (Connolly & Strupp, 1996). At the same time, the most commonly endorsed outcomes were largely similar across all three treatment arms, reflecting no clear influence of a specific treatment approach or philosophy (Connolly & Strupp, 1996). On a related note, however, it is also possible that the completion of symptom-focussed nomothetic outcome measure may have focussed the attention of young people and parents on symptom change. Indeed, as part of the quantitative outcome measurement within the IMPACT trial, young people completed a number of different symptom measures. They covered symptoms of depression, anxiety, obsessive-compulsive disorder, anti-social behaviour, suicide and self-harm, along with a measure of overall psychosocial impairment, and a measure of health-related quality of life (Goodyer et al., 2017).

Finally, this study did not examine difference in viewpoint and priorities *within* the three groups; that is, amongst young people, parents, and clinicians. As such, while mapping even rarely mentioned outcomes, it did not explicitly explore divergent experiences or perceptions. This will be the focus of the Viewpoints on Outcome Priorities study in the next Chapter.

5.4.4 Review of the Initial Outcome Taxonomy

Most categories within the initial taxonomy (devised through the Outcome Taxonomy Review and partly revised through the Outcome Measurement Review) were relevant and applicable to the accounts of change obtained from young people, parents, and clinicians. Existing categories within the domains of *symptoms* and *functioning* exhaustively accommodated the outcomes described by participants. Substance use, obsessive-compulsive symptoms and somatic complaints were rarely mentioned, and subsumed under “other comorbid difficulties” for the purpose of brevity. In relation to the category previously labelled “Behaviour and anger”, participants tended to emphasise anger and aggression over other aspects of behaviour (e.g., impulsivity). This resonates with a previous study that highlighted anger as a key theme in interviews conducted with IMPACT-ME participants at the start of treatment (Midgley et al., 2015). The category was thus relabelled to reflect this.

No new outcomes were identified in the domains of *relationships*, *personal growth*, and *parental support and wellbeing*. Several outcomes included in the preliminary taxonomy were not salient in participant narratives, including *meaning and purpose, recognising desires,* and *romantic relationships*, all of which were originally extracted from the adult-focussed Bern Inventory (Grosse Holtforth & Grawe, 2002) and appeared less relevant to participants in this study. Within the *self-management* domain, participants tended to describe young people's pro-active use of cognitive-behavioural strategies, which involved *challenging negative thought patterns and trying to approach situations differently*. The focus was thus more on deliberate cognitive-behavioural coping than on changes in cognitive patterns and processes, which are likely harder to evaluate for all three participant groups, as they are often observed using specialised tests and procedures (Fonagy, 1997).

Within the wellbeing domain, instead of calmness and composure, participants tended to describe *peace of mind*, *greater hope and optimism*, and *future orientations*. Changes in physical health were rarely discussed and might be subsumed under the domain of wellbeing. The domain of *therapeutic process* was added to account for the transformative impact that aspects of the therapeutic alliance (e.g., feeling heard, cared for, and understood) appear to have had for several young people. While these constitute facilitators of change, rather than person-level change in itself, these experiences shall be considered in Chapter 6 alongside treatment outcomes, to explore their relative importance in young people's judgements about what constitutes a 'good outcome'.

5.5 Conclusions

This study has highlighted that young people, parents, and therapists touch upon a variety of different outcomes, when reflecting on the changes observed over the course of psychotherapeutic treatment. Although *mood and affect* was the most salient outcome category, change was also commonly discussed in relation to *coping and resilience*, *family functioning and relationships*, and *academic and vocational functioning*. Young people and parents also frequency emphasised the benefit of feeling heard, cared for, and listened to in the therapeutic context. These core outcomes were salient across stakeholder groups and treatment modalities. But there was variation in the extent to which these were emphasised, as well as in the types of outcomes discussed beyond these core themes. New conventions for multidimensional outcome measurement are needed for clinical research and practice. These should reflect the multiple outcome domains and perspectives that have been highlighted here.

Chapter 6. Viewpoints on Outcome Priorities

6.1 Introduction

In the previous Chapter, the Post-Therapy Outcome Perspectives study identified a range of outcomes discussed by young people, parents, and clinicians when describing the changes observed or experienced over the course of therapy. It partly validated and partly added to the initial taxonomy developed in the Outcome Taxonomy Review. While a number of particularly salient and frequently described outcomes were identified, there was still a wide variety of other outcomes that were mentioned less frequently, but might be important to certain individuals, or subgroups of youth, parents, or clinicians.

Judgements of what constitutes a good or important outcome are influenced by considerations of what constitutes a ‘good life’ or a well-functioning child, which in turn is shaped by social and cultural norms, values, and expectations (Binder et al., 2010; Fonagy, 1997; Hoagwood et al., 1996). Research into adolescent goal setting suggests that young people have diverse aspirations that are shaped by socio-economic factors such as age, gender, family characteristics, and ethnicity, as well as the socio-political environment, cultural and gender norms (Massey et al., 2008). For example, one study has suggested that female high school students are more likely than their male peers to endorse personal goals around educational attainment and relationships, and that older students report more future trajectory goals than younger students (Massey et al., 2009). Based on this, it can be hypothesised that young people hold diverse views on what outcomes are important when accessing treatment for depression, even though little is known about such differences.

As discussed in Chapter 1, existing studies examining outcome perceptions and priorities amongst young people have generally identified dominant outcome themes, and have not investigated diverging or minority viewpoints (Bergmans et al., 2009; Cortés et al., 2018; Lavik et al., 2018). The exception is Gibson & Cartwright’s (2014) narrative analysis of young people’s experiences of school-based counselling. They identified four distinctive narratives: Profound changes to the self were described by youth demonstrating a “transformative” narrative; young people with a “supportive” narrative described counselling as mainly holding them in place; the “pragmatic” narrative underlined the utility of counselling in trying to resolve specific problems; and within the “disappointed” narrative, young people were unable to describe any changes caused by counselling. This points to considerable heterogeneity in young people’s experiences, which may be masked in qualitative inquiries attempting to draw a general picture of dominant themes and experiences.

Similarly, none of the studies examining differences in outcome priorities between youth, parents, and clinicians commented on differences *within* these groups (Garland et al., 2004; Hawley & Weisz, 2003; Jacob et al., 2016; Odhammar & Carlberg, 2015; Yeh & Weisz, 2001). Investigations of parental or clinician perceptions of outcome that provide equal depth as the youth-focussed studies by Bergmans and colleagues (2009) or Lavik and colleagues (2018) are lacking. It has been suggested that the changes considered meaningful by clinicians may vary by training and therapeutic approach, with one suggestion being that psychoanalytic psychotherapists may focus primarily on changes within the self, while cognitive behavioural therapists may focus more on symptom reduction (C. E. Hill et al., 2013). This was underscored by Odhammar and Carlberg (2015) who analysed goals defined by psychodynamic psychotherapists and found that these were highly reflective of this particular theoretical approach, with the most endorsed goals relating to children's "intrapsychic development, with a particular emphasis on attachment, empathy and reflexivity" (p. 284). The adult recovery literature, in turn, suggests that clinicians tend to adhere to symptom-focussed concepts of outcome, contrary to service users who tend to convey a more recovery-oriented perspective (Pilgrim & McCranie, 2013; Secker et al., 2002), further underscoring a sense of division between, and homogeneity in views within, these groups.

One recent study specifically examined heterogeneity in parental outcome priorities using latent class analysis (dosReis et al., 2018). In a sample of 346 parents, they identified four classes of priorities: The first two classes (making up 37% and 25%, respectively) were primarily concerned with issues relating to the child's safety behaviour towards themselves and other children; the third class focussed on the child's ability to be independent and to function well in their future adult lives; and the fourth class focussed on obtaining educational support and reduced risk of self-harm. Children had an average age of ten years, and 38% had an autism spectrum-disorder, which likely influenced the parents' principal concerns. However, the study still provides a case in point, showcasing that views and priorities are diverse.

A method that is tailored to the systematic study of distinctive viewpoints and priorities is Q-methodology (S. R. Brown, 1980; Stephenson, 1935). As part of a Q-methodological study, researchers assemble a set of items or statements that represent the broad discourse on a phenomenon of interest. Participants are then invited to sort these items according to a pre-defined ranking scheme (Stenner et al., 2003). The item configurations thus produced are then analysed using inverted ("by-person") factor analysis to explore patterns of similarity and divergence based on the correlation matrix. The technique moves beyond eliciting a majority view by aiming to identify a range of distinctive viewpoints, with special attention to minority experiences (S. R. Brown, 2006; Jedeloo et al., 2010).

Q-methodology has been deployed in a number of studies to explore diverse attitudes, experiences and preferences amongst young people in health or social care settings. For example, Jedeloo and colleagues (Jedeloo et al., 2010) examined preferences for hospital care amongst adolescent outpatients with chronic physical health conditions. They identified four preference profiles: *Conscious & Compliant, Backseat Patient, Self-confident & Autonomous,* and *Worried & Insecure*. The profiles differed with regards to young people's degree of autonomy, desire for self-management, therapeutic compliance and expectations about the clinician's role, leading the authors to conclude that there was "no 'one size fits all' approach to adolescent health care" (Jedeloo et al., 2010, p. 601). Other Q-studies involving adolescents have explored perceptions of antidepressant side effects (Cheung et al., 2003), reasons for treatment non-adherence in youth with life-threatening physical conditions (Bullington et al., 2007; Tielen et al., 2008), attitudes towards health-related lifestyles (van Exel et al., 2006), and family perceptions amongst youth in foster care (Ellingsen et al., 2011). To date, Q-methodology has not been deployed to examine outcome priorities amongst adolescents with a history of depression, or with mental health difficulties more generally.

Two relevant Q-studies have, however, been conducted with mental health professionals. Jackson-Blott and colleagues (2019) explored perceptions of recovery amongst staff and adult service users at a forensic mental health service. They identified four distinct viewpoints that respectively focussed on (a) *personal growth and psychosocial aspects of recovery*, (b) *gaining insight into symptoms and managing them more effectively*, (c) *self-reliance and life skills*, (d) and *finding redemption for past offenses, and reintegrating into the community*. A second study explored perceptions of recovery in borderline personality disorder amongst mental health professionals and identified two viewpoints (Dean et al., 2018). The first focussed on *symptoms*, *risk management,* and *coping*; the second focussed on *hope*, *personal growth*, *self-belief* and *acceptance*. Q-methodology has further been used with professionals to describe psychotherapy processes involving adolescents (Bambery et al., 2007; Calderon et al., 2017), but no study has yet examined outcome priorities amongst mental health professionals in relation to adolescent depression.

The study of heterogeneity within groups of youth or clinicians in relation to outcome priorities currently constitutes a gap in the literature. Existing qualitative studies have largely extracted one dominant voice, without exploring diversity in experiences and priorities. Existing studies are similarly limited with regards to exploring the relative importance of one outcome relative to another. The existing studies offering the 'thickest' description of outcome perspectives have not provided any indicators of their relative salience (Bergmans et al., 2009; Cortés et al., 2018; Lavik et al., 2018). In resource poor clinical contexts where respondent burden must be managed, there is a need to prioritise what is most important to measure in clinical practice, while still ensuring a person-centred approach, which covers those outcomes that matter most to service users. For these reasons, the study of diverging viewpoints on the one hand, and of outcome priorities on the other hand, can generate important insights for clinical practice.

6.1.1 The Present Study

The aim of this study was to explore viewpoints about what treatment outcomes adolescents with lived experience of depression, and mental health professionals providing support for depression consider most important. Rather than conducting open-ended interviews or focus groups as done by previous studies (Lavik et al., 2018), Q-methodology was used to (1) elicit the subjective viewpoints and preferences of participants, and (b) encourage participants to identify the outcomes they considered most important, while also considering the opportunity cost of deprioritising other outcomes in the process (Baker et al., 2006). This study aimed to answer the following research questions:

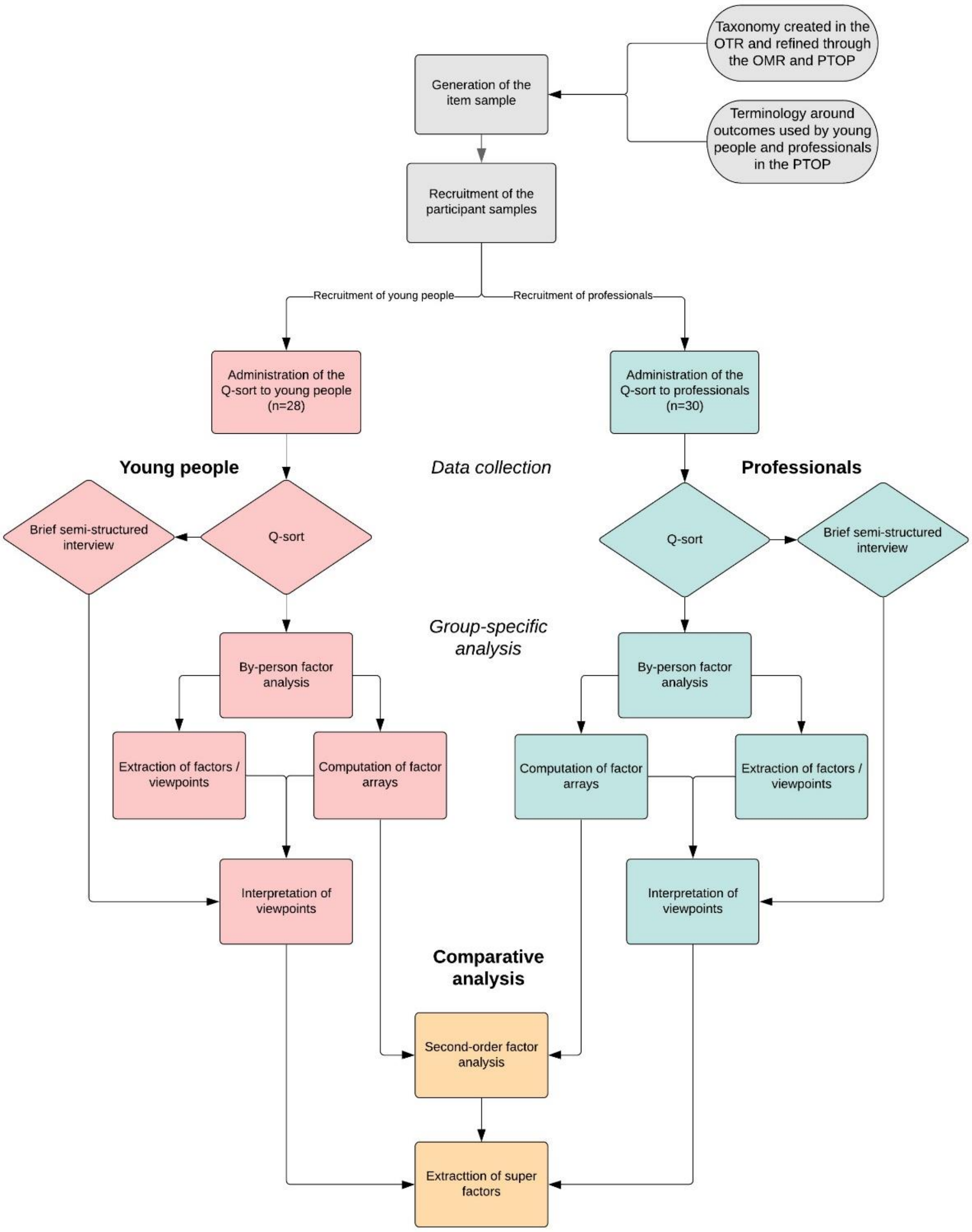
1. Which outcomes do young people and clinicians consider *most* important?
2. What distinctive viewpoints exist in relation to outcome priorities amongst young people and clinicians?
3. To what extent do the viewpoints held by young people and clinicians converge or diverge?

To answer these questions, two separate Q-studies were conducted amongst (a) young people and (b) mental health professionals, using the same Q-set and identical procedures for data collection and analysis (Watts & Stenner, 2012). Within each study, viewpoints and priorities were analysed in depth and in their own right. Next, the viewpoints extracted in each study were compared using second-order factor analysis (Kline, 1994), which is described in more detail below. The process is illustrated in Figure 6.1.

6.2 Methodology

Q-methodology has a number of strengths that have informed its choice for this study. First, it is specifically tailored to the study of subjective viewpoints and preferences, and makes diverse views both visible and “discussible” (Donner, 2001, p. 24). Second, it is considered a non-threatening methodology that lends itself well to participatory research with children and adolescents (Donner, 2001; Ellingsen et al., 2014; Owens, 2016). Contrary to in-depth interviews, Q-methodology does not require participants to provide coherent narratives of their thoughts and experiences, as narrative patterns emerge through the factor-analytical comparison of participants’ item configurations (Baker et al., 2006). Because analytic categories emerge directly from the participants' manipulation of the item cards rather than through intervention and interpretation by the researcher, Q-methodology has also been described as an approach that empowers participants’ voices in the research process (S. R. Brown, 2005, 2006; Donner, 2001). Lastly, Q-methodology allows for rigorous analysis of preferences and priorities with small samples, which is an important strength in light of barriers to recruiting large samples of young people due to requirements around parental consent and institutional gatekeeping (Heath et al., 2007).

Figure 6.1. Flowchart of the Q-study process and components



Note. OTR stands for Outcome Taxonomy Review (Chapter 3); OMR stands for Outcome Measurement Review (Chapter 4); and PTOP stands for Post-Therapy Outcome Perspectives study (Chapter 5). Elements shaded in grey relate to the study design and development of study materials; elements shaded in red relate to the youth sample; elements shaded in turquoise relate to the professional sample; and elements shaded in yellow relate to the comparative second-order factor analysis that drew on data from both participant samples.

6.2.1 The Item Sample

In Q-methodology, the item set to be sorted by participants is referred to as the Q-set. It represents a sample of items selected from a larger *concourse*, that is, the collection of all possible and relevant aspects and viewpoints concerning the topic under study (S. R. Brown, 2006; Ellingsen et al., 2010; Watts & Stenner, 2005). Assembling the concourse typically represents the first step in any Q-study, and the first step in developing the Q-set. In the case of this study, the concourse represented all outcomes that could possibly be relevant to young people and professionals in relation to adolescent depression. It was compiled successively through a number of initial workshops and consultations with young people, clinicians, and researchers; the Outcome Taxonomy Review; the Outcome Measurement Review; and the Post-Therapy Outcome Perspectives study. The initial concourse obtained through this process was structured into eight higher-level outcome domains, and 73 more specific outcome themes. It represented a more fine-grained version of the final coding frame produced as part of the Post-Therapy Outcomes Perspectives study.

It is recommended that a Q-set include between 40 and 80 items, but smaller sets may be advisable for use with children to ensure that the sorting process is engaging and enjoyable (Watts & Stenner, 2012). For this study, a Q-set of 35 items was compiled, which broadly corresponded with previous studies involving young people (Bullington et al., 2007; Ellingsen et al., 2011; Jedeloo et al., 2010; van Exel et al., 2006). Outcomes were selected from the concourse using Fisherian balanced block design, by selecting roughly equal numbers of items across all eight outcome domains (S. R. Brown, 1970, 1993). Between four and five outcomes were thus selected across the eight domains of *symptoms, self-management*, *functioning*, *personal growth*, *relationships*, *therapeutic process*, *wellbeing*, and *parental support and wellbeing.*

In order to cover as much of the concourse as possible while retaining a relatively small Q-set, similar outcome concepts were collapsed into one item (Watts & Stenner, 2012). For example, the items “being better able to concentrate” and “being more organised” were collapsed into one item – “Being better able to get things done (e.g., concentrate, be organized)” – to represent executive functioning. This was done by drawing on the terminology used by young people and clinicians in the IMPACT-ME interviews. In other cases, outcome themes that were aggregated into broader outcome categories in the Post-Therapy Outcome Perspectives study were maintained as more specific component items, to ensure an equal number of outcomes in each outcome domain. For example, the theme “parental support” was disaggregated into “My parents having a better understanding of me and my difficulties” and “My parents feeling more able to support me.”

An early draft of the Q-set was piloted with two young people with lived experience of service use for depression in 2017. Based on their feedback, redundant or ambiguous items were removed. One adolescent volunteer further reviewed the final Q-set for the clarity of the language and concepts used and further changes were made based on their feedback. Each item was then printed on a separate numbered card for sorting.

6.2.2 Participants

For conventional factor analysis, it has been suggested that the study sample should include at least twice as many participants as variables (Kline, 1994). In the context of a Q-methodological study, where the factor matrix is inverted, this would imply that the item sample should be at least twice as large as the participant sample (Watts & Stenner, 2005). At the same time, it is recommended to aim for significant factor loadings (i.e., significant correlation coefficients between individual participant Q-sorts and factors) for at least four to five participants on each extracted factor, with the number of factors extracted in Q-studies rarely exceeding six. If the extraction of up to six factors is anticipated, between 24 and 30 participants should thus be recruited. If the rule of having item samples twice the size of the participant sample was to be strictly followed, this would imply that the former would need to include at least 48 items. As suggested above, however, smaller item sets may be desirable in studies with children and young people. In light of such considerations, a less strict rule of thumb is usually applied whereby the number of participants should not exceed the number of items constituting the Q-set (van Exel & de Graaf, 2005; Watts & Stenner, 2012), with most Q-methodological studies involving samples of 20 to 50 participants.

A further point regarding sample size is of note. Q-methodological studies use qualitative, interpretative analysis to generate nuanced description of the viewpoints identified through inverted factor analysis. This sets them apart from conventional factor analytical studies (as does data collection through a card-sorting task, as opposed to more conventional survey designs). Large samples used in quantitative studies hinder such in-depth qualitative analysis, as the detection and explanation of subtle patterns and differences becomes increasingly difficult with increasing data volume (Watts & Stenner, 2005; Yardley, 2000). This is another reason why participant samples in Q-methodological studies rarely exceed 50 individuals. In light of the above-mentioned considerations, a recruitment target of 24-30 participants for each participant group (i.e., youth and clinicians) was set for the present study. Within this range, the principle of saturation was used, and after reaching 24 participants, recruitment was continued until no new viewpoints were articulated (Morse, 1995).

Participants were recruited using convenience sampling (Robinson, 2014). An open call for recruitment was advertised through the networks of the Anna Freud National Centre for Children and Families (AFNCCF) in London and other youth mental health charities, the University College London Psychology Subject Pool, as well as social media, and by soliciting peer support groups across England (advertisement materials are included in Annex D, Figure D.4). Snowball sampling (Biernacki & Waldorf, 1981) was used to further expand the reach of the recruitment call through word of mouth. Participants were not sampled according to pre-defined characteristics, as the limited evidence on outcome priorities did not enable the meaningful formulation of hypotheses about how such characteristics may influence viewpoints (Watts & Stenner, 2005). Instead, recruitment aimed to compose a diverse sample of participants that would reflect the different profiles, socio-economic circumstances, and needs of young people seeking support for depression in England, as well as different roles, placements and treatment approaches amongst mental health professionals.

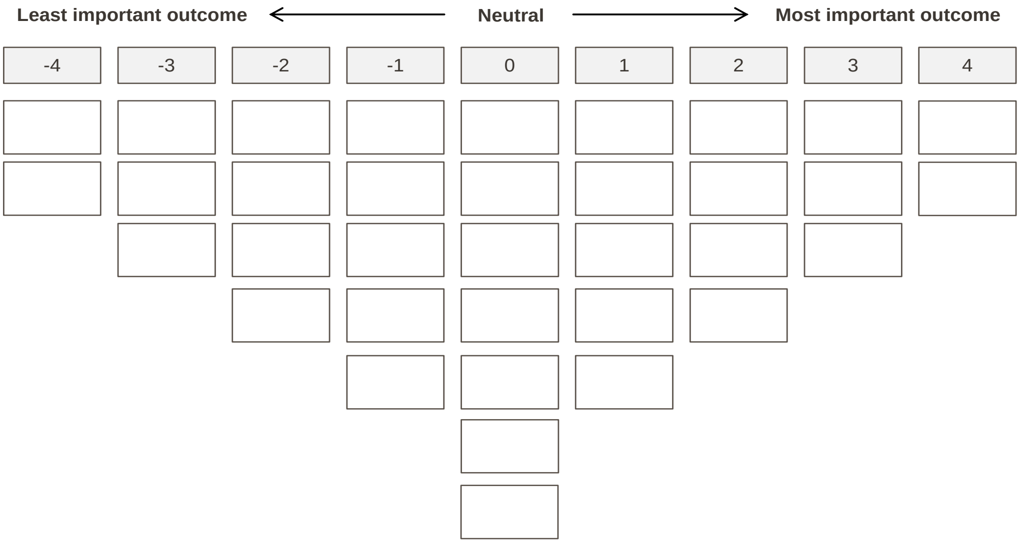
To recruit a diverse sample of young people, a broad set of inclusion criteria were defined. Any young person could participate who was aged 12 to 21 years, with current or past experience of accessing mental health support for depression. Depression did not have to be the only reason for seeking help, but had to constitute a primary presenting problem. Young people with comorbid difficulties were eligible, including young people with neurodevelopmental disorders (e.g., Autism spectrum disorder; ASD), if they were able to complete the research task in a self-directed manner. For the purpose of safeguarding, exclusion criteria included acute suicidal ideation or behaviour, or psychosis at the time of recruitment. Inclusion criteria for mental health professionals were equally broadly defined: any mental health professional with experience of providing treatment and support to depressed adolescents was invited to participate. No restrictions were placed on the type of mental health settings professionals worked in, their training, or role.

The study was originally intended to also include parents as a third participant group, thus mirroring the perspectives included in the Post-Therapy Outcomes Perspective study. However, as will be discussed in more detail in Sections 6.2.6 and 6.4.3, the recruitment of a sufficiently large parent sample was not achieved using the recruitment channels described above, and in hindsight, would have required a different recruitment approach.

6.2.3 Procedure

Participants were provided with the 35-item card deck, an A2-sized sorting grid, and a score sheet. They were instructed to begin by reading each card carefully, and by creating two piles of cards: one pile should include outcomes they considered important, while the other should include outcomes they considered unimportant. In a second step, participants were instructed to place all cards in the predefined sorting grid, along a 9-point scale of importance, from +4 (*most important outcomes*) to -4 (*least important outcomes*, see Figure 6.2). The middle category of zero marked a neutral placement. The sorting grid guided participants in ranking their items in the form of a quasi-normal distribution with a predefined number of items to be placed under each point of the scale. For example, participants could place two items each under the heading of “+4” and “-4”, three items in positions +3 and -3 and so on (see Figure 6.2, below). All items had to be distributed in this way, forcing participants to evaluate each outcome relative to all others. While Q-sorts have been conducted without restricting the distribution of the cards in this way, a forced distribution, facilitates data management and analysis for the researcher, and the sorting process for the participant (Watts & Stenner, 2005). Once participants were satisfied with their item configuration, they recorded the position of each item card in the score sheet provided (see Annex D, Figure D.5).

Figure 6.2. Sorting Grid Used by Participants



Note. The sorting grid provides 35 empty spaces – one for each item card.

The placement of the item cards was followed by a brief (5-15 minutes) semi-structured interview during which participants were asked about the rationale for their item configuration with special attention to the cards placed at the extreme ends of the distribution and in the neutral category. They were further asked to identify any outcomes that they found difficult to place or understand, or that they thought were missing from the item set. The interviews aimed to facilitate the interpretation of the viewpoints emerging from the by-person factor analysis by illustrating the participants' rationale, and by reducing the amount of interpretation added by the researcher to minimise bias and increase validity. Interpreting factors based on participant accounts rather than the literature also enables the generation of new theory as it opens up the possibility of unexpected findings and rationales (Gallagher & Porock, 2010).

Most Q-sorts were administered in person, with one conducted over the phone. Half of the young people recruited into the sample completed the sorting task individually, but in the company of peers within their peer support group, while the other half completed the sort at an individual appointment. Of the professionals recruited into the sample, 19 completed the task individually within their work hours, and at their workplace (e.g., during lunch time), while 11 participated in one of two workshop settings, where they produced their individual Q-sorts in the company of other attendants. Follow-up interviews with youth and professionals were conducted individually and in private. Young people and clinicians also filled in a brief structured questionnaire, which collected demographic data, as well as background information on young people's experience with depression, and clinicians' professional background (see Appendix D, Figure D.6 and Figure D.7). Young people were remunerated for their time with a £10 shopping voucher, and reimbursed for travel expenses.

6.2.4 Ethical Approval and Informed Consent

This study was reviewed by the ethics review committee of University College London and approved in March 2018 (UCL Ethics Project ID Number: 10567/002). All participants provided their informed consent. Since all participants were aged 16 years or older, parental consent was not required. The ethics approval letter, and examples of the participant information sheets and consent forms can be consulted in Appendix D, Figure D.1, Figure D.2, and Figure D.3.

6.2.5 Statistical Analysis

Statistical analysis and subsequent interpretation of the factor solutions were first conducted separately for each participant sample to obtain an in-depth understanding of the viewpoints existent within each group (in line with the process outlined in Figure 6.1, above). Second-order factor analysis was then conducted to compare viewpoints between young people and clinicians (Kline, 1994).

Analysis of Viewpoints Within Each Participant Group

The completed Q-sorts of all participants within each of the two samples were subjected to by-person factor analysis using the computer software PQMethod (Schmolck, 2014). Based on the correlation matrix of all the Q-sorts within each sample, patterns of convergence and divergence were identified. Using principal component analysis, the variety of Q-sort configurations was reduced to a smaller number of factors or typical ways of sorting the items, which were extracted based on the following criteria: (a) analysis of the scree plot of Eigenvalues, (b) consideration of the shared variance explained by the factor solution, (c) the number of Q-sorts loading significantly on only one factor, and (d) the correlation between factors scores. To improve the fit of the model, the unrotated correlation matrices were first subjected to Varimax rotation (Thurstone, 1947) to identify a factor solution that would maximise the amount of variance explained (Watts & Stenner, 2012). Small adjustments via hand rotation were then applied to increase the inclusivity of the factor solution by raising the number of participant Q-sorts loading significantly on a single factor (Watts & Stenner, 2012).

For each of the extracted factors, an ideal-typical Q-sort was generated by averaging the item rankings across all participants associated with this factor, using the correlation coefficients as weights (van Exel et al., 2006). The ideal-typical Q-sorts (or “factor arrays”) illustrate how an archetypical respondent loading on a factor would have sorted the outcome items. These were then used to describe the viewpoint conveyed by each factor, which involved a holistic examination of the ideal-typical item configuration, with special attention to the highest and lowest-ranked outcomes, and the outcomes that the relevant viewpoint ranked higher or lower than any other viewpoint (Watts & Stenner, 2012, p. 167). The viewpoint interpretations were then further strengthened by drawing on the qualitative data from the semi-structured interviews conducted with participants after the Q-sorts to better understand their sorting rationale (Gallagher & Porock, 2010).

Analysis of Super Factors Emerging Across Both Groups

In order to compare the viewpoints identified amongst young people and mental health professionals, a third Q-study was conducted that used the ideal-typical Q-sorts representing each viewpoint from the group-specific analyses as raw data. Such second-order factor analysis (Kline, 1994) yields a set of super factors that reflect the relationships between the group-specific factors (Watts & Stenner, 2012, pp. 53–54). Another principal component analysis with Varimax rotation was performed to examine the extent to which young people's and clinician's viewpoints represented common overarching viewpoints, or viewpoints distinctive to each group.

6.2.6 A Note on Reflexivity

Q-methodology is seen to minimise the researcher's influence on the research process, compared to other qualitative methodologies (S. R. Brown, 2006). Viewpoints are formed not primarily by interpretation, but through by-person factor analysis, based on the item configurations produced directly by participants. Nevertheless, the researcher's perspective inevitably influences the construction of the concourse and the creation of the Q-set at the design stage, as well as the interpretation of the viewpoints once extracted.

A note on the doctoral candidate's training as a sociologist has already been provided in Chapter 5. From this background stems a particular interest in the discourse[4], subjectivity, and power relations that determine what types of knowledge are considered valuable (Hook, 2001; Potter & Wetherell, 1987). The approach underpinning this study was rooted in constructivism, as outcome concepts and priorities were understood to be socially constructed. In this inquiry they were treated as valid, regardless of whether they were dominating current discourse or representing a minority viewpoint. At the same time, the doctoral candidate's placement within a unit of evidence-based practice also promoted a pragmatic approach to research problems with attention to their clinical and policy implications. This equally informed the choice of Q-methodology, which uses statistical factor analysis, more rooted in a positivist paradigm, to extract a clear mapping of viewpoints using a transparent process that initially reduces the interpretive role of the researcher relative to other qualitative and constructivist methodologies. As such, the choice of Q-methodology reflects both the pragmatist and transformational research paradigms underpinning this thesis.

The doctoral candidate had limited influence on the item configurations produced by participants, as they completed the Q-sort independently. The candidate's expectations and preconceptions did, however, come to the fore at the analytic stage, when the best factor solution was to be identified. Different factor solutions were examined in relation to statistical criteria, as well as their theoretical significance (Watts & Stenner, 2005, pp. 105–110). The

---

[4] Discourse can be defined as "a specific ensemble of ideas, concepts, and categorizations that are produced, reproduced and transformed in a particular set of practices and through which meaning is given to physical and social realities" (Hajer, 1997, p. 44).

candidate drew on her knowledge of the different viewpoints she had encountered in the literature, in previous stakeholder consultations, and amongst participants while administering the Q-sorts, to identify solutions that were theoretically meaningful. In addition, analysis aimed to identify a manageable number of viewpoints with clearly interpretable differences, which could be articulated to clinicians with relative ease to promote the clinical relevance and applicability of the study findings.

Two additional points are worth noting. First, the doctoral candidate approached the design of the Q-set anticipating that parents would form one of the participant groups. Special attention was given to including outcomes frequently discussed by parents in the Post-Therapy Outcome Perspectives study, leading to the inclusion of four outcomes around parental support and wellbeing. As mentioned above, however, it was impossible to recruit a sufficient parent sample within the time allocated for recruitment. When administering the Q-sort to young people and clinicians, the four parent-related outcomes were often (but not always) ranked as least important, and it is a limitation of this study that the parental perspective could not be added.

Second, the item *being able to do the same things other adolescents do* (item 11) was devised to represent global functioning, based on the narratives provided by young people, parents, and professionals in the Post-Therapy Outcome Perspectives study. However, during the Q-sort, participants across all viewpoints frequently interpreted this as meaning that youth should conform with social expectations – an idea that most rejected on the basis that therapy should enable young people to accept their individual preferences and needs and empower them to find their own path in life. This was unexpected and highlights the importance of fine-grained item-level piloting, which in this case was not fully possible due to time constraints.

6.3 Results

6.3.1 Results – Young People

Participant Characteristics

A total volunteer sample of 28 young people completed the Q-sort, who ranged in age from 16 to 21 years (mean age of 18.7 years, see Table 6.1). The sample included 18 young women, 9 young men, and one young person identifying their gender as non-binary. The sample was diverse with regards to their experiences of service use. Twelve participants (42.9%) were receiving treatment at the time of the research, while the remainder had previously completed treatment. Fifteen participants (53.6%) had accessed services on multiple occasions for recurring or non-remitting depression, while the remainder had thus far only started or completed one course of treatment. Seven participants (25%) reported having been admitted to emergency care in relation with their depression, and three (10.7%) had spent time in inpatient care.

Table 6.1. Demographic Characteristics – Youth Sample

Variable	*N* (%)
Female	18 (64%)
Currently receiving treatment	12 (43%)
Have started more than one cycle of treatment	15 (54%)
Ever admitted to A&E in relation to depression	7 (25%)
Ever hospitalised overnight in relation to depression	3 (11%)
Comorbid presenting problems *(based on self-report)*	
Anxiety	23 (82%)
Sleep	18 (64%)
Self-harm	17 (61%)
Eating	16 (57%)
Neurodevelopmental disorder (e.g., ADHD, ASD, Dyslexia)	10 (36%)
Anger and violence	6 (21%)
Obsessions or compulsions	5 (18%)
Substance use	5 (18%)
Psychosis	5 (18%)
Trauma	4 (14%)
Other	2 (7%)
Types of treatment received	
Individual therapy or counselling	24 (86%)
Medication	16 (57%)
Family therapy	13 (46%)
Group therapy	8 (29%)
Other treatment format	3 (11%)
	M (*SD*)
Mean age (in years)	18.7 (1.83)
Age range	16-21

Note. *N* = 28.

While data on formal diagnoses was not collected, young people were asked to self-identify any comorbid presenting problems that had been a focus of their treatment. Young people mentioned four additional problems on average, ranging from none to as many as eight. The most commonly mentioned comorbid difficulties were anxiety (82.2%), disrupted sleep (64.3%), self-harm (60.7%) and disordered eating (57.1%). Six young people (21.4%) reported having a learning difficulty, and four each stated struggling with Attention Deficit Hyperactivity Disorder (ADHD, 14.3%) or ASD (14.3%). Participants resided in a mix of urban and rural areas in Southern and Northern England, and came from a variety of socio-economic and ethnic backgrounds.

Factor Solution

Through principal component analysis, four distinct viewpoints were identified from the 28 Q-sorts completed by young people (see Table 6.2 and Table 6.3, below). The overall variance explained by these four factors was 48.7%, which can be considered satisfactory (Kline, 1994). Following Varimax rotation, 20 Q-sorts had significant factor loadings on one of the four extracted factors (i.e., factor loadings exceeded ±0.44, with the significance threshold set at the 0.01 level); seven Q-sorts did not have significant loadings on any factor; and one had significant loadings on more than one factor (i.e., was confounded). Small adjustments per hand rotation (see Appendix D, Table D.2 for details) yielded a more inclusive factor solution, with 25 participants loading significantly one of the four factors. Two Q-sorts did not load significantly on any factor, and one remained confounded. These three Q-sorts were not considered for further analysis (Watts & Stenner, 2012). There was no significant inter-correlation between the factors which suggest that each factor represents a distinct viewpoint (see Appendix D, Table D.3). Characteristics of the rotated factor solution are displayed in Table 6.2 (below).

Table 6.2. Factor Characteristics – Youth Sample

Factor Characteristics	Factors			
	1	2	3	4
Explained variance (%)	13.0	14.6	13.5	7.6
Cumulative explained variance (%)	13.0	27.6	41.4	48.7
No of defining sorts (*n*)	6	8	8	3
Average reliability coefficient	0.80	0.80	0.80	0.80
Composite reliability	0.96	0.97	0.97	0.92
Standard Error of Factor Scores	0.20	0.17	0.17	0.28

The rotated factor matrix is shown in Table 6.3. The ideal-typical outcome ranking for each viewpoint (or factor array) is shown in Figure 6.3. These ideal-typical Q-sorts informed the interpretation of the viewpoints, which also drew on the explanations provided by young people in the post-sort interviews. Each viewpoint description refers back to the relevant ideal-typical Q-sort by indicating item numbers in parentheses, alongside the rank assigned to the item within the relevant ideal-typical Q-sort. Young people’s names have been replaced by pseudonyms to ensure confidentiality.

Table 6.3. Factor Loadings Following Rotation – Youth Sample

Viewpoint name	Pseudonym	Factor loadings			
		Y.A	Y.B	Y.C	Y.4
Y.A: Becoming a ‘healthier’ person	Becca (17 yrs)	**.80**	.35	.14	.13
	Marco (19 yrs)	**.79**	.18	-.09	.23
	Ellie (17 yrs)	**.69**	.09	.18	-.19
	Soraya (21 yrs)	**.62**	.16	-.04	-.11
	Samuel (17 yrs)	**.61**	.30	.37	.25
	Josh (18 yrs)	**.49**	.44	.33	.07
Y.B: Learning to help myself	Melody (20 yrs)	-.28	**.77**	.14	-.04
	Adam (17 yrs)	.09	**.68**	-.39	.24
	Ameera (16 yrs)	-.05	**.63**	.30	.13
	Jacob (21 yrs)	.37	**.54**	.10	-.43
	Hannah (17 yrs)	.26	**.52**	-.24	-.02
	Liam (21 yrs)	-.07	**.51**	.16	.00
	Taylor (21 yrs)	.08	**.47**	.28	.36
	Boris (21 yrs)	.26	**.44**	.32	.16
Y.C: Making sense of the past to embrace the future	Lauren (20 yrs)	-.01	.06	**.74**	.05
	Chelsea (17 yrs)	.04	-.12	**.67**	.35
	Imogen (21 yrs)	.03	-.02	**.61**	-.20
	Liz (18 yrs)	-.23	.15	**.53**	-.04
	Connor (21 yrs)	-.03	.44	**.51**	.02
	Jade (19 yrs)	.39	.28	**.50**	-.22
	Chloe (16 yrs)	.26	.33	**.49**	-.18
	Amber (18 yrs)	.18	.29	**.44**	.32
Y.D: Being able to do what other adolescents do	Lewis (17 yrs)	.04	-.05	-.05	**.66**
	Georgia (17 yrs)	-.27	.19	.00	**.65**
	Meghan (21 yrs)	.25	-.07	016	**.51**
Not assigned	Karimah (19 yrs) [a]	-.09	.55	.54	-.09
	Faizah (18 yrs) [b]	.29	-.05	.06	.04
	Lien (21 yrs) [b]	.03	.25	.11	-.25

Note. Y.A. = youth viewpoint A; Y.B. = youth viewpoint B, etc. The extraction method was principal component analysis with varimax rotation and adjustment via hand rotation. Factor loadings printed in bold were significant at the $p < 0.01$ level (i.e., loadings equal to or above 0.44). The corresponding q-sort contributed to computing the ideal-typical factor array for this factor.

[a] This Q-sort had significant loadings on two factors and was therefore excluded from the computation of the factor arrays. [b] These q-sorts did not load significantly on any of the four factors and were excluded from the computation of the factor arrays.

Figure 6.3. Ideal-Typical Q-sorts – Youth Viewpoints

#	Q-sort item	Item rank			
		Y.A	Y.B	Y.C	Y.D
Symptoms					
1	Being less angry and not losing my temper as much.	-2	-4	-1	0
2	Feeling less down and depressed.	4	2	2	3
3	Feeling happier and enjoying things more.	4	4	1	4
4	Feeling more loved.	3	1	-2	-2
5	Engaging less in behaviour that can be harmful.	3	-4	4	0
Self-management					
6	Being more active and engaged in things.	0	-2	0	-2
7	Knowing ways to cope with my emotions.	3	3	1	-4
8	Having a better understanding of my feelings and thoughts.	0	3	0	-1
9	Being able to challenge negative thoughts and approach situations differently.	2	4	1	-3
Functioning					
10	Being better able to get things done (e.g. concentrate, be organized).	1	0	-2	0
11	Being able to do the same things other adolescents do.	-2	-2	-3	4
12	Working more effectively in school (e.g. being more motivated and focused).	0	-1	-1	0
13	Attending school more regularly.	-2	0	-3	-3
14	Being more sociable and better able to be around other people.	2	-1	-3	1
Personal growth					
15	Feeling more confident.	1	0	0	3
16	Being better able to stand up for my needs and opinions.	-1	-3	0	-2
17	Being more independent and able to take responsibility for my life.	-1	2	0	0
18	Being able to make sense of things that have happened in the past, or that are still happening.	1	2	3	1
19	Having a better sense of who I am and how to be myself around others.	0	1	-1	1
Relationships					
20	Feeling more able to talk about my feelings and thoughts.	-1	3	-1	2
21	Getting on better with my family	0	0	2	-4
22	Getting on better with my friends or having made new friends.	1	-3	-4	-1
23	Getting on better with my peers in school (e.g. not feeling bullied).	-2	-3	-4	-3
Therapeutic space					
24	Having a space where someone listens and cares about me.	-1	-2	3	-2
25	Having a space where I can let out my feelings.	-1	0	-1	2
26	Having a space where I can talk about anything without being judged.	0	1	1	-1
27	Having a space to reflect and think about things differently.	-3	0	0	-1
Wellbeing					
28	Having greater peace of mind (e.g. feeling calmer, more balanced).	2	1	2	-1
29	Feeling more optimistic and positive about life and the future.	2	2	4	2
30	Feeling physically healthier.	0	-2	-2	0
31	Being able to make plans for the future and have goals.	1	0	2	2
Parental support and wellbeing					
32	My parents feeling happier and less stressed and worried.	-3	-1	0	1
33	My parents having a better understanding of me and my difficulties.	-3	1	3	0
34	My parents feeling more able to support me.	-4	-1	1	1
35	My parents feeling less guilty.	-4	-1	-2	3

● Y.A: Becoming a 'healthier' person

● Y.B: Learning to help myself

● Y.C: Making sense of the past to embrace the future

● Y.D: Being able to do what other adolescents do

Note. Gradient colour coding has been applied to highlight items ranked in positions +4, +3 and +2. Y.A. = youth viewpoint A; Y.B. = youth viewpoint B, etc.

Youth Viewpoint A: Becoming a 'Healthier' Person

Youth Viewpoint A focussed on overcoming typical symptoms of depression, such as low mood, anhedonia, and feeling unloved; as well as improved wellbeing. Six young people represented this viewpoint, three of whom were female. Half had completed several cycles of treatment, as opposed to one cycle. One participant had accessed emergency care, but none had spent time in inpatient care in relation with their depression. On average, participants in this group reported comorbid difficulties in four areas, which included anxiety in all six cases, and problems with sleeping, eating and self-harm in four cases.

A primary desire for these young people was for key depression symptoms to resolve, as explained by one young person in the following words:

> I'd been feeling it for so long it was like something that I just wanted to get rid of. And especially because like "feeling less down and depressed", like for a depressed person that seems like… heaven (laughs) … you know? (Becca, 17 years).

They highly endorsed outcomes such as *feeling less down and depressed* (item 2, rank +4), *feeling happier and enjoying things more* (item 3, rank +4), *feeling more loved* (item 4, rank +3), and *engaging less in risky and harmful behaviour* (item 5, rank +3). Alongside these, young people endorsed outcomes that expressed a broader sense of wellbeing, but were related to depression symptoms, such as greater *peace of mind* (item 28, rank +2) and *optimism about life and the future* (item 29, rank +2). As such, recovery for these young people was a quest for improved wellbeing, which one young person described as a journey towards "*becoming a more healthy person"* (Samuel, 17 years).

Knowing ways to cope with emotions (item 7, rank 3), and being able to *challenge negative thoughts and approach situations differently* (item 9, rank +2) were other important outcomes. Several young people reported engaging in self-harm or other risky forms of coping such as substance use or gambling. While providing relief in the short-term, these strategies were seen to entail greater difficulties in the long term. Other outcomes related to self-management and personal growth were considered less important.

There was a tendency amongst young people in this viewpoint to perceive their depression as a barrier to connecting with others, leaving them feel isolated. Consequently, *being more sociable* (14, +2), and *getting on better with friends* (item 22, rank +1) were ranked more highly by this group than by other viewpoints. Youth in this viewpoint were generally reluctant to rely on family support to cope with their depression. Some preferred dealing with their difficulties at their own pace, while others doubted that their parents would be able to understand and support them. Consequently, outcomes related to parental support and parental wellbeing were assigned the lowest importance, compared to all other outcomes (32, -3; 33, -3; 34, -4; 35, -4).

> My parents didn't know until after a good year or so until I started medication. […] I kept them out literally as long as possible, and so to me it didn't really matter whether they understood. (Marco 19 years)

In summary, these young people described a good outcome as liberating themselves from the negative feelings associated with depression, which they perceived as a solitary quest that could be mastered with the help of therapy, but little reliance on their environment.

Youth Viewpoint B: Learning to Help Myself

The second youth viewpoint emphasised the importance of coping and self-management outcomes. Seven young people represented this viewpoint, three of whom were female. These young people reported the lowest burden from depression and other mental health problems, compared with other viewpoints, and had no experience of visiting emergency or inpatient care in relation with their depression. They described comorbid difficulties in three areas, on average, with five out of eight experiencing anxiety, half reporting eating or sleeping problems, three reporting self-harm, and one each reporting other difficulties.

Young people representing this viewpoint considered that the most important outcomes were related to *feeling happier and enjoying things more* (item 3, rank +4), and *being able to challenge negative thoughts and approach situations differently* (item 9, rank +4), followed by *knowing ways to cope with emotions* (item 7, rank +3), and *having a better understanding of their feelings and thoughts* (item 8, rank +3). Contrary to Youth Viewpoint A (*Becoming a 'healthier' person*), they did not believe in a cure for their depression symptoms, and anticipated that a vulnerability to depression would remain a feature of their lives. *Making sense of their past and current experiences* (item 18, rank +2) and *becoming more independent and able to take responsibility for their lives* (item 17, rank +2) were important outcomes from the personal growth domain, that underscored a desire to take back control, and actively manage depression symptoms. By gaining a better understanding of their emotions and possible triggers of low mood and self-doubt, and by learning practical coping strategies, young people hoped to be able to break the momentum of depressive cycles:

> Like when the first wave of sadness hits, normally you don't have the strategies, you will just like, again, be snowballing, but […] it's important to find ways to kind of break that momentum and stop that snowball before it just gets worse. (Jacob, 21 years)

Young people representing this viewpoint appreciated having a therapeutic space where they could talk about feelings and thoughts without being judged (item 26, rank +1), but were reluctant to become dependent on therapy.

> When I first went to CAHMS, it was a case of I just wanted to not feel this way anymore. But when I kept going back to CAHMS, then I thought, this isn't sustainable, I need to be able to function without CAHMS, so the more times I cycled through getting help, the more important sort of resilience or being able to help myself became. (Hannah, 17 years)

They described a notion of wanting to transpose the therapeutic space into their daily lives, by internalising it within themselves, and by relying on their social networks for support. Indeed, *being more able to talk about feelings and thoughts* was another highly ranked outcome (item 20, rank +3). Outcomes related to family and parental support were not highly endorsed, as most young people felt sufficiently supported already.

The outcomes ranked lowest by this viewpoint related to *anger management* (item 1, rank -4), and *self-harm and risky behaviour* (item 5, rank -4), which these young people did not generally struggled with. Similarly, they did not feel that their depression impacted negatively on their relationship with friends and peers (items 22 and 23, rank -3), or on their confidence and assertiveness (item 15, rank 0, item 16, rank -3). In summary, young people representing this viewpoint appeared comparatively well resourced to confront life's challenges, and considered therapy as a temporary source of support that could help develop their self-management skills.

Youth Viewpoint C: Making Sense of the Past to Embrace the Future

The third youth viewpoint focussed on finding safe outlets for emotions, making sense of past experiences, and gaining a more positive outlook into the future. Eight young people were associated with this viewpoint, seven of whom were female. Half had completed several courses of treatment, three had visited emergency care in relation with their depression, and one had spent time in inpatient care. On average, they described comorbid problems in five areas, with anxiety being most frequent (7 out of 8), followed by self-harm (6 out of 8) and sleeping problems (5 out of 8). Four participants reported learning difficulties, three an ASD, and two ADHD.

A considerable share of young people representing this viewpoint had sought mental health support not just for depression, but also in relation with learning difficulties, ADHD, ASD, or trauma. They felt that growing up with these difficulties had set their experiences apart from those of peers or family members, and they themselves often struggled to make sense of their own experiences. For some, living with ASD created anxieties about the future, such as their ability to access higher education or employment, and exacerbated problems with their mood. In this context, young people endorsed a mix of outcomes that revolved around calming some of the anxieties and confusion that stemmed from experiencing the world differently. Amongst the most highly ranked outcomes were *feeling more optimistic about life and the future* (item 29, rank +4), and *having greater peace of mind* (item 28, rank +3). Young people representing this viewpoint felt that unresolved issues from the past were holding them back from embracing their future, and they felt that they needed to make sense of past and current experiences (item 18, rank +3), in order to be able to move on with their lives.

> I kind of wanna get all my thoughts in order and there's a lot of stuff that has happened in the past that I wanna deal with before I start dealing with stuff now. (Chelsea, 17 years)

> I do have PTSD and that still comes back to me and I can't figure that out, I can't figure out why that has happened, or why something has happened on so many occasions, why it always happens to me, it's just, there's no explanation for it and I find that occupies my mind a lot. (Imogen, 21 years)

> With the Asperger's I don't really understand emotions in general […] I can never tell if I'm sort of truly feeling something or if I'm just thinking I'm feeling that. (Jade, 19 years)

Another crucial outcome for this group was *engaging less in harmful behaviour* (item 5, rank +4). Young people described self-harm as an "outlet" for their emotions, but acknowledged that it created a high amount of stress for themselves and their families. They explained their need for an outlet by feeling frequently overwhelmed by their emotions and worries, and unable to articulate them to their families. This was reflected in young people endorsing outcomes such as *having a space where somebody listened and cared about them* (item 24, +3), *parents gaining a better understanding of their difficulties* (item 33, rank +3), *parents learning to support them more effectively* (item 34, rank +1). They also hoped to improve their relationships with their families (item 21, rank +2), which were often strained by both sides struggling to understand their difficulties and experiences:

> My dad… cause he's so fixated on me being this perfect child that he doesn't quite grasp that the things that he can do, I may not be able to do. For example, because I've got Autism, Asperger's, high-functioning, and I've got other learning disabilities and things that make it difficult for me which he doesn't quite understand. (Connor, 21 years)

Outcomes around peer relationships (items 22 and 23, rank -4), and daily functioning (items 11, 13 and 14, rank -3) were assigned the lowest importance by this group. They did not consider that engaging in the same activities as other adolescents or being more sociable were relevant goals of treatment, partly because they accepted that they did not experience the world quite like 'typical' adolescents, and felt at ease with this.

Youth Viewpoint D: Being Able to Do what Other Adolescents Do

The fourth viewpoint represented an experience marked by a constant struggle with depression and other mental health difficulties, and a desire to recover a sense of normality. Three participants loaded significantly on this factor, two of whom were female. Two each had been in treatment repeatedly, accessed emergency care, or inpatient care. On average, young people in this group reported comorbid difficulties in seven additional areas, which represented the highest burden experienced by any viewpoint. All three reported anxiety, self-harm, eating and sleeping problems; two each reported issues with anger, trauma, and psychosis; and one had a learning difficulty.

Young people in this small group experienced a complex set of difficulties, which interfered considerably with their daily lives, and had done so for years. They described having to interrupt school or reduce their subjects, being unable to go out with friends, or struggling to move about town using public transport. To this group, *feeling happier and enjoying things more* (item 3, rank +4) and *being able to do the same things other adolescents do* were the two most important outcomes of treatment (item 11, rank +4). *Feeling less down and depressed* was also highly ranked (item 2, rank +2).

> It’s affected everything, like literally everything. And so it means that I’m not like the average person my age, like they are all [inaudible] and I sit at home and think, ‘well actually, if I went out and did that, like what could the consequences be?’ (Georgia, 17 years).

With complex mental health difficulties being a constant presence and defining feature of their daily lives, young people in this group struggled to envisage an identity and future beyond these difficulties, and described their sense of self and outlook into life as fluctuating with their mood: “Who I am can feel quite dependent on my mood at that moment, and if I’m feeling very low then I’m like […] nothing’s ever gonna be worth it…” (Meghan, 21 years). They endorsed outcomes that related to recovering confidence, hope and optimism about the future (item 15, +3; item 29, rank +2; item 31, rank +2). Beyond their own lives, they felt that their mental health difficulties had also affected the lives of those around them, most notably their parents. Contrary to all other viewpoints, these young people considered that their parents feeling less guilty (35, +3), and less stressed and worried (item 32, rank +1) were important outcomes in their own right.

Contrary to other youth viewpoints, this group did not consider improved coping and self-management to be important outcomes (item 6, rank -2; item 7, rank -4; item 8, rank -1; item 9, rank -3). They expressed scepticism that such strategies could be deployed at will, and described their feelings as often too overwhelming to apply such strategies:

> I guess if my mood was better then-, like I don’t think everyone in their life consciously kind of has a negative thought and then it’s like ‘but I must challenge this’, it’s more that these negative thoughts don’t get in the way so much. So, I’d hope that if my mood was better…and I guess, just “approach situations differently” just seems… I don’t know, that whole thing seems very much like what a professional would want for my life rather than what I want? (Meghan, 21 years).

Common Outcome Themes Across Youth Viewpoints

Some outcomes were frequently endorsed across all four youth viewpoints. *Feeling happier and enjoying things more* (item 3) was ranked in the most important category (rank +4) by three out of four viewpoints, and *feeling less down and depressed* (item 2) was considered important or very important (ranks +2 to +4) by all. This showcased a general agreement on the importance of reducing depressive symptoms, despite differences in prioritising outcomes in other domains. Another outcome considered important by all four viewpoints was *feeling more optimistic and positive about life and the future* (item 29, ranks +2 to +4).

Consistently ranked around the neutral middle of the importance scale (ranks -1 to +1) was *working more effectively in school* (item 12), *having a better sense of who I am and how to be myself around others* (item 19), and *having a space where I can talk about anything without being judged* (item 26). Most consistently ranked as an unimportant outcome of treatment was *getting on better with my peers in school* (item 23, ranks -2 to -4), with many young people explaining that they did not struggle with peer relationships or bullying, and thus considered that this item did not apply to them.

6.3.2 Results – Clinicians

Participant Characteristics

The Q-study with mental health professionals involved a final sample of 30 volunteers, aged 22 to 65 years (*mean* 43.3 years, *SD* = 11.5, see Table 6.4). The majority (*n* =19, 63%) were female. On average, participants had 16.7 years (*SD* = 10.1) of experience working in child mental health (ranging from 1 to 37 years). Thirteen identified as psychologists or psychotherapists, five as psychiatrists, and five as mental health nurses. The remaining five professionals included mental health support workers, service coordinators and commissioners. Professionals worked at statutory outpatient CAMH services within the National Health Service (NHS; *n* = 13), CAMH charities (*n* = 8), school or university settings (*n* = 4), statutory in-patient services (*n* = 3), or other settings (*n* = 2). Participants reported employing a range of treatment modalities, with Cognitive Behaviour Therapy (*n* = 15) and Systemic Therapy (*n* = 14) being the most commonly used.

Table 6.4. Demographic Characteristics – Professional Sample

Variable	*N* (%)
Female	19 (63%)
Professional role	
Psychologist / psychotherapist	14 (47%)
Mental health nurse	5 (17%)
Psychiatrist	5 (17%)
Another role [a]	5 (17%)
Setting	
Statutory CAMHS (outpatient)	13 (43%)
CAMH Charity	8 (27%)
School or university	4 (13%)
Statutory CAMHS (inpatient)	3 (10%)
Other setting [b]	2 (6%)
Treatment modalities	
Cognitive Behavioural Therapy	15 (50%)
Systemic Therapy	14 (47%)
Pharmacological treatment	7 (23%)
Behavioural activation	7 (23%)
Supportive counselling	6 (20%)
Psychoanalytic or psychodynamic psychotherapy	6 (20%)
Social skills therapy	4 (13%)
Problem solving therapy	4 (13%)
Dialectical Behaviour Therapy	3 (10%)
Mentalisation based Therapy	3 (10%)
Psychoeducation, guided self-help and signposting	3 (10%)
Other treatment modality	10 (30%)
	Mean (*SD*)
Mean age (in years)	43.3 (11.5)
Age range	22 – 65 years
Years of professional experience in CAMH	16.7 (10.1)

[a] Other roles included support worker, commissioner, clinical social worker, and service coordinator. [b] Other settings included an adolescent outreach team and a clinical commissioning group. [c] Other treatment modalities included case work, group work, inpatient care, mindfulness, crisis support, school-based counselling, trauma-focussed CBT.

By-person factor analysis of the 30 Q-sorts identified a three-factor solution, which explained a total common variance of 48.0%. Following Varimax rotation, 27 Q-sorts loaded significantly on one of the three factors (i.e., factor loadings exceeded ±0.44, with $p < .01$). Two Q-sorts did not have significant loadings on any factor, and one had significant loadings on more than one factor (i.e., was confounded). Small adjustments by hand rotation (see Appendix D, Table D.4) resolved the confounding of the third Q-sort, so that the final factor solution had 28 out of 30 participants loading significantly on one of the three factors. Inter-factor correlations varied between 0.26 and 0.33, but none reached statistically significant levels (see Appendix D, Table D.5).

The first factor was comprised of significant positive and negative loadings. In Q-methodology, such factors convey two viewpoints that represent polar opposite of one another. Outcomes considered most important by participants loading positively on this factor were considered least important by those with negative loadings. For further analysis, this factor was split into two viewpoints, consisting of the positively and negatively loading Q-sorts, respectively. This was done by duplicating the existing factor and reversing the polarity of one of them (Schmolck, 2008). Ideal-typical Q-sorts were computed for each of these two viewpoints, and each was interpreted in its own right. The Q-study involving professionals thus revealed four distinctive viewpoints, which are discussed below. Characteristics of the rotated factor solution are displayed in Table 6.5. The rotated factor matrix is shown in Table 6.6. The ideal-typical outcome ranking for each viewpoint (or factor array) is shown in Figure 6.4.

Table 6.5. Factor Characteristics – Professional Sample

Factor Characteristics	Factors		
	1 [a]	2	3
Explained variance (%)	17.5	18.0	12.5
Cumulative explained variance (%)	17.5	35.5	48.0
No of defining sorts (*n*)	13	8	7
Average reliability coefficient	0.80	0.80	0.80
Composite reliability	0.98	0.97	0.96
Standard Error of Factor Scores	0.14	0.17	0.19

Note. The characteristics shown refer to the extracted factors prior to the splitting of factor 1.

[a] Factor 1 was subsequently split into two viewpoints, one representing positively loading Q-sorts and one representing negatively loading Q-sorts.

Table 6.6. Factor Loadings Following Rotation – Professionals

Professional viewpoint	ID	Role	Factor loadings			
			P.A	P.B	P.C	P.D
P.A: Managing risk and reducing symptoms	P1	Psychiatrist	**.72**	-.72	-.10	.31
	P2	Psychotherapist	**.69**	-.69	.29	.38
	P3	Psychiatrist	**.66**	-.66	.17	.32
	P4	Psychotherapist	**.62**	-.62	.37	-.05
	P5	Psychotherapist	**.61**	-.61	.03	.16
	P6	Psychiatrist	**.61**	-.61	.44	.30
	P7	Mental health nurse	**.55**	-.55	.27	.23
	P8	Psychotherapist	**.52**	-.52	.03	.19
	P9	Psychotherapist	**.48**	-.48	.36	-.01
	P10	Role missing	**.48**	-.48	.02	-.07
	P11	Psychiatrist	**.45**	-.45	.29	-.05
P.B: Empowering youth and parents	P12	Support Worker	-.72	**.72**	.25	.34
	P13	Psychiatrist	-.49	**.49**	-.37	.37
P.C: Building skills in young people to aid coping	P14	Psychotherapist	-.05	.05	**.88**	.05
	P15	Psychotherapist	.05	-.05	**.81**	-.04
	P16	Mental Health nurse	.09	-.09	**.74**	.08
	P17	Mental Health nurse	.07	-.07	**.72**	-.05
	P18	Support Worker	.00	-.00	**.64**	.39
	P19	Mental Health nurse	-.02	.02	**.60**	.06
	P20	Psychotherapist	.12	-.12	**.58**	.31
	P21	Mental Health nurse	.17	-.17	**.56**	.31
P.D: Building family support around young people.	P22	Psychotherapist	.43	-.43	.17	**.68**
	P23	Psychotherapist	-.42	.42	-.12	**.66**
	P24	Service coordinator	-.01	.01	.24	**.62**
	P25	Psychotherapist	.19	-.19	.38	**.62**
	P26	Psychotherapist	.27	-.27	.14	**.58**
	P27	Psychotherapist	.07	-.07	.23	**.51**
	P28	Clinical Social Worker	.20	-.20	-.28	**.45**
Not assigned [a]	P29	Psychotherapist	.35	-.35	.30	.25
	P30	Commissioner	-.05	.05	.23	-.27

Note. P.A. = professional viewpoint A; P.B. = professional viewpoint B, etc. The extraction method was principal component analysis with varimax rotation and adjustment via hand rotation. Factor loadings printed in bold were significant at the $p < .01$ level (i.e., loadings equal to or above 0.44). The corresponding q-sort contributed to computing the ideal-typical factor array for this factor.

[a] These q-sorts did not load significantly on any of the factors and were excluded from the computation of the factor arrays.

Figure 6.4. Ideal-Typical Q-sorts – Professional Viewpoints

#	Q-sort item	Item rank			
		P.A	P.B	P.C	P.D
Symptoms					
1	Being less angry and not losing their temper as much.	1	-3	-1	0
2	Feeling less down and depressed.	3	-4	2	0
3	Feeling happier and enjoying things more.	4	-3	1	-1
4	Feeling more loved.	0	-4	-3	1
5	Engaging less in behaviour that can be harmful.	4	-1	3	4
Self-management					
6	Being more active and engaged in things.	1	-2	-1	2
7	Knowing ways to cope with their emotions.	3	0	4	2
8	Having a better understanding of their feelings and thoughts.	0	1	4	3
9	Being able to challenge negative thoughts and approach situations differently.	0	2	3	0
Functioning					
10	Being better able to get things done (e.g., concentrate, be organized).	2	-2	-2	-1
11	Being able to do the same things other adolescents do.	0	0	-2	-2
12	Working more effectively in school (e.g., being more motivated and focused).	0	-3	-3	-1
13	Attending school more regularly.	-1	3	-1	1
14	Being more sociable and better able to be around other people.	1	-1	-2	0
Personal growth					
15	Feeling more confident.	2	0	0	0
16	Being better able to stand up for their needs and opinions.	-1	2	-2	1
17	Being more independent and able to take responsibility for their life.	-1	4	0	0
18	Being able to make sense of things that have happened in the past, or are still happening.	-1	0	2	-1
19	Having a better sense of who they are and how to be themselves.	-2	3	1	2
Relationships					
20	Feeling more able to talk about their feelings and thoughts.	0	2	2	0
21	Getting on better with their family	1	1	0	3
22	Getting on better with their friends or having made new friends.	2	0	-1	1
23	Getting on better with their peers in school (e.g., not feeling bullied).	1	0	-1	1
Therapeutic space					
24	Having a space where someone listens and cares about them.	-2	-1	0	-4
25	Having a space where they can let out their feelings.	-4	1	0	-2
26	Having a space where they can talk about anything without being judged.	-2	1	1	-3
27	Having a space to reflect and think about things differently.	-3	1	3	-4
Wellbeing					
28	Having greater peace of mind (e.g., feeling calmer, more balanced).	0	0	0	-2
29	Feeling more optimistic and positive about life and the future.	3	-1	2	-1
30	Feeling physically healthier.	-3	-2	-4	-2
31	Being able to make plans for the future and have goals.	2	-1	1	2
Parental support and wellbeing					
32	Parents feeling happier and less stressed and worried.	-3	-2	-3	-3
33	Parents having a better understanding of their child and their difficulties.	-2	2	1	3
34	Parents feeling more able to support their child.	-1	3	0	4
35	Parents feeling less guilty.	-4	4	-4	-3

● P.A: Managing risk and reducing symptoms

● P.B: Empowering youth and parents

● P.C: Building skills in young people to aid coping

● P.D: Building family support around young people.

Note: P.A. = professional viewpoint A; P.B. = professional viewpoint B, etc. Gradient colour coding has been applied to highlight items ranked in positions +4, +3 and +2.

Professional Viewpoint A: Managing Risk and Reducing Symptoms

This first viewpoint conveyed a focus on outcomes related to symptom reduction and risk management. Eleven Q-sorts loaded positively on the first extracted factor, which explained 17.5% of the study variance. Of the professionals associated with this viewpoint, five identified as psychologists, four as psychiatrists and one as a mental health nurse; 64% were female. On average, professionals had 20.5 years of experience working in CAMH. Commonly employed treatment types included CBT, systemic therapy, and behavioural activation.

One of the two outcomes considered most important by this viewpoint was young people *engaging less in behaviour that could be harmful* (item 5, rank +4). Professionals considered keeping young people safe their principal responsibility, stressing that preventing suicide had to be the most important goal of depression treatment. The second most important outcome was *feeling happier and enjoying things more* (item 3, rank +4), closely followed by young people *feeling less down and depressed* (item 2; +3). While the two items are related, professionals considered that feeling happier was more than an absence of depression, as it extended to not feeling numb, not feeling anxious, and being able to enjoy life.

> The ability to taste and enjoy pleasure – that seems, to me, fantastic and important and […] with the sort of future facing optimism, there’s hope, there’s agency. (Professional #3)

In addition to these symptom-focussed outcomes, this viewpoint also prioritised *knowing ways to cope with emotions* (item 7, rank +3), young people *feeling more optimistic about life and the future* (item 29, rank +3), and *being able to make plans for the future and have goals* (item 31, rank +2), thus endorsing outcomes related to coping and wellbeing in a similar vein as Youth Viewpoint A (*Becoming a healthier person*). Restoring hope and a sense of the future was considered key to promoting young people’s safety:

> So, to me […] feeling hopeful about the future is a ‘keeping them alive’ outcome, which I think is great. (Professional #2)

Other outcomes considered important included young people *being able to get things done* (item 10, rank +2), *getting on better with friends* (item 22, rank +2). In contrast, a number of outcomes were assigned low importance because professionals considered them to be procedural or secondary to the ultimate goals of symptom and risk reduction. They included the benefit of the therapeutic space (items 24 and 27, rank -2; item 25, rank -3; item 26, rank -4), gaining a stronger sense of self, (item 19, -2), understanding feelings and thoughts (item 8, 0), and making sense of past or present experiences (item 18, -1).

> Lots of young people I work with say that they value that a lot in therapy – having that safe space – but I don’t think it’s an outcome. It’s not what they’re coming to therapy for specifically. They’re coming to therapy because they want to feel better. (Professional #5)

Another set of outcomes that were assigned low importance related to parental wellbeing and support (item 31, rank -3; item 33, rank -2; item 34, rank -1; item 35, rank -4), as professionals saw their role in focussing on the young person's wellbeing first and foremost.

Professional Viewpoint B: Empowering Youth and Parents

The second professional viewpoint represented the mirror image of the first, based on splitting the first factor. Only two participants represented this viewpoint, including one mental health support worker and one psychiatrist based at an inpatient unit. This viewpoint was characterised by a double focus on empowering young people to become more independent and self-reliant, and empowering their parents to feel more confident in supporting their child.

These two professionals did not consider a reduction in symptoms to be a primary outcome of treatment (item 2, -4; item 3, -3; item 5, -1). Instead, they described symptom reduction as only an initial step in a longer journey towards recovery, which should eventually lead to greater independence and ability to take responsibility for one's life (item 17, +4), and considered that more profound stabilisation was required to promote functioning in important areas such as school attendance (item 13, rank +3). They suggested that a stronger sense of identity (item 19, rank +3), assertiveness (item 16, +2), confidence (items 15, rank +2) and ability to talk about their feelings and thoughts (item 20, rank +2) were more meaningful indicators of change. There was thus a strong focus on promoting personal growth, and empowering young people to become confident agents of their own recovery.

> Rather than just thinking about the week to week changes in their depression stages […] if I am discharging a young person who is more independent and more able to take responsibility for their life, I think I'm setting them up for a good outcome moving on from hospital. (Professional #13)

The second most important outcome in the eyes of these professionals was *parents feeling less guilty* (item 35, rank +4), which was in stark contrast to the low rankings that this outcome received from other viewpoints. Other highly ranked outcomes included an improved parental understanding of the young person's difficulties (33, +3), and improved ability to support them (34, +3). This viewpoint considered that empowering parents was a crucial counterpart to empowering young people, as parents would struggle to support their children effectively, unless they felt confident about their role as parents.

Professional Viewpoint C: Building Skills in Young People to Aid Coping

This third viewpoint emphasised the importance of teaching young people coping and self-management skills to promoting their resilience. Eight professionals represented this viewpoint, including two psychologists working in school settings, one psychologist working with a CAMHS charity, four mental health nurses and one support worker. Professionals employed a range of approaches, including systemic therapy, supportive counselling, psychological crisis support, social skills therapy, pharmacological treatment, and CBT. On average, participants had 11.7 years of experience working in CAMHS. Half of the participants representing this viewpoint were female.

These professionals emphasised that challenges and difficult emotions formed an inherent part of life, and that some young people might struggle to become entirely symptom-free. On this basis, these professionals considered that outcomes related to self-management were the most important. They prioritised outcomes related to *learning ways to cope* with residual symptoms (item 4, +4), *gaining a better understanding of feelings and thoughts* (item 8, +4), and *learning to challenge negative thoughts and approach situations differently* (item 27, +3). They interpreted their role as enabling young people to help themselves: “I think it's around creating skills in young people, rather than treating them” (Professional #19).

This viewpoint emphasised that learning to challenge negative thinking styles associated with depression was crucial in enabling young people to feel happier (item 3, rank +1), more confident (item 15, rank 0), and more able to make friends (item 22, rank -1). Another outcome considered important was for young people to *be able to make sense of their past and present experiences* (item 18, +2), and how these may have contributed to their difficulties. Some professionals expressed that focussing solely on symptoms as part of outcome measurement risked conveying an incomplete picture of actual progress. These professionals thought of therapy as a collaborative space and assigned high importance to the idea of young people using this space for reflection and for adopting new perspectives (item 27, rank +3). However, similarly to youth viewpoint II (*Learning to help myself*), they stressed that young people would need to internalise this reflective space in order to become more resilient in the longer term, which underscored a focus on enabling young people to help themselves.

> If they’re able to have a space to reflect, to have a better understanding of their feelings and thoughts, then gradually, even without the treatment, they would hopefully continue to improve and grow and develop as people. (Professional #14)

Outcomes related to the domains of functioning (item 10, -2; item 11, -2; item 12, -3; item 13, -1; item 14, -2) and relationships with friends or family (21, 0; 22, -1; 23, -1) were ranked lower than by other viewpoints, as participants did not consider these to be essential outcomes. While they did value improvements in these areas, improvements were expected to occur naturally as a result of young people coping more successfully.

Professional Viewpoint D: Building Family Support Around Young People

The fourth professional viewpoint endorsed the idea of promoting young peoples’ resilience by strengthening family support. Seven professionals represented this viewpoint, six of whom were female. On average, participants had 16.2 years’ experience working in CAMH. Five were psychologists working at CAMHS outpatient services (*n* = 2) and a CAMHS charity (*n* = 2), and two had other roles. Treatment modalities included CBT, psychoanalytic or psychodynamic approaches, behavioural activation, systemic therapy, and Mentalisation.

To this group, improved parental support (item 34, rank +4) and insight into young people's difficulties (item 33, rank +3), as well as improved family relationships (item 21, rank +3) were amongst the most important treatment outcomes. They converged with Professional Viewpoint C (*Building skills in young people to aid coping*) on the notion that ups and downs were a natural feature of life, especially during adolescence, and that therapy should aim to foster resilience. But rather than focussing on individual resilience resources, they privileged strengthening family support and acceptance within young peoples' environments.

> If the relationship with the parent, the family improves, that often has more impact than anything that's going on in the room. This is the critical thing for me, how the family can construct a young person's depression in a way that's non-critical and supportive and caring. (Professional #22).

> A lot of the time, in order to help them help themselves in the future, they need to have a decent relationship with people around them. (Professional #27)

Despite the strong emphasis on family support, change in parental well-being was not considered a key purpose of treatment with professionals sharply distinguishing between outcomes related to parenting practices on the one hand (item 33, rank +3; item 34, rank +4) and parental wellbeing (item 32, rank -3; item 35, rank -3) on the other hand. This distinguished this viewpoint from Professional Viewpoint B (*Empowering young people and parents*), which considered parental wellbeing and empowerment a precondition for strengthening family support. Participants ranked the importance of symptom-related outcomes in the neutral area of the scoring distribution (item 2, rank 0; item 3, rank -1; item 1, rank 0), considering, like Professional Viewpoint C (*Building skills in young people to aid coping*), that being entirely symptom-free or having greater peace of mind (item 28, rank -2) constituted unrealistic goals:

> I think adolescence or life isn't really sort of like that. We just don't have peace of mind and it's about coming to terms with that, I think – that's part of what therapy does. You learn to accept that life's a struggle and that in itself already can be very calming. (Professional #23)

However, as most other viewpoints they considered reducing risky and harmful behaviour to be imperative (item 5, +4): "*Self-harming or drugs and alcohol, that's a bit of an absolute that one needs to really stop*" (Professionals #26). In stark contrast with Professional Viewpoint B (*Empowering young people and parents*), this viewpoint did not consider the therapeutic process to form an important outcome in itself (item 24, rank -4; item 27, rank -4, item 26, rank -3). Instead, they explained that therapy should create a safe and non-judgmental space within young people's families – endorsing an idea that was also expressed by Youth Viewpoint B (*Learning to help myself*):

> I usually want to move towards the young person getting a sense of agency to be able to find another space, other than CAMHS […] I think naturally, if they're able to get on better with their family, so the mum or dad or even a sibling about issues without being judged, I think that's more important. (Professional #27)

Consensus Outcomes Amongst Professionals

Some outcomes were endorsed by all professional viewpoints with the exception of viewpoint B (*Empowering young people and parents*), which represented a distinctive minority perspective. Viewpoints A, C, and D agreed that a *reduction in risky behaviour* (item 5) was one of the most important treatment outcomes, assigning rankings between +3 and +4. They did, however, suggest disaggregating this outcome further into *preventing suicide* on the one hand, which they considered imperative; and *reducing non-suicidal self-injury, substance use or other risk-taking behaviours*, which they considered less urgent and likely to occur as a result of other changes. *Knowing ways to cope* was another strongly endorsed outcome across all three viewpoints, with rankings between +2 and +4.

Similar to young people, professionals did not have strong views about *improved peer relationships* and *friendships* (items 22 and 23), with rankings ranging from -1 to +2. These relationships were seen as valuable, but not a primary contributing factor to young peoples' recovery, or a primary indicator of treatment success. Again, professionals considered that these would naturally improve as a result of young people getting better. *Reducing parental stress and worries* was consistently ranked as one of the least important outcomes across all viewpoints (item 32, rankings between -2 and -3). Three out of four also ranked reduced *parental guilt* as not important, with some professionals explaining that improving parental wellbeing was beyond their remit (item 35, rankings between -3 and -4). Viewpoint B (*Empowering young people and parents*) presented a distinctively different interpretation, stressing the importance of parental confidence and wellbeing as a prerequisite for strengthening family support. Improved *physical health* was another outcome that professionals of all viewpoints deprioritised (item 30, rankings from -2 to -4). While many stressed that mental health should be placed in the context of a broader view on health, they considered that improving physical health was not a principal goal of psychotherapy for depression.

6.3.3 Comparing Viewpoints Between Youth and Professionals

To compare viewpoints between young people and mental health professionals, the ideal-typical Q-sorts computed for each of the eight viewpoints described above were treated as raw data in a second-order principal component analysis (Kline, 1994). A four-factor solution was identified, which explained 82.0% of the common variance between the eight viewpoints. As in the group-specific analysis for professionals, the first super factor had both positive and negative loadings, with the two negative-loading Q-sorts being those associated with *Professional Viewpoint B: Empowering Youth and Parents.* As previously, this factor was split in two, and the loadings for one of them were inverted. Because the two Q-sorts associated with this viewpoint also loaded positively on another super factor, they were assigned to this one, and the split-out fifth factor was dropped. To enhance the inclusivity and clarity of the factor solution, small adjustments were performed via hand rotation (see Table D.6 in Appendix D). There was no significant inter-correlation between any of the super factors, suggesting that each represented a distinctive overarching viewpoint (see Appendix D, Table D.7 for details). The characteristics of the extracted super viewpoints following rotation are shown in Table 6.7 (below).

Table 6.7. Characteristics of Extracted Factors – Second Order Analysis

Factor Characteristics	Super Viewpoints			
	S.A	S.B	S.C	S.D
Explained variance (%)	27.0	27.0	16.0	13.0
Cumulative explained variance (%)	27.0	54.0	70.0	83.0
No of defining sorts (*n*)	3	3	1	1
Standard Error of Factor Scores	0.28	0.28	0.45	0.45

Note. S.A. = super viewpoint A; A.B. = super viewpoint B, etc.

The factor matrix and factor loadings for the second order factor analysis are shown in Table 6.8 (below). Amongst both young people and professionals, the first round of analysis identified a viewpoint prioritising symptom change, with some additional emphasis on coping, and restoring a sense of hope and future orientation. Together, these two viewpoints (Y.A: Becoming a healthier person and P.A: Managing risk and reducing symptoms) significantly loaded on an overarching super factor, conveying this symptom-focussed perspective: Super Viewpoint A: Symptoms - Feeling better. This first super factor explained 27% of the overall variance in the second-order factor analysis model.

A second super factor *(Self-management – Resilience through coping skills)* was formed from the ideal-typical Q-sort of the three youth and professional viewpoints that focussed on fostering self-management, coping skills, and improving young people's understanding of their feelings and thoughts (i.e., Y.B: *Learning to help myself; Y.C: Making sense of the past to embrace the future; P.C: Building skills in young people to aid coping*). While this super viewpoint also endorsed outcomes around symptom reduction, these were not ranked as prominently as by Super Viewpoint A (*Symptoms – Feeling better*), with self-management outcomes taking priority.

Professional Viewpoints B *(Empowering young people and parents*) and D (*Building family support around young people*) did not have equivalents amongst young people, and constituted their own super factor (*Super viewpoint C: Parental support – Resilience through family support*). While professionals within this viewpoint assigned relatively high importance to individual coping skills, they considered fostering parental support and understanding to be even more important. However, like Super Viewpoint B, they endorsed a notion that treatment should ultimately improve young people's resilience. Finally, Youth Viewpoint D (*Being able to do what other adolescents do*), which conveyed a desire for depression symptoms to interfere less with daily life, had no equivalent in the professional sample. It constituted its own super factor in the second-order analysis (Super Viewpoint D: *Functioning – less interference with daily life*).

Table 6.8. Factor Loadings Following Rotation – Second-Order Factor Analysis

Super viewpoint	Constituting viewpoints		S.A	S.B	S.C	S.D
S.A: Symptoms – Feeling better	Y.A:	Becoming a healthier person	**.86**	.32	-.21	-.03
	P.A:	Managing risk and reducing symptoms	**.86**	.11	.24	.09
S.B: Self-management – Resilience through coping skills	Y.B:	Learning to help myself.	.20	**.70**	.01	.01
	Y.C:	Making sense of the past to embrace the future.	.16	**.85**	.27	-.09
	P.C:	Building skills in young people to aid coping.	.06	**.82**	-.16	.09
S.C: Parental support – Resilience through family support	P.B:	Empowering young people and parents	-.80	.28	**.51**	-.05
	P.D:	Building family support around young people.	.24	.07	**.90**	-.15
S.D: Functioning – less interference with daily life	Y.D:	Being able to do what other adolescents do	.02	.02	-.14	**.99**

Note. P.A., = professional viewpoint A; S.A. = super viewpoint A; Y.A. = youth viewpoint A. Factor loadings printed in bold were significant at the *p* < .01 level (i.e., loadings above 0.44).

Across the super viewpoints, the most consistently endorsed outcomes related to *feeling less down and depressed*, *feeling happier*, and *learning ways to cope*, which were each ranked in places +2 to +4 by three of the four viewpoints. The only outcome that was consensually ranked important (rank +1 or +2) by *all four* super viewpoints was *being able to make plans for the future and have goals*, which related to symptoms and risk management, but also represented a broader concept of wellbeing and hope. The full ideal-typical Q-sorts for each super viewpoint are displayed in Figure 6.5 (below) and are discussed in further detail in the next section.

Figure 6.5. Ideal-Typical Q-sorts – Super Viewpoints

#	Q-sort item	Item ranks S.A	S.B	S.C	S.D
Symptoms					
1	Being less angry and not losing my temper as much.	0	-2	0	0
2	Feeling less down and depressed.	3	2	0	3
3	Feeling happier and enjoying things more.	4	2	-1	4
4	Feeling more loved.	1	-1	0	-2
5	Engaging less in behaviour that can be harmful.	4	0	4	0
Self-management					
6	Being more active and engaged in things.	0	-1	2	-2
7	Knowing ways to cope with my emotions.	3	4	2	-4
8	Having a better understanding of my feelings and thoughts.	0	3	3	-1
9	Being able to challenge negative thoughts and approach situations differently.	1	4	0	-3
Functioning					
10	Being better able to get things done (e.g., concentrate, be organized).	2	-1	-1	0
11	Being able to do the same things other adolescents do.	-1	-3	-2	4
12	Working more effectively in school (e.g., being more motivated and focused).	0	-2	-1	0
13	Attending school more regularly.	-2	-1	1	-3
14	Being more sociable and better able to be around other people.	1	-2	0	1
Personal growth					
15	Feeling more confident.	2	0	0	3
16	Being better able to stand up for my needs and opinions.	-1	-2	1	-2
17	Being more independent and able to take responsibility for my life.	-1	1	1	0
18	Being able to make sense of things that have happened in the past, or are still happening.	0	3	-1	1
19	Having a better sense of who I am and how to be myself around others.	-1	0	2	1
Relationships					
20	Feeling more able to talk about my feelings and thoughts.	0	2	0	2
21	Getting on better with my family	0	0	3	-4
22	Getting on better with my friends or having made new friends.	2	-3	1	-1
23	Getting on better with my peers in school (e.g., not feeling bullied).	0	-3	1	-3
Therapeutic space					
24	Having a space where someone listens and cares about me.	-2	0	-4	-2
25	Having a space where I can let out my feelings.	-3	0	-2	2
26	Having a space where I can talk about anything without being judged.	-1	1	-3	-1
27	Having a space to reflect and think about things differently.	-3	1	-4	-1
Wellbeing					
28	Having greater peace of mind (e.g., feeling calmer, more balanced).	1	1	-2	-1
29	Feeling more optimistic and positive about life and the future.	3	3	-1	2
30	Feeling physically healthier.	-2	-4	-2	0
31	Being able to make plans for the future and have goals.	2	1	2	2
Parental support and wellbeing					
32	My parents feeling happier and less stressed and worried.	-4	-1	-3	1
33	My parents having a better understanding of me and my difficulties.	-3	2	3	0
34	My parents feeling more able to support me.	-2	0	4	1
35	My parents feeling less guilty.	-4	-4	-3	3

- S.A: Symptoms – Feeling better
- S.B: Self-management – Resilience through coping skills
- S.C: Parental support – Resilience through family support
- S.D: Functioning – Less interference with daily life

Note. S.A. = super viewpoint A; A.B. = super viewpoint B, etc. Gradient colour coding has been applied to highlight items ranked in positions +4, +3 and +2.

6.4 Discussion

6.4.1 Summary of Findings

This study explored outcome priorities amongst young people and mental health professionals in relation to adolescent depression. The aim was to identify which treatment outcomes were considered most important, and whether there were distinctive viewpoints on this question amongst and between young people and professionals.

The Q-study revealed considerable diversity in priorities amongst young people and professionals. Four distinctive viewpoints were identified in each group, which converged to form four cross-cutting super viewpoints. Super Viewpoint A (*Symptoms – Feeling better*) represented young people and professionals for whom the most important outcomes were reduced depression symptoms and improved wellbeing. Super Viewpoint B (*Self-management – Resilience through coping skills*) represented young people and professionals who suggested that treatment should focus on fostering coping skills, cognitive-behavioural techniques, and a deeper understanding of feelings and thoughts to promote young peoples' capacity for self-management. Super Viewpoint C (*Parental support – Resilience through family support*) was represented by professionals only and suggested treatment should promote resilience by building strong support networks around young people, with special attention to family and parental support. This viewpoint reflected a systemic understanding of how young people may best be strengthened. Super Viewpoint D (*Functioning – less interference with daily life*) represented a small group of young people who were struggling to cope, and whose primary desire was for their mental health difficulties to interfere less with their daily lives and aspirations.

The adult recovery literature suggests that a symptom-focussed understanding of good outcome is typically endorsed by mental health practitioners, while service users value more personalised outcomes (Pilgrim & McCranie, 2013; Schrank & Slade, 2007; Stotland et al., 2008). The same has been suggested by existing qualitative research with young people, conducted following their experience of treatment (Bergmans et al., 2009; Cortés et al., 2018; Gibson & Cartwright, 2014; Lavik et al., 2018). This dualism was, however, not reflected in this study's findings, which revealed a more complex picture.

Second order factor-analysis identified a super viewpoint amongst professionals and young people focussed on symptom reduction that explained 27% of the overall variance in the second order factor analysis. Amongst young people, *feeling happier and enjoying things more* was the most consistently endorsed outcome. While professionals conveyed a stronger focus on risk management than young people did, this was still also one of the most highly ranked outcomes in two out of four youth viewpoints. At the same time, three out of four professional viewpoints conveyed broader perspectives on outcome, focussing on empowerment, coping, and family support, respectively.

The strong focus on coping and self-management conveyed by Super Viewpoint B *(Self-management – Resilience through coping skills*)*,* and the frequent endorsement of coping as an outcome across other viewpoints is consistent with existing research. Young people have frequently endorsed the importance of learning to cope with challenges, and of better understanding their feelings and thoughts in studies examining treatment goals (Jacob et al., 2016), outcome concepts (Lavik et al., 2018), and perceptions of recovery (Bergmans et al., 2009). A qualitative study involving adult service users who had completed a course of CBT for depression equally emphasised the importance of developing cognitive coping strategies, and of learning to break negative thought cycles (Glasman et al., 2004). Indeed, a common theme was that study participants began to describe themselves as “self-therapists”. They demonstrated a sense of self-efficacy that youth in this study also emphasised. Coping and self-management were considered less important outcomes by youth representing Viewpoint D (*Being able to do what other adolescents do*), who experienced the highest burden from mental health difficulties. They described that their low mood would frequently become so overwhelming that deliberate coping was no longer possible (e.g. due to low self-efficacy) or effective. This resonates with findings from adult research, which suggest that service users struggle to apply cognitive-behavioural strategies when feeling particularly low or overwhelmed (De Smet et al., 2019; Glasman et al., 2004). Indeed, those who most strongly endorsed coping and self-management outcomes in the present study were the youth with the lowest comparable burden from depression and other mental health difficulties.

In the Post-Therapy Outcome Perspectives study, 62% of young people and 29% of clinicians discussed outcomes in relation to the therapeutic space and process, with young people frequently emphasising the importance of feeling listened to and cared for, which for many constituted a positive outcome in and of itself. Young people also emphasised the value a therapeutic safe space in a number of other studies investigating helpful and unhelpful aspects of therapy (Binder et al., 2011; Dunne et al., 2000; Freake et al., 2007; Gibson et al., 2016). In this Q-study, however, outcomes related to therapeutic process were generally not ranked highly. When participants had to choose between items related to therapeutic process and items related to therapeutic outcome, they tended to prioritise the latter, frequently describing the therapeutic process as a means to achieve higher-level outcomes. Super Viewpoint B *(Self-management – Resilience through coping skills*) was the only one to rank these process outcomes neutrally to slightly positively (with ranks between 0 and 1), while all other super viewpoints ranked process outcomes as unimportant.

A further unexpected finding of this study was that neither young people nor clinicians emphasised the outcome categories of *friendships* and *peer relationships* as particularly important. This was contrary to findings by Lavik and colleagues (2018), whereby having deeper and more meaningful relationships with others constituted one of five key outcome themes. The importance of connectedness with others was equally emphasised in the adult recovery literature (Binder et al., 2010; Leamy et al., 2011; Timulak & Creaner, 2010). While *friendships* was one of the seven most frequently discussed outcome categories amongst young people in the Post-Therapy Outcome Perspectives study, participants in the Q-study did not frequently prioritise this outcome.

Several of the viewpoints identified in this study reflected a developmental perspective on treatment outcomes, emphasising that therapy should help young people become more resilient and resourceful. This resonates with suggestions by Friesen (2007) that resilience may be an important concept to consider alongside recovery, in child mental health. It also resonates with calls for more developmentally informed outcome frameworks that reflect both the dynamic nature of childhood and the importance of the young peoples' family environment (Fonagy, 1997; Hoagwood et al., 1996).

6.4.2 Implications for Clinical Practice

Viewpoints prioritising outcomes related to coping and symptoms were identified amongst young people and clinicians, suggesting that the divide between both groups is not as stark as suggested by the adult recovery literature (Pilgrim & McCranie, 2013; Secker et al., 2002). This existence of shared viewpoints amongst young people and clinicians is encouraging. It suggests that some core outcomes can be identified that resonate with both groups. At the same time, two super viewpoints were identified that each appeared rooted in only one of the two stakeholder groups. In a clinical context, a young person prioritising coping may well encounter a clinician focussing on symptoms or family support. Embedding open communication and shared decision-making about outcome priorities in clinical practice is crucial to ensure that each party's expectations are made clear, and that outcomes can be measured in a way that is meaningful to both. A shared decision-making tool presenting different possible outcomes may help guide such conversations. When administering the Q-sort to professionals as part of the present study, several participants suggested using the Q-sort stimuli for this purpose in clinical practice.

Most viewpoints identified by this study conveyed a multidimensional understanding of outcome, endorsing a variety of changes beyond the reduction of depression symptoms. In clinical practice, service users and professionals endorsing outcomes beyond symptom change may be sceptical about unidimensional approaches to measurement. Research with service users and their parents suggests that outcome measurement is only considered helpful if it is more than a box-ticking exercise (Moran et al., 2012). Outcome measurement should thus assess changes that matter to service users. However, this study has also showcased that service users are heterogeneous in their hopes and aspirations. To ensure that outcome measurement is meaningful to all of them, one way forward would be to routinely measure a broad range of outcomes. This, however, poses the risk of imposing an excessive burden by measuring outcomes that are not relevant to everybody (Kiresuk & Sherman, 1968). Instead, a balance might be struck between identifying a minimal core set of outcomes relevant to most stakeholders, and enabling a degree of personalisation to flexibly include other outcomes that are meaningful to service users, or in the context of testing a specific treatment mechanism. This might be done through idiographic outcome measures, as discussed in Section 1.7.1.

As mentioned in Chapter 1, several initiatives are currently underway to develop core outcome sets for depression in children and young people to harmonise measurement in clinical practice and research. One project is led by ICHOM (see www.ichom.org) and has assembled an international working group of experts by profession and experience to devise a core set for practical use (Krause et al., 2020). The second initiative is led by Monga and colleagues (2019) at TORCH, Hospital for Sick Children Research Institute, and the Cundill Centre for Child and Youth Depression, Centre for Addiction and Mental Health, in Toronto (see http://www.comet-initiative.org/studies/details/1122). It focusses on clinical trials. Both initiatives use Delphi-type consultation processes (Linstone & Turoff, 1975) to generate consensus on a core outcome set, with members of a reference group repeatedly voting on the types of outcomes and measures to recommend for inclusion (Bolger et al., 2011; Landeta et al., 2011; Linstone & Turoff, 2011; Rowe & Wright, 2011). These initiatives can make an important contribution to streamlining measurement, and generating wider consensus about core outcomes that are generally considered important. By focusing on consensus outcomes, however, they risk masking diversity in viewpoints, such as those identified by this study. It is important to emphasise that such sets present a *minimum recommendation*, and that measurement should be tailored further. In clinical research, additional outcomes may be assessed that relate to key treatment mechanisms (Fonagy, 1997; Kazdin, 1999a). In clinical practice, measurement should be personalised, for example through the use of idiographic outcome measures.

6.4.3 Limitations

The above-mentioned findings should be interpreted in the context of a number of limitations. The first set of caveats relates to the possibility of recruitment bias. This study employed opportunistic snowball sampling (Biernacki & Waldorf, 1981), meaning that participants self-selected into the study. It is possible that young people who participated voluntarily shared an interest in outcome research, and a certain level of confidence in speaking about their experiences. The inclusion of young people in peer-support groups moderated this risk to an extent, as they were approached as a group, and did not have to pro-actively contact the research team to arrange a meeting, which may have required a considerable level of personal motivation and confidence. All youth participants had managed to access some variation of mental health support, and were willing to speak about this experience without fear of stigma. Depressed young people who do not access support may have distinctively different viewpoints, and demographic groups that tend to be under-represented in child mental health services are also likely to be underrepresented in this study. Finally, views of what constitutes a good outcome may fluctuate, depending on the symptom burden young people are experiencing at any given time, and the stage that they are at in their recovery journey. Young people who participated in this study felt well enough to do so. Young people in crisis might have expressed different outcome perceptions and priorities. Similarly, professionals who volunteered their time for the study often expressed an interest in outcome measurement, and may not be representative of the general population of mental health professionals working with young people in the UK. Those with reservations against routine outcome measurement, in particular, may be underrepresented.

Second, the Q-set for this study was derived using a rigorous multi-stage process that drew on the literature, as well as accounts of different stakeholder groups (young people, clinicians, parents), and consultations. Nevertheless, the final Q-set represents a subset of possible outcomes, while others may have been missed. Participants suggested the inclusion of additional symptoms. Four participants suggested *improved sleep and energy levels*, two suggested a separate item on *suicidal ideation and behaviour*, and others suggested *improved appetite and nutrition*. Other suggestions included *feeling personally safe*, *having more structure in life*, *being able to feel compassion for others*, and *coping more effectively with academic pressure*. Within the *relationship* domain, participants suggested adding *young people feeling less rejected by their parents*, *young people having a support network in their extended family*, and *having identified a trusted adult*. At the process level, outcomes suggested to be missing included *young people actively engaging in decision-making about their care*, the *strengthening of institutional support networks* (e.g., involving social workers and after care), and *therapy achieving better communication and understanding about the child's difficulties between school and parents*.

Third, although originally designed to include young people from the age of 12 years, the study was unable to recruit participants under the age of 16. Younger adolescents were rarely present in the organisations and structures through which recruitment took place, and none responded to the study advertisements on social media. The peer support groups visited, while open to younger adolescents, generally did not have younger people in attendance when the research was conducted. The need to obtain parental consent ahead of time was another barrier to engaging adolescents under the age of 16 years (Heath et al., 2007). In future studies, a different recruitment strategy may be more successful at obtaining the views of this group, who might be easier to engage through recruitment in clinical settings.

Fourth, as mentioned in Section 6.2.6, this study was originally designed to include parents as a third stakeholder group, and the item set was developed with this objective in mind. However, recruitment of parents through similar channels as those used for young people and professionals was unsuccessful. The number of parent peer support groups and mental health associations in the UK is limited, especially outside of groups that are attached to statutory services. Parents were less responsive to calls for participation that were disseminated via social media, or through advocacy networks. Another study aspiring to recruit a UK-based parent sample in parallel with this study met similar challenges (Liverpool et al., 2019). Longer recruitment times may be required to achieve desired parent samples in this context.

Fifth, Q-methodology is not suited for informing generalisations about the distribution of viewpoints in the wider population. Q-studies do not produce large-scale datasets and do not judge the validity of a viewpoint in relation to the number and representativeness of those subscribing to it (Baker et al., 2006). Nevertheless, it allows for generalisation “with respect to the subjectivity at issue", that is about the existence of these particular viewpoints in one segment of the population, providing “unequivocal evidence” of its existence in these groups (S. R. Brown, 2005, p. 202). This study does not claim that the viewpoints identified are exhaustive or representative of all young people seeking help for depression in the UK, or all professionals supporting them. However, it provides a basis upon which new and more informed hypotheses can be built (Stenner et al., 2003).

Sixth, this Q-study asked participants to sort their outcomes into a fixed quasi-normal distribution, with a limited number of slots provided under each rank. This approach facilitates the sorting for participants, and analysis for the researcher. However, forced rankings have been criticised for artificially increasing the correlations between ranks across respondents by the forced nature of the instruments, rather than by participants freely placing items in similar positions (Kampen & Tamás, 2014). Another complication is that principal component analysis considers each item rank as equally valid and informative in identifying sorting patterns across participants. If a Q-set includes items that participants consider redundant or duplicative, they may choose to assign one of these items to a high rank of importance, while deprioritising the other. Principal component analysis would then interpret the latter’s low position as face-valid, and create a spurious association with participants who considered this outcome to be truly unimportant. Nevertheless, the factors identified by this Q-study were theoretically plausible, and their interpretation drew closely on the post-sort interviews conducted with participants to increase their validity.

6.4.4 Future Research

To further explore the generalisability of the present findings, future research may incorporate descriptions of the viewpoints identified in this study into a quantitative questionnaire, along with relevant demographic variables, to examine their prevalence in a larger, representative sample of youth with lived experience of depression (S. R. Brown, 2002; Danielson, 2009). While the influence of gender or ethnicity on outcome perceptions was not the focus of the present study, examining similarities or differences in viewpoints for different demographic groups, and with special attention to hard-to-reach groups, such as refugee or migrant populations, is another important area for future research.

Many participants in this study reported that they enjoyed the Q-sorting process, and several commented on the potential benefit of using this method in clinical practice to clarify outcome priorities and expectations in a playful and non-threatening way. There is potential to further develop the existing Q-set and to convert it into a decision aid for use in service settings. Further research is required to explore the effectiveness of such a decision-aid at improving communication about outcomes amongst clinicians and service users, and at making outcome measurement more meaningful and relevant. Further research is also needed to better understand how outcomes in different domains could best be measured, which existing tools might be suited, and whether using idiographic measures might be a preferred solution for personalising outcome measurement in the absence of existing relevant tools.

6.5 Conclusion

This study found diversity in viewpoints on what treatment outcomes were most important in the eyes of young people and clinicians. While the adult recovery literature has suggested that clinicians tend to focus on symptom change and functioning, this study found considerable diversity in professionals' perspectives, with three viewpoints emphasising the importance of other outcomes, such as empowerment, self-management, family support and resilience building, alongside symptom change. Inversely, young people consistently highlighted symptom change as one of the most important outcomes.

To ensure that outcome measurement in clinical practice captures change that matters to young people and professionals, a multidimensional perspective may be required when selecting outcomes for measurement. Given that there appears to be a high degree of agreement on the importance of symptom change and self-management, it may be possible to create a core set of outcome measures to guide and harmonise such approaches. At the same time, the diversity in viewpoints and the range of other outcomes variantly considered important suggest that an additional element of personalisation may need to be built into such a core set in order to ensure that what is being measured reflects what is most meaningful to young people and those working with them.

This study examined agreement between young people and clinicians with regards to the importance of different outcome domains. The next study, presented in Chapter 7, moves on to the actual assessment of outcomes, and examines agreement in change metrics obtained using different youth-reported instruments, and across different outcome domains.

Chapter 7. Comparing Change Across Outcome Domains

7.1 Introduction

If symptom change was an adequate proxy for change in other outcome domains, this could justify a continued focus on symptoms to minimise the burden of routine measurement in clinical practice and promote harmonisation in research studies. Existing studies, however, suggest that symptom change does not necessarily translate into changes in other outcome domains (Becker et al., 2011; Brookman-Frazee et al., 2006; Edbrooke-Childs et al., 2015; Karpenko & Owens, 2013; Shapiro et al., 1997). Two outcome domains that have frequently been suggested for additional measurement alongside symptom change are functioning and idiographic (i.e., personalised) outcomes.

7.1.1 Functional Impairment

Functioning is the second most frequently measured outcome domain after symptoms in recent clinical research on treatment effectiveness for adolescent depression (see Outcome Measurement Review, Chapter 4), and in child mental health more broadly (Becker et al., 2011; Hoagwood et al., 2012; P. S. Jensen et al., 1996; Weisz et al., 2005). Several reasons for measuring functioning alongside symptom change have been mentioned in Chapter 1, including identifying a diagnosis of depression, influencing help-seeking, and the allocation of services (Hodges, Kay; Doucette-Gates, Ann; Kim, 2000; Striley et al., 2003). In addition, findings from the Post-Therapy Outcome Perspectives study (Chapter 5) suggest that academic functioning in particular is a salient outcome for youth, parents and clinicians. Findings from the Viewpoints on Outcome Priorities study (Chapter 6) indicate that functional impairment may be a particularly important outcome for youth with a high burden from depression and comorbid difficulties, who struggle to cope. A literature review suggests moderate to strong correlations between measures of depressive symptoms and functioning across clinical samples of different ages (McKnight & Kashdan, 2009). This suggests that while these measures capture some common variance, a considerable part of variance is not shared, which points to the complementarity of measuring both outcome concepts (McKnight & Kashdan, 2009).

Two studies have examined the degree of convergence between change in symptoms and change in functioning in child mental health. In a sample of 112 adolescent outpatients, Brookman-Frazee and colleagues (2006) found minimal agreement in individual-level change ratings across the three domains of symptoms, functioning and family relationships. Amongst the 31 young people who showed meaningful improvement in any of these three domains, only six improved in another domain, and two improved across all three domains. A systematic review of randomised controlled trials testing psychosocial and combined treatments in child mental health assessed the strength of the evidence they provided in favour of a treatment effect on symptoms and on functional impairment, respectively (Becker et al., 2011). Where both outcomes were measured, the evidence in favour of a change in functioning was considerably weaker than that for symptom reduction. The authors concluded that it was more difficult to provide even minimal empirical support for positive changes in functioning, compared to changes in symptoms. Several studies with adult populations have further found that impairment in areas such as social or vocational functioning resolves more slowly than depressive symptoms (Bothwell & Weissman, 1977; Hirschfeld et al., 2002, 2000; McKnight & Kashdan, 2009; Scott et al., 2000). This existing research suggests limited convergence between change in symptoms and in functioning. However, the study by Brookman Frazee and colleagues (2006) included only a small sample and did not present separate analysis on the extent of convergence between symptoms and functioning, regardless of changes in family functioning. Becker and colleagues (2011) commented on the strength of the evidence base, but did not examine convergence of change ratings at the individual level. None of these previous studies focussed specifically on adolescent depression.

7.1.2 Idiographic Outcomes

As discussed in Chapter 1, idiographic outcome measures track change in relation to issues that individual service users themselves determine. They complement nomothetic measures by ensuring a person-centred assessment focussed on each service user's individual priorities (King et al., 1999; Sales & Alves, 2012). As such they can help reduce the burden that would stem from administering batteries of standardised items, of which only some may be relevant to individual service users (Kiresuk & Sherman, 1968; Sales & Alves, 2016). A characteristic feature of idiographic outcome measures is their increased sensitivity to change, compared with nomothetic measures, by nature of centring around issues that are meaningful to service users and that treatments are likely to target (Ashworth et al., 2004; Godfrey et al., 2019; Lacasse et al., 1999).

Superior sensitivity to change was demonstrated in a child mental health context by Edbrooke-Childs and colleagues (2015) who analysed naturalistic outcome data of 137 children and young people who had accessed CAMHS in the UK. The study compared parent-reported changes on the idiographic Goal Based Outcome Measure (GBO; Law, 2006) and on nomothetic outcome measures of functioning and psychosocial difficulties (i.e., the Strengths and Difficulties Questionnaire; R. Goodman et al., 1998). Based on comparisons of effect sizes, change score correlations and rates of reliable change in the aggregate sample, Edbrooke-Childs and colleagues concluded that the GBO showed higher sensitivity to change and suggested higher rates of goal progress, compared with improvements in the domains of symptoms and functioning. This is consistent with studies in adult populations that have equally demonstrated higher sensitivity to change in idiographic outcome measures (Ashworth et al., 2005; Elliott et al., 2016; Paterson, 1996). Indeed, a systematic review of studies using idiographic goal attainment scaling (Kiresuk & Sherman, 1968) with adult patients in physical and neurological rehabilitation found that seven studies provided evidence of goal attainment scaling exceeding nomothetic measures in their sensitivity to change (Hurn et al., 2006).

While the above-mentioned studies examined convergence by correlating mean change scores, or by comparing effect sizes or reliable change at the group level, Karpenko and Owens (2013) assessed the convergence of two idiographic measures and one nomothetic measure at the individual level. Drawing on self-reported data from 117 adolescent outpatients with a range of presenting problems, they compared change measured on the nomothetic Problem Severity subscale of the Ohio Scales (Benjamin et al., 2001), the idiographic Target Complaints (Battle et al., 1966), and three items measuring subjective perceptions of change (Karpenko & Owens, 2013). While there was a significant moderate correlation between change in youth-rated symptoms and perceived change (r = -0.38) as well as symptoms and target problems (r = 0.40) over a 3-month period, there was considerable divergence at an individual level. Cohen's kappa indicated low agreement of ratings of clinically significant change on the symptom measure and meaningful change on the two idiographic measures (κ = 0.33 and 0.09 for perceived change and target problems, respectively). One third of young people who did not show significant change in symptoms did demonstrate meaningful change based on ratings of perceived change and target problems.

Both of these existing studies were limited by relatively small samples, which included youth with a range of presenting problems. The study led by Edbrooke-Childs and colleagues (Edbrooke-Childs et al., 2015) further included young children, and only considered parent-report, at the expense of exploring convergence or divergence between outcome domains as rated by young people themselves. None of the two studies provided a detailed assessment of divergence or convergence at individual level using the reliable change index.

7.1.3 A Real-World Example of Multidimensional Outcome Measurement

In the UK, multidimensional outcome measurement across the domains of symptom change, functioning, and progress towards idiographic goals was encouraged as part of a large-scale service transformation initiative led by the National Health Service (NHS) between 2011 and 2015. The Children and Young People's Improving Access to Psychological Therapy (CYP IAPT) programme aimed to embed elements of good practice in participating CAMHS in England, including the routine measurement of outcomes based on child- and parent-reported measures (Law & Wolpert, 2014; Wolper, Fugard, et al., 2012; Wolpert et al., 2016). Services were encouraged to select from 13 nomothetic measures across the domains of symptoms, functioning, and quality of life, and to also consider measuring progress towards self-defined goals (Law & Wolpert, 2014; Wolpert, Fugard, et al., 2012).

Within the domains of symptoms and functioning, a number of alternative measures were suggested from which clinicians could select those based on the needs of individual cases (Law & Wolpert, 2014). Measures of symptom change relevant to young people with depression included the *Revised Children's Anxiety and Depression Scale* (RCADS; Chorpita et al., 2000) and the *Strengths and Difficulties Questionnaire* (SDQ; R. Goodman, 1997). Measures of functioning included the *SDQ Impact Supplement* (R. Goodman, 1999) and the

Outcome Rating Scale (Miller et al., 2003) in its child (CORS) and youth versions (ORS). In addition, the GBO were recommended as an idiographic measure of progress towards self-defined goals (Law, 2006). Services and clinicians were not mandated to use specific measures or combination of measures, based on the understanding that despite varying strengths and weaknesses, these measures could be treated as broadly equivalent in assessing the outcome domains of interest (Wolpert, Cheng, et al., 2015; Wolpert, Fugard, et al., 2012). To facilitate comparisons across different measures, the initiative recommended the reporting of outcomes at an individual level in terms of the reliable change index. This assumption of equivalence has since been built into the development of a new NHS Mental Health Services Data Set (MHSDS) outcome metric, which will aggregate reliable change on different nomothetic measures, as well as the idiographic GBO, into one overall change metric (Jacob, 2019).

At the same time, however, it has been acknowledged that there is limited evidence on the extent to which measures that purport to measure the same thing actually do so, as well as the extent to which change in one domain can be seen as interchangeable with change in another domain (Wolpert, 2017; Wolpert, Görzig, et al., 2015). Similarly, evidence on the convergence of change measured through nomothetic and idiographic instruments is scarce. Only two studies examined this issue, both of which were limited by small sample sizes and populations with mixed presenting problems. The study by Karpenko and Owens (2013), while assessing divergence at an individual level, did not consider the domain of functioning. Edbrooke-Childs and colleagues (2015) did consider functioning, however based on parent report and without undertaking detailed analysis of divergence and convergence at an individual level.

7.1.4 The Present Study

This study drew on the naturalistic outcomes data collected through the CYP IAPT initiative from CAMHS in England. Building on previous research, it examined the extent of convergence between different measures and outcome domains in assigning ratings of reliable improvement at an individual level. In contrast to previous studies, this study only considered outcome data collected through child self-report. It is the first such study to focus specifically on adolescent depression.

This study addressed the following three research questions:

1. How do levels of reliable change compare between two symptom measures?
2. How do levels of reliable change compare between two functioning measures?
3. How do levels of reliable change compare across the domains of symptoms, functioning, and goal progress?

Good outcome was defined as reliable or meaningful change at an individual level (Evans et al., 1998; S. A. Jensen & Corralejo, 2017), as described in Chapter 1. Two measures of emotional problems (RCADS and SDQ Emotions), two measures of global functioning (SDQ Impact and C/ORS), and one measure of progress towards self-defined goals (GBO) were considered (see below for further details). All five tools are widely used across a range of mental health settings in England, with between 40% (GBO and C/ORS) and 56% (SDQ) of recently surveyed practitioners reporting to have used them (Costa da Silva & Wolpert, 2018).

7.2 Methods

7.2.1 Participants and Process

This study drew on naturalistic outcome data submitted as part of the CYP IAPT initiative between 2011 and 2015 by CAMHS in England. The data were collated and stored by the Child Outcomes Research Consortium (CORC) – a professional learning collaboration of children’s mental health services in England. The full dataset included 96,325 case records of which 23,373 were seen for more than just assessment (Wolpert et al., 2016). For inclusion in this study, cases had to meet the following inclusion criteria: (a) aged 12-21 years at assessment, (b) an initial rating of moderate or severe depression on the clinician-rated Current View screening tool, and (c) could contribute paired data to at least one of the three comparative analyses conducted as part of this study. The latter was true if a service user had paired data on the SDQ Emotion and the RCADS (for comparison within the symptom domain); or on the SDQ Impact and the C/ORS (for comparison within the functioning domain); or on at least one of the two symptom measures, at least one of the two functioning measures, and on the GBO (for comparison across domains). “Paired data” refers to the same measure having been completed on at least two separate occasions.

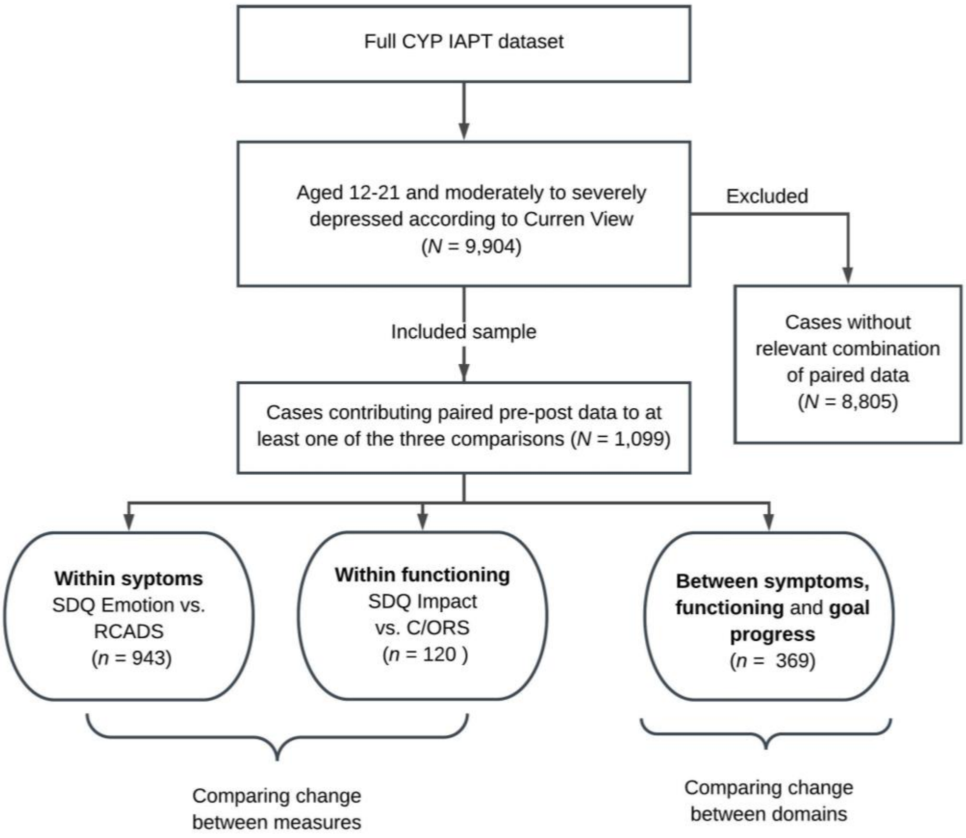
Overall, 9,904 case records referred to adolescents with moderate to severe depression ratings on the Current View and were thus eligible for inclusion in this analysis. Because services were not mandated to use particular measures, different tools were administered in a variety of combinations, and reasons for selecting one measure over another were not systematically recorded. As a result, comparisons requiring the availability of paired data on several measures or in several domains are faced with small subsamples (see Appendix E for more detail on missing value patters in the dataset of eligible cases). Of the 9,904 eligible cases, 1,099 could contribute paired data to at least one of the three comparative analyses conducted as part of this study (see Appendix E, Table E.3 for a detailed overview of data availability patterns within this included sample).

- 943 cases had paired data on the SDQ Emotion and on the RCADS and were considered for comparing change within the symptom domain.
- 120 cases had paired data on the SDQ Impact and on the C/ORS and were considered for comparing change within the functioning domain.

- 369 cases had paired data in the domains of symptoms, functioning, and goal progress and were considered for comparing change across outcome domains.

To obtain a sufficiently large analytical sample for the comparison of change across the three outcome domains, and to use available data in the most efficient way possible, a composite reliable change index was computed for symptoms and functioning, by drawing on available data from at least one of the two symptom measures and one of the two functioning measures (see Section 7.2.6). Of the 369 cases considered for the cross-domain comparison, some had data on both symptom and/or both functioning measures and could also be considered for the within-domain comparisons, while others had data on only one symptom measure and/or or only one functioning measure, and were considered for this comparison only. The analytical process and a breakdown of the overall study sample into the three analytical subsamples are illustrated in Figure 7.1.

Figure 7.1. Flowchart of the Analytical Process



7.2.1 Participant Characteristics

Of the 9,904 adolescents who showed signs of moderate to severe depression at assessment, 1,099 could contribute paired data to at least one of the three change comparisons. Of these, 77.5% were female, 86.3% identified as white British, and the mean age was 14.8 years (*SD* = 1.43). On the Current View, 82.7% were rated as moderately depressed and 17.3% as severely depressed (see Table 7.1). Data originated from 57 CAMH services and the average length of contact was 31.6 weeks. Another 8,805 depressed adolescents did not have a pre-post data on a relevant combination of measures to be considered for analysis. The included and excluded samples differed slightly but significantly in composition: Service users in the included sample were more likely to be female (χ^2 = 19.5, p < .001), to identify as white British or other white (χ^2 = 9.4, p = .002), and to be classified as severely rather than moderately depressed (χ^2 = 18.3, p < .001). The included sample was also slightly younger, on average, than the excluded sample (t = 3.4, p < .001).

Three subsets of the included sample of 1,099 cases were considered for the comparisons of change within the symptom domain (*N* = 943), within the functioning domain (*N* = 120) , and between the domains of symptoms, functioning, and goal progress (*N* = 369). The share of youth rated as severely depressed was slightly larger in the sample for the comparison of symptom measures (18.2%) than in the other two subsamples (14-15%). The small sub-sample of youth considered for the comparison of functioning measures had a longer average contact length (35.1 weeks) than the other two subsamples (around 31 weeks), and a lower share of females (74.2%, compared with roughly 77%).

7.2.2 Ethical Review

As this study was a secondary analysis of routinely collected data, ethical review was not required (National Health Service, 2018).

Table 7.1. Characteristics of the Included and Excluded Samples

Characteristics	Excluded Sample	Included Sample	Symptom comparison	Functioning comparison	Domain comparison
	N = 8,805	*N* = 1,099	*N* = 943	*N* = 120	*N* = 369
Sex (% female)	74.1%	77.5%	77.1%	74.2%	77.2%
Ethnicity (% White British) [a]	83.9%	86.3%	86.4%	83.7%	85.1%
Current View depression rating					
% Moderately depressed	84.5%	82.7%	81.8%	85.0%	85.9%
% Severely depressed	15.46%	17.3%	18.2%	15.0%	14.1%
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Age (in years)	15.0 (1.51)	14.8 (1.43)	14.9 (1.44)	14.8 (1.38)	14.9 (1.40)
Mean contact length (in weeks) [b]	N/A	31.6 (20.78)	31.3 (21.01)	35.1 (20.58)	31.2 (20.02)
No of services	75	57	56	25	46

[a] Due to non-response, the denominator for calculating ethnic composition was 6,670 in the excluded sample; 896 in the included sample; and 763, 98, and 315 for the three subsamples, respectively.

[b] Length of contact was computed based on the dates of the very first and very last assessment completed on the RCADS, the SDQ or the C/ORS.

7.2.3 Measures of Symptoms and Presenting Problems

While the other studies conducted as part of this doctoral thesis focused specifically on depression, the present study considered two measures of broader emotional problems within the symptom domain, thus covering depression and anxiety symptoms (i.e. the RCADS and SDQ Emotions, see below). While the depression-specific Patient Health Questionnaire-9 (PHQ-9; Spitzer et al., 1999) was recommended as part of the CYP IAPT initiative, this was only filled in by a small number of young people and therefore did not lend itself for analysis. The SDQ Emotion and RCADS were selected given higher levels of data quality. While the RCADS can be disaggregated into a separate depression sub-scale, this is not possible for the brief, 5-item SDQ Emotion. For this reason, the symptom domain was defined here with regards to a broader concept of emotional problems, using the total scores of both measures.

Current View

The Current View is a brief clinician-rated screening tool designed to provide a snapshot assessment of presenting problems, complexity factors and contextual problems (Jones et al., 2013). Clinicians complete the measure at first contact, drawing on all available information, including other completed measures. The assessment of presenting problems includes 30 problem descriptions which map onto ICD-11 diagnostic criteria that the developers considered relevant to children and young people. Clinicians rate each presenting problem according to the associated severity of distress and functional impairment, compared with the general population. Severity is rated on a scale from 0 (*none*) to 3 (*severe*). Depression is identified through problem description number 9: “Depression/low mood (Depression).”

Revised Children's Anxiety and Depression Scale (RCADS) – Total Score

The RCADS is a 47-item self-report scale for youth aged 8-18 years, measuring the frequency of symptoms associated with depression and anxiety (Chorpita et al., 2000). It consists of six subscales measuring symptoms of depression (10 items), generalized anxiety disorder (6 items), separation anxiety disorder (7 items), social phobia (9 items), panic disorder (9 items), and obsessive-compulsive disorder (6 items), in line with DSM-IV dimensions. Symptom frequency is scored on a four-point Likert scale ranging from 0 (*never*) to 3 (*always*). Subscale scores can be summed to compute overall anxiety and depression scores, as well as a total score of internalising symptoms. The RCADS has demonstrated good internal consistency, test-retest reliability, and construct validity (Chorpita et al., 2005; Chorpita et al., 2000; de Ross et al., 2002; Esbjorn et aj., 2012; Kösters et al., 2015).

Strengths and Difficulties Questionnaire (SDQ) – Emotional Problems Subscale

The Strengths and Difficulties Questionnaire (SDQ) is a 25-item self-report measure originally designed as a screening tool for behavioural, emotional and relationship difficulties in children and young people aged 4 to 16 years (R. Goodman et al., 1998; R. Goodman, 1997). It includes a five-item emotional symptoms subscale, which taps into unhappiness (*I* *am often unhappy, down-hearted or tearful*), worries (*I worry a lot*), clinginess (*I am nervous in new situations. I easily lose confidence*), fears (*I have many fears, I am easily scared*), and somatic symptoms (*I get a lot of headaches, stomach-aches or sickness*) (A. Goodman et al., 2010). Each problem description is scored on a 3-point scale, from 0 (*not true*) to 2 (*certainly true*) and can be summed to obtain a total subscale score ranging from 0 to 10. While the SDQ as a whole has been widely used and validated with regards to its internal consistency, test-retest reliability, construct validity and predictive validity (e.g., Achenbach et al., 2008; A. Goodman & R. Goodman, 2009; R. Goodman et al., 2000; R. Goodman & Scott, 1999; Mullick & R. Goodman, 2001; Muris et al. 2003; Yao et al., 2009), internal consistency for the emotional symptoms subscale has been shown to be questionable, with a Cronbach's Alpha of 0.66 (R. Goodman, 2001).

Hypotheses About the Equivalence of the two Symptom Measures

The RCADS and the SDQ Emotion subscale both measure internalising symptoms. The RCADS provides a more detailed and sophisticated assessment. It covers symptoms related to major depression with 10 dedicated items, and five specific anxiety disorders with six to nine items each. The SDQ Emotion, in contrast, consists of just one item measuring low mood; three items capturing fears, worries, and clinginess; and one capturing somatic symptoms. Several symptom clusters that are covered by the RCADS, are omitted by the SDQ Emotion (e.g., symptoms related to obsessions and compulsions, or social phobia). As a result, change in these areas may be missed by the shorter measure. Previous research suggests that brief and broadly defined measures such as the SDQ are less likely to capture treatment effects than more focussed measures developed to assess symptoms associated with a specific disorder (Lee et al., 2005).

Since both measures are purported to capture change in internalising symptoms, it is hypothesised that their reliable improvement ratings should converge. The brevity of the SDQ Emotion may, however, reduce its sensitivity to change, compared with the RCADS. A second hypothesis therefore states that if there is a lack of convergence, this is likely to be driven by the inferior sensitivity to change of the SDQ Emotion.

7.2.4 Measures of Functioning

Strengths and Difficulties Questionnaire (SDQ) – Impact Supplement

The SDQ Impact Supplement includes five optional questions assessing the impact of a child's difficulties on their daily lives (R. Goodman, 1999). If young people indicate difficulties in any of the four problem areas assessed by the general SDQ they can then answer these additional questions about the duration and degree of distress caused by these difficulties, and their impact on home life, friendships, classroom learning and leisure activities (four items rated on a 4-point scale from 0 (*not at all*), to 3 (*a great amount*). The items on distress and impact are summed to compute a total score ranging from 0 to 10 (for this, the response scale is converted to a 3-point scale with the original response categories 0 and 1 being recoded as

0, 2 being recoded as 1, and 3 being recoded as 2). The measure's developers report good internal validity, with a Cronbach's Alpha of 0.81 (R. Goodman, 2001).

Child Outcome Rating Scale (CORS) and Outcome Rating Scale (ORS) – Total Score

The Child Outcome Rating Scale (CORS) and the Outcome Rating Scale (ORS) are four-item self-report measures of clinical progress across four domains of functioning: symptom-related distress, interpersonal wellbeing, social role, and overall wellbeing (Miller et al., 2003). Responses to each question are recorded as markings on a 10-cm visual analogue scale, with the right-hand side of the scale representing high functioning and the left-hand side representing low functioning. Scoring is done by measuring the length between the starting point of the visual analogue scale and the marker, and by converting the measurement from centimetres into score points (e.g., ten is the highest score, given the line's 10-cm length). A total functioning score is computed by summing the four subscale scores. The CORS were designed for use with children aged 6-12 years; the ORS were designed for use with young people aged 13 and older. The tools are identical in format and scoring, but the CORS uses more child friendly language than the ORS, as well as smiley/frowny faces instead of descriptions of high and low functioning (Duncan et al., 2003). While the C/ORS has not yet been widely validated, the few existing studies have reported good internal consistency (between $\alpha = 0.81$ and 0.97; Bringhurst et al., 2006; Campbell & Hemsley, 2009; Casey et al., 2019; Miller et al., 2003), whereas findings for test–retest reliability have been mixed ($r = .66$ to 0.81; Bringhurst et al., 2006; Miller et al., 2003).

Given the similarity of the CORS and ORS, they have been combined into a composite score for the purpose of this study. Composite C/ORS scores and reliable change indicators were computed by drawing on all available data, while prioritising age-appropriate versions wherever both had been completed.[5]

Hypotheses About the Equivalence of the two Functioning Measures

Both the SDQ Impact and the C/ORS are brief measures of psychosocial functioning. They are of comparable length and cover functioning in the family, peer, and school context. Both also inquire about global notions of distress or wellbeing. But while the SDQ Impact probes into functional impairment caused by the mental health problems previously identified through the SDQ, the C/ORS asks young people *how well* they were generally doing. It has been suggested that functioning measures probing about disorder-specific impairment may be more sensitive to change than generic measures (Patrick & Deyo, 1989). On this basis, it can

---

[5] Of the 9,904 eligible cases, only a fraction had paired data on the C/ORS or ORS. Of these, a considerable share of young people filled in the ORS for adolescents aged 13 or older, despite being 12 years old at assessment, and inversely, a number of youth aged 13 or older completed the CORS for younger children. Within the eligible sample, incorrect versions were filled in by 765 young people at first assessment and by 451 at follow-up (some of these also filled in the correct form). As explained above, the two versions differ only in terms of their readability and child-friendliness: The CORS requires a reading level of third grade, while the ORS requires a reading level of eighth grade (Duncan et al., 2006). Discarding data collected on age-mismatched tools would reduce the sample from 458 to 282, leading to a considerable loss of information.

be hypothesised that both measures should broadly converge in their assessments of reliable improvement, but that the problem specific SDQ Impact may indicate higher levels of change than the generic C/ORS.

7.2.5 Idiographic Measure

Goals Based Outcomes (GBO)

The GBO tool (Law, 2006) is an idiographic outcome measure, designed primarily with clinical utility in mind (Wolpert et al., 2015). Young people, ideally jointly with their clinician, parents and/or teachers define a number of goals at the beginning of treatment, which they would like to work towards (Law & Jacob, 2015). While more goals can be defined, it is recommended that the top three goals are used for outcome reporting. Progress is then rated periodically on a scale from 0 (*no progress*) to 10 (*goal has been reached*).

The reliable change index for the GBO has been defined as a movement by at least 2.45 along the goal progress scale, where progress scores are aggregated across the three goals (Edbrooke-Childs et al., 2015). However, as service users are free to define less than three goals, the incidence of missing values is high in the CYP IAPT data. An alternative ‘meaningful change’ index has been suggested that uses all available GBO data, but does not require complete measurements on all three goals: ‘Meaningful change’ is defined as an improvement by at least 3 scale points on any completed goal, without equivalent deterioration on any other available goal (Jacob, 2019). This is the approach chosen for this study. Data on the reliability of the GBO is not currently available (Jacob et al., 2018).

Hypotheses About the Sensitivity to Change of the GBO

As mentioned above, idiographic outcome measures have tended to indicate higher levels of change than nomothetic measures, in a number of previous studies (Ashworth et al., 2004; Edbrooke-Childs et al., 2015; Godfrey et al., 2019). They typically focus on the most important target complaints, presenting problems, or treatment goals identified by service users. Since treatment should ideally focus on alleviating these issues, these are also likely to be the areas in which change is most likely or tangible (Lacasse et al., 1999). As such, it can be hypothesised that progress towards self-defined goals, as tracked by the GBO in this study, will exceed change measured by the four nomothetic measures of symptoms and functioning.

7.2.6 Statistical Analysis

Assessing Reliable Change

This study assessed the convergence between change assessed on different measures and in different outcome domains at an individual level. The criterion used to determine the salience of individual-level change was the reliable change index, as clinically significant change could only have been determined for the symptom measures, for which clinical norms were available (Jacobson & Truax, 1991). As described in Chapter 1, Section 1.7.2, the reliable change index determines the amount of change required to demarcate improvement beyond fluctuations that can be attributed to measurement error (Jacobson et al., 1999; Jacobson & Truax, 1991). The Reliable Change Index (RCI) is calculated by dividing the difference between scores at first (Time 1 or T1) and second (Time 2 or T2) measurement by the standard error of the difference between the two measurements:

$$RC = \frac{x_1 - x_2}{S_{\text{diff}}}$$

where x_1 and x_2 are an individual's scores at T1 and T2, and S_{diff} is the standard error of the difference between these scores. S_{diff} can be calculated from the standard error of measurement (S_E) according to the following formula:

$$S_{diff} = \sqrt{2}(S_E)^2$$

$$S_E = S_1\sqrt{1} - r_1$$

where S_1 is the standard deviation of the score measured at T1 and, r_1 is the reliability of the measure, defined here as Cronbach's Alpha, at T1. A magnitude of change exceeding 1.96 times the S_{diff} is unlikely to be due to measurement error alone in more than 5% of cases.

Rather than drawing on published standard deviations and reliability coefficients, RCIs were computed based on the psychometric properties and score distributions identified in the study sample. For each of the four nomothetic measures, the RCI was computed based on the standard deviation of the mean T1 score and the measure's internal consistency at T1, in the sub-sample that contributed paired data on the relevant measure either to the comparison of change across measures, or to the comparison of change across domains, or both. For example, 1053 cases contributed paired data on the SDQ Emotion, either to the within-domain comparison or to the cross-domain comparison. The RCI for the SDQ Emotion was computed based on this sample and the same procedure was followed for the three other nomothetic measures.

The SDQ Emotion and SDQ Impact demonstrated questionable internal consistency (α = 0.63 to 0.64, respectively). Internal consistency was good on the C/ORS (α = 0.87) and excellent on the RCADS (α = 0.95). The resulting Reliable Change Indices are displayed in Table 7.2. To be considered as indicating reliable change, individuals had to demonstrate a pre-post difference in scores that exceeded RCI thresholds of 15.39 for the RCADS, 3.49 for the SDQ Emotions, 4.02 on the SDQ Impact Scale, and 8.33 on the C/ORS. As explained above, the meaningful change threshold for the GBO was defined as 3. For the sake of brevity, the short-hand term “improved” will be used to describe both reliable improvement and meaningful improvement.

Table 7.2. Parameters Used to Determine the RCI for Each Measure

Measure	*N*	M_{T1} (*SD*)	$M_{T1} – M_{T2}$ (*SD*)	Alpha	RCI
SDQ Emotion	1053	7.37 (2.08)	-1.49 (2.62)	0.63	3.49
RCADS	960	73.37 (24.7)	-18.27 (27.4)	0.95	15.39
SDQ Impact	429	4.71 (2.43)	-1.79 (2.92)	0.64	4.02
C/ORS	138	19.35 (8.20)	6.41 (10.25)	0.87	8.33

Assessing the Convergence of Improvement Between Measures

In response to the first and second research question, change levels were compared between the SDQ Emotion and the RCADS within the symptom domain, and between the SDQ Impact and the C/ORS within the functioning domain. Across measures, few cases showed reliable deterioration, with percentages ranging from 1.7% on the SDQ Emotion to 7.7% on the RCADS, see Table E.2 in Appendix E). Analysing this category separately led to very small cell sample sizes, which frequently fell short of the minimum reportable cell sample size (Office for National Statistics, 2006). Therefore, cross tables of dichotomised reliable change ratings were computed instead, distinguishing only between reliable improvement versus no reliable improvement (including deterioration). McNemar’s test of correlated proportions (McNemar, 1947) was computed to assess the likelihood of no agreement between reliable improvement classifications achieved by both measures within a domain. In addition, Cohen’s Kappa for chance-corrected agreement was computed to estimate the level of agreement between two measures (J. Cohen, 1960). In line with recommendations (Kirkwood & Sterne, 2003), kappa results were interpreted according to the guidelines proposed by Landis and Koch (1977). Kappa values were interpreted as follows: $\kappa \leq 0$ poor agreement, $\kappa = 0.01–0.20$ slight agreement, $\kappa = 0.21–0.40$ fair agreement, $\kappa = 0.41–0.60$ moderate agreement, $\kappa = 0.61–0.80$ substantial agreement, and $\kappa = 0.81–1.00$ almost perfect agreement (Landis & Koch, 1977, p. 165). It has been suggested, however, that any kappa below 0.60 suggests inadequate agreement when assessing inter-rater reliability (McHugh, 2012).

Assessing the Convergence of Improvement Between Domains

In response to the third research question, the convergence of reliable improvement ratings was assessed across domains, in a subsample that had paired data on at least one symptom measure, one functioning measure, and on the GBO. Levels of reliable change were compared between each of these domains in a series of paired comparisons, as well as across all three domains. Composite reliable change indices were computed for the symptom and functioning domains by defining as reliably improved those who showed reliable change on at least one of the two measures within each domain, and no deterioration on the other. Table E.1 in Appendix E provides a break-down of the extent to which each measure contributed to these composite reliable change indices. McNemar's test and Cohen's Kappa were computed for paired comparisons, and Fleiss kappa was computed to assess agreement across all three domains (Fleiss, 1981).

7.3 Results

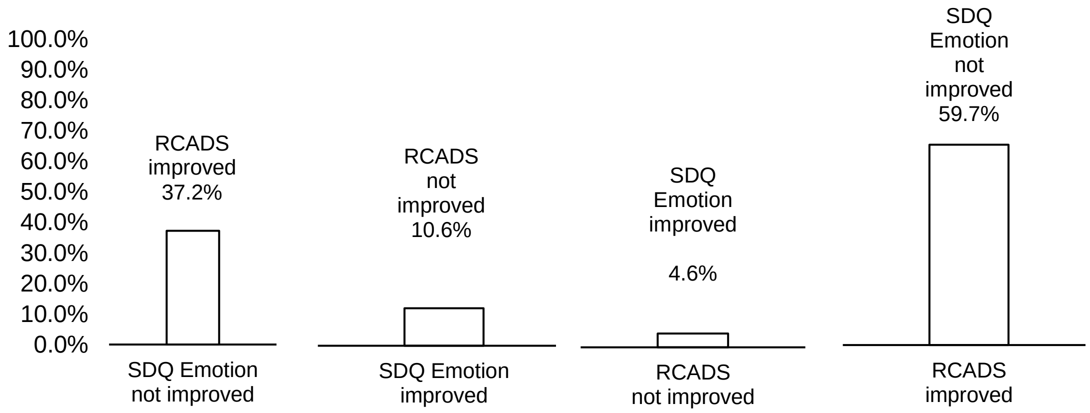
7.3.1 Comparing Change Within the Symptom and Functioning Domains

Convergence of Improvement Ratings Between Symptom Measures

A subsample of 943 young people had paired data on both the RCADS and the SDQ Emotion (see Table 7.3). Improvement was considerably higher on the RCADS (48.7%) than on the SDQ Emotion (22.0%, see Table E.2, Appendix E). Of all cases considered, 49.0% did not improve on any of the two measures, 2.3% improved solely on the SDQ Emotion, 29.1% improved solely on the RCADS, and 19.6% improved on both symptom measures. Reliable change ratings were thus discrepant in close to one third of cases (31.4%).

There was some divergence in the improvement ratings produced by each measure (see Figure 7.2). Of all cases who did not improve on the SDQ Emotion, more than one third (37.3%) did improve on the RCADS. Of the cases who improved on the RCADS, 59.7% failed to improve on the SDQ Emotion. McNemar's test of correlated proportions showed a significant difference in improvement ratings ($\chi^2 = 214.5$, $p < .001$), and Cohen's kappa indicated fair agreement ($\kappa = 0.36$, $p < .001$) between the two measures.

Figure 7.2. Disagreement Between the Symptom Measures



Note. *N* = 943.

Table 7.3. Disagreement Between Measures and Domains (Bivariate Comparisons)

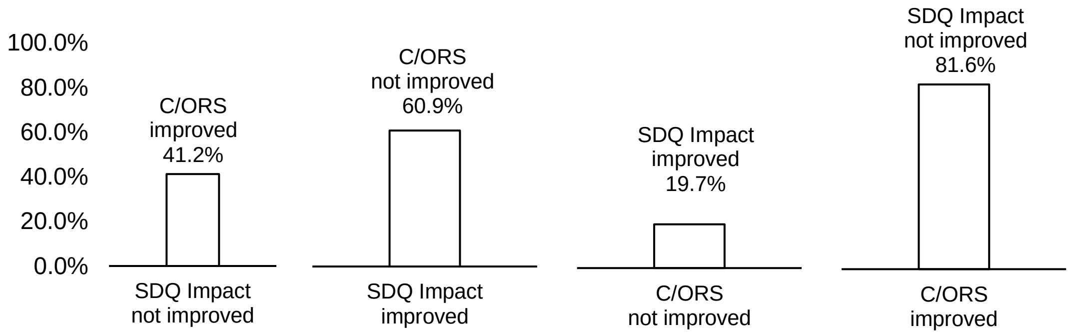
First measure/ domain		Second measure / domain		
Within the symptom domain				
		RCADS		
SDQ Emotion		Not improved *n* (%)	Improved *n* (%)	Total *n* (%)
	Not improved	462 (49.0)	274 (29.1)	736 (78.0)
	Improved	22 (2.3)	185 (19.6)	207 (22.0)
	Total	484 (52.3)	459 (48.7)	943 (100)
Within the functioning domain				
			C/ORS	
SDQ Impact		Not improved *n* (%)	Improved *n* (%)	Total *n* (%)
	Not improved	57 (46.5)	40 (33.3)	97 (80.8
	Improved	14 (11.6)	9 (7.5)	23 (19.2)
	Total	71 (58.1)	49 (40.8)	120 (100)
Across the symptom, functioning, and goal progress domains (paired comparisons)				
			Functioning	
Internalizing Symptoms		Not improved *n* (%)	Improved *n* (%)	Total *n* (%)
	Not improved	190 (55.3)	24 (6.5)	214 (58.0)
	Improved	93 (27.6)	62 (17.0)	155 (42.0)
	Total	283 (76.7)	86 (23.3)	369 (100)
			Goal progress	
Internalizing Symptoms		Not improved *n* (%)	Improved *n* (%)	Total *n* (%)
	Not improved	94 (25.5)	120 (32.5)	214 (58.0)
	Improved	30 (8.1)	125 (33.9)	155 (42.0)
	Total	124 (33.6)	245 (66.4)	369 (100)
			Goal progress	
Functioning		Not improved *n* (%)	Improved *n* (%)	Total *n* (%)
	Not improved	107 (29.0)	176 (47.7)	283 (76.7)
	Improved	17 (4.6)	69 (18.7)	86 (23.3)
	Total	124 (33.6)	245 (66.4)	369 (100)

Convergence of Improvement Ratings Between Functioning Measures

A subsample of 120 cases had paired functioning data on both the SDQ Impact and the C/ORS. Improvement was considerably higher on the C/ORS (40.8%) than on the SDQ Impact (19.2%, see Table E.2, Appendix E). Of all cases, 47.5% did not improve on any of the two measures, 11.6% improved solely on the SDQ Impact, 33.3% improved solely on the C/ORS, and 7.5% improved on both functioning measures (see Table 7.3). Reliable change ratings were thus discrepant in 44.9% of cases.

Of the cases that improved on the SDQ Impact, 60.9% did not improve on the C/ORS (see Figure 7.3, below). In turn, of all cases not improving on the SDQ Impact, over a third (41.2%) still improved on the C/ORS. Of all cases improving on the C/ORS, the large majority (81.6%) showed no improvement on the SDQ Impact. Of those showing no improvement on the C/ORS, 19.7% did improve on the SDQ Impact. McNemar's test of correlated proportions ($\chi^2 = 12.52$, $p < .001$) indicated statistically significant disagreement. Cohen's kappa (κ = -0.01, $p = .573$) was not statistically significant, likely due to the small sample size.

Figure 7.3. Disagreement Between the Functioning Measures



Note. *N* = 120.

7.3.2 Comparing Change Across Outcome Domains

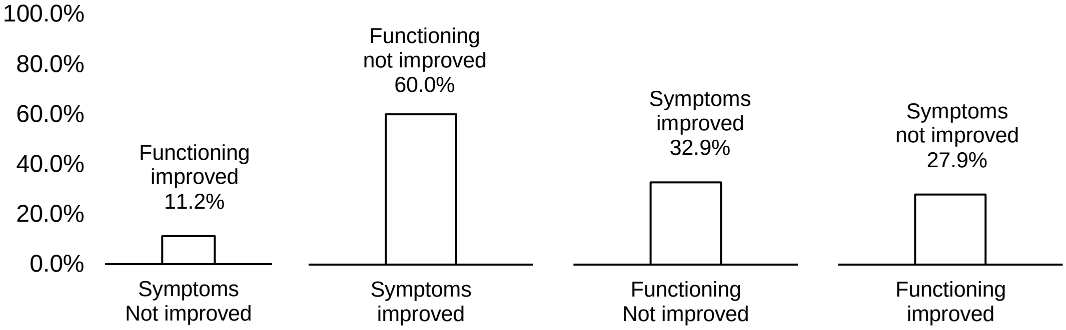
A subsample of 369 cases had paired data on at least one measure in the symptom domain, one measure in the functioning domain, and on the GBO measure of goal progress. Of these, 66.4% showed improved their goal progress, 41.0% improved their symptoms, and 23.3% improved their functioning (see Appendix E, Table E.2). Paired comparisons were conducted between the changes observed in different domains as well as across all three domains.

Comparing Change Between the Symptom and Functioning Domains

The first paired comparison focussed on the domains of symptoms and functioning. Of all cases considered, 55.3% improved in none of the two domains, 27.6% improved only their symptoms, 8.9% improved only their functioning, and 21.4% improved in both domains (see Table 7.3).

Of all cases that improved in the symptom domain, 60.0% did not improve their functioning (see Figure 7.4). Of all cases that failed to improve their symptoms, 11.2% did, however, improve their functioning. Of those improving their functioning, 27.9% did not improve their symptoms, while of those not improving their functioning, 32.9% improved their symptoms. McNemar's test showed a significant difference in improvement ratings between the symptom and functioning domains ($\chi^2 = 40.69$, $p < .001$). Cohen's kappa indicated fair agreement (κ = 0.31, $p < .001$).

Figure 7.4. Disagreement Between the Symptom and Functioning Domains



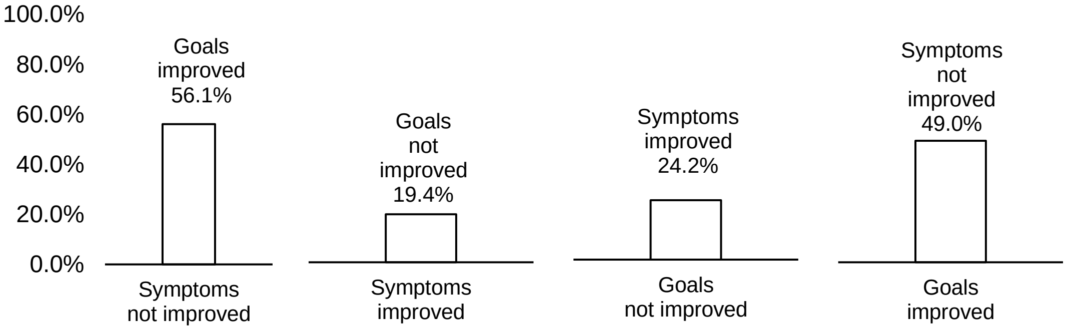
Note. N = 369.

Comparing Change Between the Symptom and Goal Progress Domains

Next, change was compared between the domains of symptoms and goal progress. Of all cases considered, 25.5% improved in none of the two domains, 32.5% exclusively improved their goal progress, 8.1% improved exclusively their symptoms, and 33.9% improved in both domains (see Table 7.3).

Of the cases that failed to improve their symptoms, 56.1% still improved their goal progress (see Figure 7.5, below). Of all cases that did improve their symptoms, 19.4% did not improve their goal progress. Of those not improving on the GBO, 24.2% still improved their symptoms, and of those who did improve their goal progress, 49.0% showed no improvement in symptoms. McNemar’s test of correlated proportions showed a significant difference in improvement ratings between the symptom and goal progress domains ($\chi^2 = 54.00$, $p < .001$) while Cohen’s kappa indicated fair agreement ($\kappa = 0.23$, $p < .001$).

Figure 7.5. Disagreement Between the Symptom and Goal Progress Domains



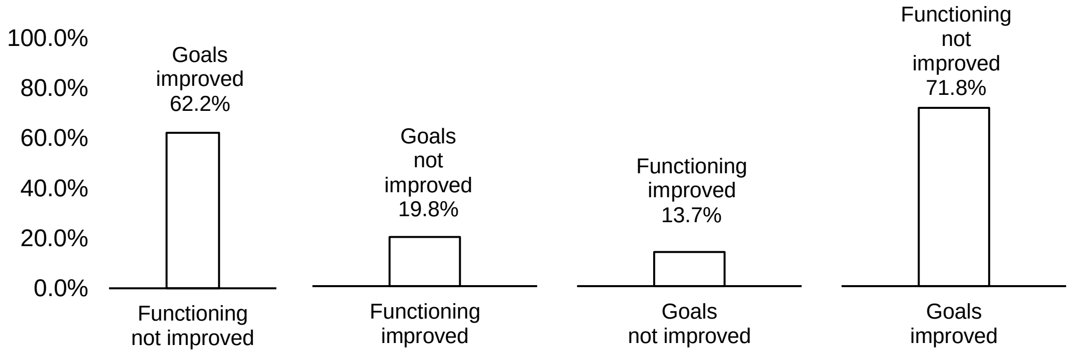
Note. *N* = 369.

Comparing Change Between the Functioning and Goal Progress Domains

Next, change was compared between the domains of functioning and goal progress. Of all cases considered, 29.0% improved in none of the two domains, 47.75% exclusively improved their goal progress, 4.6% improved exclusively their functioning, and 18.7% improved in both domains (see Table 7.3).

Of all cases that showed no improvement in functioning, 62.2% still improved their goal progress (see Figure 7.6, below). Of those who did improve their functioning, 19.8% showed no improvement in goal progress. Inversely, of all cases improving their goal progress, 71.8% failed improve their functioning. Of those not improving their goal progress, 13.7% still improved their functioning. McNemar's test of correlated proportions showed a significant difference in improvement ratings between the functioning and goal progress domains (χ^2 = 130.99 *p* < .001). Cohen's kappa (κ = 0.11, *p* = .001) indicated slight agreement.

Figure 7.6. Disagreement Between the Functioning and Goal Progress Domains



Note. N = 369

Comparing Change Between the Symptom, Functioning and Goal Progress Domains

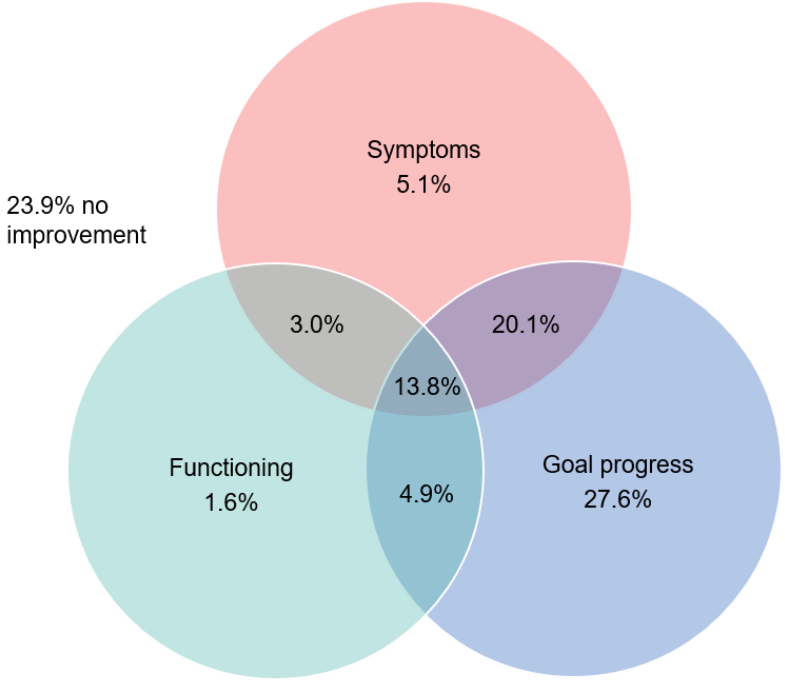
Lastly, change was compared across all three outcome domains. Of all cases considered, 23.8% did not improvement in any domain, 27,6% improved exclusively their goal progress, 20.1% improved their symptoms and goal progress but not their functioning, and 13.8% improved across all three domains (with 2-5% of cases showing other combinations of change, see Table 7.4 and Figure 7.7).

Table 7.4. Disagreement Between Symptoms, Functioning, and Goal Progress

		Goal progress	
Symptoms	Functioning	Not improved	Improved
Not improved	Not improved	88 (23.8%)	102 (27.6)
	Improved	6 (1.6%)	18 (4.9%)
Improved	Not improved	19 (5.1%)	74 (20.1%)
	Improved	11 (3.0%)	51 (13.8)

Note. N = 369 (100%).

Figure 7.7. Venn Diagram of Improvement Across all Three Domains



Note. N = 369. Figures within the circles indicate percentages of reliable or meaningful improvement. The circles are not proportionate in size to the level of improvement observed in each domain.

Overall, change ratings across the three domains were discrepant for 62.3% of cases. Fleiss' kappa showed only slight agreement across reliable improvement ratings across the three domains ($\kappa_f = 0.16$, $p < .001$).

7.4 Discussion

7.4.1 Summary of Findings

This study assessed levels of change observed by five commonly used self-report measures of symptoms, functioning, and progress towards self-defined goals, as well as the convergence in reliable improvement within and between these three domains. Rates of reliable improvement were highest in goal progress, followed by symptoms, and functioning. Ratings of reliable change diverged considerably between measures within the same domain (e.g., symptoms or functioning), and across the three outcome domains of symptoms, functioning and progress towards goals. Reliable change ratings across the three outcome domains were discrepant for two thirds of the cases considered.

Lack of Convergence Between the RCADS and the SDQ

Within the symptom domain, the RCADS showed considerably higher levels of reliable change than the SDQ Emotion (48.7% and 22.0%, respectively). More than half of young people who reliably improved on the former failed to improve on the latter, and the two measures produced discrepant findings for close to one third of cases. Given these important discrepancies, our initial hypothesis whereby these two measures would converge, can be refuted. Our second hypothesis, however, whereby discrepancies might be driven by a lower sensitivity to change of the SDQ Emotion appears to hold. This might be due to the brevity of the scale, but it is also driven by its inferior internal consistency, which implies that a higher magnitude of change is needed to be considered reliable. Clinicians have highlighted the broad focus of the SDQ as a disadvantage in clinical practice, and according to a recent survey, only 44% consider the tool helpful, compared with 74% for the RCADS (Costa da Silva & Wolpert, 2018; Wolpert, Cheng, et al., 2015; Wolpert & Rutter, 2018). At the same time, poor convergence in depression symptom measures at the individual level has also been observed for adults (Zimmerman et al., 2012).

Lack of Convergence Between the SDQ Impact and the C/ORS

Within the domain of functioning, the C/ORS showed twice as much reliable change as the SDQ Impact (40.8% and 19.2%, respectively). More than 80% of cases who reliably improved on the C/ORS failed to show reliable change on the SDQ Impact. Overall, reliable change ratings were discrepant in 44.9% of cases. Our hypothesis that these two measures should yield equivalent ratings of reliable improvement in functioning must be refuted. Our second hypothesis, whereby the problem specific SDQ Impact scale would be more sensitive to change than the generic C/ORS must also be refuted. It is important to note that the C/ORS had a higher internal consistency than the SDQ Impact, which influenced the computation of the reliable change index, and meant that a lower magnitude of change was required to be considered reliable on the C/ORS than on the SDQ Impact.

Lack of Convergence Between the Symptom and Functioning Domains

Reliable improvement was considerably higher in symptoms (41.0%) than in functioning (23.3%). Close to two thirds of young people who reliably improved in the symptom domain failed to show reliable improvement in functioning. This finding is consistent with previous research showing low levels of convergence (Brookman-Frazee et al., 2006) and less evidence of treatment effectiveness in the functioning domain than in the symptom domain (Becker et al., 2011). There are several possible explanations for the observed lack of convergence. First, it has been argued that associations between depression symptoms and functioning measures may vary based on the types of symptoms assessed, as some explain a larger variation in functional impairment than others (Fried & Nesse, 2014; Tweed, 1993). Second, several studies in adult populations have found changes in social and global functioning to lag behind changes in depressive symptoms (Bothwell & Weissman, 1977; Hirschfeld et al., 2002, 2000; McKnight & Kashdan, 2009; Scott et al., 2000).

These findings suggest that symptom change is a poor proxy for change in functional impairment, and that conclusions about the latter may be inflated if inferred from symptom-focussed measurement. Symptom change may be an early sign of treatment response, while functioning may be slower to manifest but could indicate deeper or more wide-reaching treatment effects (McKnight & Kashdan, 2009). Another possible reason for the lower rates of reliable improvement observed in functioning, however, may be that functioning measures for children and young people have received less attention from psychometricians, compared with symptom measures. They tend to have weaker psychometric properties, and possibly weaker sensitivity to change (Becker et al., 2011; Canino et al., 2013; McKnight & Kashdan, 2009). In both cases, the present findings indicate that symptom and functioning measures currently recommended for routine tracking in the UK cannot be considered as providing equivalent assessments of change, but rather provide complementary evidence.

Lack of Convergence Between Symptoms, Functioning, and Goal Progress

Convergence between change in symptoms and functioning on the one hand, and progress towards self-defined goals on the other hand, was poor. Only 13.8 of youth with data in all three domains also showed reliable improvement in all three. In line with our initial hypothesis, reliable improvement in goal progress was considerably higher than in symptoms and functioning. More than half of young people who did not show reliable improvement in symptoms or functioning still improved their goal progress (56.1% and 62.2%, respectively). This was consistent with two previous studies in child mental health that compared change on nomothetic and idiographic measures, using parent-rated GBO data (Edbrooke-Childs et al., 2015), and data from the self-reported Target Complaints (Karpenko & Owens, 2013). It was also consistent with a number of studies evidencing the superior sensitivity to change of idiographic measures in adult mental health (Ashworth et al., 2005; Elliott et al., 2016; Hurn et al., 2006; King et al., 1999; Paterson, 1996).

Previous studies have assessed convergent validity at the group level by examining correlations between the GBO and nomothetic measures. While parent-reported GBO have been shown to correlate moderately with parent-reported SDQ total difficulty scores (r = 0.3 – 0.4) and clinician-rated functioning (r = 0.4 – 0.5), no significant correlation has yet been found for child-reported measures (Edbrooke-Childs et al., 2015; Wolpert, Ford, et al., 2012). The GBO may thus capture changes that are uniquely different from those assessed by nomothetic measures of symptoms and functioning. It has been suggested that idiographic measures may show a higher sensitivity to change, because they are tailored more specifically to capturing change in the issues most salient to service users, and which – ideally – treatment would focus on (Law, 2006; Law & Jacob, 2015; Rockwood et al., 1997). At the same time, limited psychometric testing of the GBO has been done to date and more research is required to confirm that the currently used threshold for meaningful change is appropriate (Jacob et al., 2018). More generally, further research examining the psychometric properties of goals and other idiographic measures is needed. The high levels of change measured by the GBO might also stem from services defining goals that are ‘too easy’ to achieve. To explore this further, goals will also need to be analysed with explicit reference to their content (Kiresuk & Sherman, 1968; Rockwood et al., 1997).

7.4.2 Clinical Implications

The findings from this study showcase that the types of outcomes services choose to monitor, and the instruments they select for this purpose, influence judgements about treatment effectiveness. The equivalence of instruments designed to capture a similar concept cannot be assumed, even when a standardised change metric like the Reliable Change Index is computed to enable comparisons. Differences in scale reliability and sensitivity to change, as well as differential item content mean that apples may be compared with oranges, which can have important implications where discrepancies are taken at face-value, and interpreted as true differences in clinician or service performance.

This study's findings corroborate existing evidence whereby symptom reduction does not constitute an adequate proxy for treatment effectiveness in other domains, as it may lead to either over- or under-estimating the overall change achieved. Symptom change may fail to translate into sizeable improvements in functioning; while youth failing to improve their symptom scores may still show progress towards personal treatment goals. This provides another argument in favour of establishing more multidimensional approaches to outcome measurement, including in routine practice. At the same time, multidimensional measurement is complex and poses challenges in terms of how to interpret and reconcile discrepant results. Aggreging metrics such as reliable improvement across instruments and outcome domains, as considered by the NHS as a future approach, simplifies the interpretation and synthesis of results, while enabling the detection of change across various domains. At the same time, however, nuanced differences in the types of outcome domains that do or do not show reliable change are masked, and valuable information may be lost.

Standards for outcome measurement in clinical practice are needed in order to promote harmonisation and enable fair comparisons and benchmarking. Multidimensional core outcome sets are crucial in this effort. As mentioned in previous sections, one set developed specifically for tracking outcomes for adolescent depression in clinical practice is currently being finalised (Krause et al., 2020). This study suggests that such efforts should not only focus on nomothetic measures, but consider including idiographic tools too. The high rate of meaningful improvement measured by the GBO in this study suggests that such tools can capture improvements in areas that matter to young people, but that are missed by nomothetic measures of symptoms and functioning. The latter are more appropriate for benchmarking and comparison, but idiographic tools may be particularly relevant to clinical decision-making and progress tracking with individual service users (Sales, 2017; Sales et al., 2014). Research suggests that goal-setting, as one approach to idiographic measurement, can improve treatment retention; and young people have associated it with improving self-awareness, problem-solving, and goal achievement (Cairns et al., 2019; Duong et al., 2016). Research in adult mental health indicates that the subjective perception of moving towards personal goals has a positive effect on subjective wellbeing (Brunstein, 1993; Koestner et al., 2002; Wiese, 2007). However, a caveat must be raised about aggregating change from idiographic measures when their actual item content is not fully understood. Aggregation can mask ambiguity in goal or problem content, and variation in the domains these pertain to.

7.4.3 Limitations

The above-mentioned findings should be considered in the context of a number of limitations. First, the CYP IAPT dataset analysed for this study has been described elsewhere as an example of naturalistic data that are *flawed*, *uncertain*, *proximate* and *sparse* (“FUPS”) (Wolpert & Rutter, 2018). The data is *flawed* in that there is a great amount of missing data, which cannot be assumed to be missing entirely at random. Young people who completed measures at assessment and follow-up may be systematically different from young people who provided incomplete data, in characteristics that have not been observed. They may be more motivated and engaged, or they (or their clinicians) may have filled in measures that felt particularly relevant, or where progress was most likely to be reflected. Future research into possible observed determinants of drop-out and non-response would add valuable insight and might enable imputation based on a missing at random assumption in the future. In this study, imputation was not possible due to the high share of missing paired data on each outcome measure and the lack of suitable auxiliary variables that could have been used to estimate missing data using multiple imputation modelling.

Second, the data is *uncertain* and *proximate* (Wolpert & Rutter, 2018) in that it remains uncertain whether observed differences in reliable change ratings are due to diverging psychometric properties, or because rates of change truly differ in the underlying concept of interest. They share the characteristics of ‘arbitrary metrics’ (Blanton & Jaccard, 2006), in that their scores and change metrics cannot be intuitively interpreted in terms of their real-life significance. It is also worth noting that the two assessment time points included in the CYP IAPT provide only a snapshot view of change achieved over the course of treatment. No information was available about the extent to which reliable change converged or diverged at a later stage, or about the temporal patterns with which change unfolded on each measure and within each domain. In addition, different measures were often completed at different time points, which may further add to the lack of convergence observed.

Third, for reasons related to the structure of the dataset, it was impossible to consider the qualitative content of goals defined by young people using the GBO. Goal content was collected through free text responses, and during data cleaning, it was discovered that these contained identifying information in a number of cases. To ensure the anonymity of the data, all goal content was removed. This is an important limitation. It meant that the GBO’s meaningful change metric had to be analysed “blindly” without reference to the actual goal content. A separate thematic mapping of self-defined goals collected from English child and adolescent mental health services found that these goals revealed themes that were not covered by commonly used nomothetic measures (Jacob, Edbrooke-Childs, Law, & Wolpert, 2017). They included improvements related to independence, confidence, self-reflection, communicating feelings, and understanding anger. This suggests that the GBO does indeed complement nomothetic symptom and functioning measures, by capturing distinctly different types of change.

This study constitutes a first step towards developing a broader research programme that should look into the overlap or distinctions between idiographic and nomothetic measures, both in terms of their item content, and in terms of how change is scored. It would be of great value for future research to consider the item content of personalised measures (e.g., of individual goal or problem descriptions) by (a) systematically mapping the outcome domains covered and how their content relates to that of established nomothetic measures; (b) exploring whether item content varies by developmental stage or other demographic characteristics; and (c) examining the extent to which the change metrics of self-defined items in specific outcome domains converge or diverge from change metrics of nomothetic measures capturing similar concepts.

Fourth, to be able to compare change across domains in a sufficiently large sample, composite reliable change indicators were computed that pulled available data from the four symptom and functioning measures. This approach resembled that which the NHS is planning to use when computing a general metric of meaningful change across different measures and domains. However, as suggested above, differential sensitivity to change of individual measures contributing to this metric must be considered. For example, in this study, reliable change ratings that informed the composite metric for functioning were pulled exclusively from the SDQ Impact for 83.7% of the cases considered (see Annex E, Table E.1). Given that the SDQ Impact indicated only about half the amount of reliable change as the C/ORS in the within-domain comparison, the cross-domain comparisons might have shown slightly less discrepant results (due to higher rates of improvement in the functioning domain), had a larger share of the data been pulled from a functioning measure with a higher sensitivity to change.

Finally, the study reflected some of the limitations in encouraging multidimensional outcome measurement in clinical practice without requiring a certain degree of harmonisation. As part of the CYP IAPT initiatives, services were encouraged to select from 13 different measures, which covered a broad range of outcome domains including measures of service experience and satisfaction[6], wellbeing[7], family functioning[8], and parental self-efficacy[9]. However, these measures were rarely completed and could not be considered for the analysis presented in this chapter. This limited the possibility of analysing convergence in reliable change ratings across a broader range of domains.

---

[6] The Experience of Service Questionnaire (ESQ; A. Brown, Ford, Deighton, & Wolpert, 2014), Session Rating Scale (SRS; Duncan et al., 2003), and Session Feedback Questionnaire (Child Outcomes Research Consortium, 2012).

[7] The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) long or short form (Haver et al., 2015; Tennant et al., 2007).

[8] The Systemic Clinical Outcome and Routine Evaluation-15 (SCORE-15; Stratton et al., 2010).

[9] Brief Parental Self Efficacy Scale (BPSES; Woolgar et al., 2013)

7.4.4 Future Research

An important area for future research is to further explore the differences in sensitivity to change between the measures examined here, as well as between other similar measures. Such inquiries should consider sensitivity to change and convergence at an item or sub-scale level (Fried, 2016; Fried & Nesse, 2014). In particular, the validity, reliability, and sensitivity to change of youth-reported functioning measures should be examined further, to help understand whether brief scales such as the SDQ Impact and C/ORS provide the best possible avenue for tracking change in functioning in clinical practice, or whether more granular measures would be preferable. In addition, it is currently not well understood why change in functioning is more difficult to evidence than change in the symptom domain (Becker et al., 2011; Canino et al., 1999, 2013). Future research should focus on comparing change trajectories for symptoms and functioning; and examine whether these outcomes tend to converge at certain time points following treatment, for specific subgroups, or based on initial levels of functioning (McKnight & Kashdan, 2009; Napoleone et al., 2019).

Personalised measures may show a higher sensitivity to change because they are designed to track the issues most salient to service users, which should be central themes during therapy (Law, 2006; Law & Jacob, 2015; Rockwood et al., 1997). However, high levels of change measured by the GBO and other idiographic tools may also stem from service users defining goals (or target problems) that are ‘too easy’ to achieve or address (Rockwood et al., 1997). More research is needed that examines sensitivity to change and convergence with other measures in relation to specific goal content, and to ascertain the psychometric properties of goals and other personalised measures. This should also seek to confirm whether the currently used threshold for meaningful change is appropriate (Jacob et al., 2018).

7.5 Conclusions

Routinely collected outcome data plays a growing role in decision-making about service organisation, allocation of funds, and policy priorities in child mental health in the UK (e.g., Garralda, 2009). In clinical practice there is an urgent need for outcome measures that provide reliable, valid, but also meaningful and personalised assessments of treatment effectiveness. If outcome measures designed to capture similar constructs cannot be used interchangeably and often yield discrepant results for the same individuals, measurement approaches may need to be streamlined to enable unbiased comparisons. This study further shows that aggregating change across outcome domains risks foregoing the benefits of multidimensional outcome measurement by masking differences in treatment impact on different domains. Making maximum use of all available data and exploring inconsistencies in terms of their meanings and drivers may enable more nuanced insights into what treatments work; for whom; and with regards to which outcome (De Los Reyes & Kazdin, 2008). While measuring outcomes across a mix of domains using a mix of nomothetic and idiographic goals is complex, harmonisation can help focus resources and interest on a set of recommended measures, which can then be calibrated thoroughly against one another, and for which reliability, validity, and sensitivity to change can be studied in detail (Kazdin, 1999b; Sechrest et al., 1996). Where practice-based evidence drives decisions about the allocation of care in the real-world, ambiguity is to be avoided. The stakes are high, and this will need to be reflected in the careful calibration of outcome measures and metrics for this purpose.

Chapter 8. General Discussion and Conclusions

8.1 What Did This Thesis Set out to Do?

Based on the growing influence of outcome data on the design and funding of mental health support for young people with depression, this doctoral thesis set out to explore what constitutes a ‘good outcome’ of treatment, through a mixed-methods approach involving five studies. These studies explored outcome concepts and priorities, and examined the extent to which judgements of good outcome converged when assessed using different measures and across different outcome domains.

- **Chapter 1** provided an overview of key concepts, discussed the existing empirical literature on outcome perceptions and priorities in child mental health, and introduced a number of relevant measurement issues.
- **Chapter 2** set out the research design and methodological approach underpinning the doctoral thesis.
- The **Outcome Taxonomy Review** (Chapter 3) involved a narrative review of existing outcome taxonomies in child mental health and their synthesis into a new, more comprehensive taxonomy.
- The systematic **Outcome Measurement Review** (Chapter 4) aimed to establish the types of outcomes measured in recent treatment efficacy and effectiveness studies for adolescent depression.
- The study of **Post-Treatment Outcome Perspectives** (Chapter 5) presented qualitative content analysis of semi-structured interviews conducted with youth, parents, and clinicians following treatment. It had three aims: First, to identify the types of outcomes discussed by young people, parents, and clinicians, when reflecting on the changes achieved over the course of therapy; second, to compare the salience of different outcome domains and categories in stakeholder interviews with the frequency of their measurement in the treatment efficacy and effectiveness studies established by the Outcome Measurement Review; and third, to explore differences in the types of outcomes discussed, and their relative salience between participant groups and treatment modalities.
- The study of **Viewpoints on Outcome Priorities** (Chapter 6) used Q-methodology to explore outcome priorities amongst youth with experience of depression and clinicians, with special attention to heterogeneity in viewpoints both within and between these two groups.
- The **Comparing Change Across Outcome Domains** study (Chapter 7) assessed the extent to which judgements of treatment success converged or diverged for individual cases, when assessing reliable change across different measures and outcome domains.

8.2 Summary of Findings

8.2.1 Outcome Taxonomy Review

Table 8.1 displays a revised version of the initial taxonomy developed in the Outcome Taxonomy Review in Chapter 3, which incorporates the changes applied based on findings from the Outcome Measurement Review (Chapter 4) and the Post-Treatment Outcome Perspectives study (Chapter 5), as well as some of the insights from the Viewpoints on Outcome Priorities study (Chapter 6). The final taxonomy includes seven higher-level outcome domains and 32 more specific outcome categories.

The majority of outcome categories identified in the initial taxonomy in Chapter 3 were relevant and applicable to the outcomes measured in treatment efficacy and effectiveness studies for adolescent depression (Outcome Measurement Review), and to the change narratives of young people, parents, and clinicians following treatment for depression in the IMPACT trial (Post-Therapy Outcome Perspectives study). The outcome categories identified in Chapter 3 were exhaustive at classifying change within the domains of *symptoms*, *functioning*, *relationships*, *personal growth*, and *parental support and wellbeing*. New outcome categories were added to the domains of *self-management* and *wellbeing*.

The domains of *self-management* and *personal growth* had no equivalents in the conceptual framework proposed by Hoagwood and colleagues (1996). Outcomes within both domains were, however, frequently mentioned by young people, parents, and clinicians in post-treatment interviews, and were also covered by a number of recent treatment efficacy and effectiveness studies. Several outcomes classified within the *self-management* domain are typically the focus of CBT, such as restructuring cognitive processes, learning ways to cope, and behavioural activation (Kennard, Clarke, et al., 2009). The personal growth domain, on the other hand, includes changes more typically targeted by psychoanalytic psychotherapy, (Fonagy, 2010; Odhammar & Carlberg, 2015). Both domains are important and theoretically grounded components of this final taxonomy. However, they do not necessarily sit at the same level as change in symptoms, functioning, or wellbeing, which can be described as higher-level outcomes based on which health status is determined. Improvements in self-management, personal growth, and relationships *can* constitute ultimate goals of treatment; but they may also constitute *intermediate* outcomes, considered instrumental in bringing about those higher-level clinical outcomes. Whether they are interpreted as one or the other is likely to vary by stakeholder group, within groups based on personal preferences and needs, and depending on the theoretical framework through which treatment is approached.

Table 8.1. Revised Taxonomy of Treatment Outcomes for Adolescent Depression

Outcome domains	OTR	OMR	PTOP
Symptoms			
Depression symptoms	✓	✓	✓
Suicidality	✓	✓	✓
Self-harm	✓	✓	✓
Anxiety	✓	✓	✓
Anger and aggression	✓	✓	✓
Eating and weight	✓	✓	✓
Sleeping problems	✓	✓	✓
Substance use	✓	✓	✓
Other comorbid problems	✓	✓	✓
Self-management			
Behavioural and physical activation	—	✓	✓
Cognitive processes	—	✓	✓
Coping with mood, thoughts and feelings	✓	✓	✓
Self-efficacy and empowerment	✓	—	✓
Understanding feelings and thoughts	✓	—	✓
Functioning			
Academic functioning	✓	—	✓
Executive functioning	✓	✓	✓
Global functioning	✓	✓	✓
Social functioning	✓	✓	✓
Personal growth			
Assertiveness	✓	✓	✓
Autonomy and responsibility	✓	✓	✓
Processing personal history	✓		✓
Self-concept and Identity	✓	✓	✓
Attitudes towards self (e.g., self-esteem)	✓	✓	✓
Relationships			
Being able to talk about feelings and thoughts	✓		✓
Family functioning and relations	✓	✓	✓
Friendships and peer relationships	✓	✓	✓
Wellbeing			
Peace of mind	✓	—	✓
Hope and optimism	—	—	✓
Future orientations	—	—	✓
Physical health	✓	✓	—
Parental symptoms			
Parental wellbeing	✓	✓	✓
Parental support	✓	—	✓

Note. ✓ indicates that an outcome was identified in the relevant study; — indicates that the outcome was not identified. OTR stands for Outcome Taxonomy Review. OMR stands for Outcome Measurement Review. PTOP stands for Post-Therapy Outcome Perspectives study.

The relevance of several outcomes, which had been extracted from the adult-focussed Bern Inventory during the Outcome Taxonomy Review, could not be validated and have been removed from the final taxonomy. These included symptoms in relation to *sexuality*, *meaning and purpose*, and the *recognition and fulfilment of wishes*. However, the latter was arguably replaced with the category of *future orientations* based on the Post-Therapy Outcome

Perspectives study. The Outcome Measurement Review mapped a number of service-level outcomes, such as service user satisfaction, treatment adherence, and the use of other services. These were, however, not incorporated into this final taxonomy, which focuses on gains for young people and their families, rather than service performance.

In the Post-Therapy Outcome Perspectives study, young people and parents frequently emphasised the value of feeling heard, cared for, and listened to without being judged. These are aspects of a well-developed therapeutic alliance, and arguably represent facilitators of change, rather than changes within the young person or their lives. These were tentatively included as outcomes in the Viewpoints on Outcome Priorities study, to examine how highly young people valued these experiences as outcomes in their own right. However, when they had to choose between these and actual changes in symptoms, coping skills, relationships, personal growth, or wellbeing, young people did not prioritise them. While they constitute important facilitators, these aspects of therapeutic alliance have therefore been omitted from the final taxonomy, which focusses firmly on person-centred treatment *outcomes.*

A final point of note is the ambiguity of the wellbeing domain. Based on the revisions made following the Outcome Measurement Review and the study of Post-Therapy Outcome Perspectives (Chapters 4 and 5), the domain covers positive changes in mindset, outlook and physical health that go beyond specific symptoms of mental health difficulties. However, a clear concept of wellbeing outcomes relevant to youth depression has not consistently emerged over the course of this research, because of considerable conceptual overlap between features of wellbeing and depressive symptoms. A systematic review of wellbeing definitions in the child health literature concluded that “well-being is a complex, multi-faceted construct that has continued to elude researchers’ attempts to define and measure it” (Pollard & Lee, 2003, p. 60). The review authors suggested that wellbeing comprises five domains: (a) *physical health*; (b) *mental health, including anxiety, emotions and self-esteem*; (c) *cognition and feelings about academic performance*; (d) *social relationships*; and (d) *availability of socio-economic resources*. According to this definition, wellbeing spans several domains included *alongside it* in the outcome taxonomy developed here (i.e., symptoms, personal growth, relationships), which underscores the difficulty of demarcating it as a concept in its own right. At the same time, a study of the correlates of mental health difficulties and wellbeing in children suggests that these are “largely distinct”, which suggests that these two concepts should be considered separately (Patalay & Fitzsimons, 2016). Further research is needed to better understand whether wellbeing outcomes can be distinguished from the symptom domain in youth with depression, and how best to demarcate and define them.

8.2.2 Outcome Measurement Review

The Outcome Measurement Review indicated that outcomes were primarily measured with respect to symptom change in recent treatment efficacy and effectiveness studies for adolescent depression. With the exception of functioning, other outcome domains were rarely covered. This was in line with previous studies that had covered a broad range of ages and presenting problems, but only considered controlled study designs (Becker et al., 2014, 2011; Hoagwood et al., 2012; P. S. Jensen et al., 1996; Weisz et al., 2005). Multidimensional outcome measurement was more common in evaluations of psychotherapeutic treatments, and less common in studies assessing combined treatments or routine care. It was also more common in qualitative or single-case study designs, than in chart reviews or follow-ups to clinical trials. This may indicate a legacy of unidimensional outcome measurement, as follow-ups and chart reviews tended to report on data that had been collected years earlier. Indeed, there was an upward trend in the average number of domains covered by the most recently published studies (i.e., between 2013-2017). While more than half of all studies considered youth self-report, clinician-report was typically used for assessing primary outcomes, while youth report was considered for secondary outcomes. The reviewed studies provided very limited conceptual justification for outcome measurement, and little discussion of findings within a developmental or multidimensional framework.

8.2.3 Post-Therapy Outcome Perspectives

The Post-Therapy Outcome Perspectives study revealed outcome narratives that were inherently multidimensional. The most salient outcome across perspectives and modalities was an improvement in mood and affect, but close to half of participants also described changes in family functioning, coping and resilience, and academic and vocational functioning; or emphasised the value of the therapeutic process as a transformative experience in and of itself. The types of outcomes described beyond these core domains varied by stakeholder perspective and treatment modality. Parents and clinicians demonstrated a greater concern with academic and vocational functioning than young people did. In addition, parents were more likely to discuss changes in young people's health behaviours at home, as well as indications that young people became more able to envisage and pursue goals for the future. Therapists showed a particular concern with self-harm, as well as changes in identity, self-image and self-confidence, especially within brief psychoanalytic psychotherapy.

While core outcomes were discussed consistently across treatment arms, their relative importance (as well as that of other outcome themes) reflected the different treatment approaches, with a particular focus on cognition and behaviour in the CBT arm, on healthy life-styles in the BPI arm, and on changes in the sense of self in the STPP arm. Of the different outcome categories identified as salient to youth, parents, and therapists, only changes in mood and affect (i.e., depressive symptoms) have been routinely reported in recent treatment effectiveness studies, according to the Outcome Measurement Review. This suggests that a shift towards considering multiple dimensions and perspectives is urgently needed, to meet stakeholder needs and enable more nuanced insight into change processes.

8.2.4 Viewpoints on Outcome Priorities

The Q-methodological study suggested that there are different views on what outcomes are *most* important to young people and mental health professionals. Four viewpoints on how to prioritise outcomes were identified in each of the two groups, which converged to form four cross-cutting super viewpoints. Two of these were shared by young people and professionals, and focussed on (a) symptom change, safety, and wellbeing; and (b) promoting resilience through improved coping skills, self-management, and understanding of emotions and experiences; respectively. A third super viewpoint focussed on promoting resilience systemically through family support, and was identified amongst mental health professionals only. A fourth super viewpoint, focussed on the interference of depression with daily functioning, was identified in a small group of young people with a high reported burden from depression and comorbid difficulties.

In line with findings from the Post-Therapy Outcome Perspectives study, changes related to mood and affect (i.e., core depressive symptoms) were endorsed as important by seven out of the eight viewpoints identified in the two groups. While youth viewpoints consistently emphasised the importance of *Feeling happier and enjoying things more,* professionals consistently prioritised youth e*ngaging less in behaviour that could be harmful*, thus showing a focus on risk management that was also observed in the Post-Therapy Outcome Perspectives study. Coping and self-management was another outcome that emerged as important across both studies. In contrast, academic and vocational functioning and the value of the therapeutic process were frequently discussed in the Post-Therapy Outcome Perspectives study, but rarely prioritised by young people or professionals in the Q-methodological study of Viewpoints on Outcome Priorities.

8.2.5 Comparing Change across Outcomes

This study compared reliable improvement between two measures of symptoms, two measures of functioning, and the three domains of symptoms, functioning and goal progress. The two measures of internalising symptoms yielded conflicting ratings of reliable improvement for one third of cases; and the two measures of psychosocial functioning for over 40%. There were discrepancies in 62.3% of cases when change was considered across all three domains simultaneously. Reliable improvement was highest in goal progress (66.4%) and lowest in functioning (23.3%). It was observed consistently across all three domains for only 13.8% of cases. This study suggests that judgements based exclusively on symptom metrics risk over- or underestimating actual impact, and that aggregating results from different measures or domains into a single composite indicator can mask informative differences, such as the number and type of outcome domains in which improvement was achieved. Great caution is warranted when comparing aggregate change indices obtained from different measures or outcome domains, as differences in the measures sensitivity to change, or in the temporal change trajectories observed for different outcome domains may be misinterpreted as a true difference in service performance or treatment effectiveness.

8.2.6 Integrated Summary of Key Findings

In considering all five studies, several key findings can be highlighted. First, theoretical frameworks that are informed by developmental psychopathology stress the multidimensional and dynamic nature of outcome in child mental health, where standards of 'normal' functioning are relative to a young person's developmental status, as well as to the socio-cultural norms and expectations of their environment; and where the interactions between the young person and their family and peer environment are important to consider. The Outcome Taxonomy Review highlighted the range of possible outcomes and thematic multidimensionality.

The qualitative research suggests that many of these outcomes matter to young people, parents, and clinicians when reflecting on a treatment experience. Different outcomes were emphasised by different stakeholder groups and treatment arms, although a number of generally salient 'core outcomes' could be identified. The Viewpoints on Outcome Priorities Study showed that outcome priorities vary *within* stakeholder groups too. These finding are in stark contrast with the persistent focus on clinicial outcomes – symptoms first and foremost, and functioning to a lesser extent – in treatment efficacy and effectiveness research, as highlighted by the Outcome Measurement Review. While symptom change in particular is important to key stakeholders across treatment modalities, focussing exclusively on symptom metrics does not reflect their holistic perspectives. The Comparing Change Across Outcome Domains study has demonstrated that symptom change is not an adequate proxy for change other domains; and that symptom-focussed measurement not only fails to reflect stakeholder priorities, but also risks under-or over-estimating the magnitude of change actually achieved.

8.3 Implications for Clinical Practice and Research

While there are a range of policy calls to move towards more person-centred care (see Section 1.1), the extent to which this is meaningfully delivered in practice is currently uncertain. A shift towards multidimensional measurement is warranted to advance the scientific and clinicial utility, relevance, and acceptability of outcome measurement in clinical practice and research (The Lancet Psychiatry, 2020).

Given the variety of possible outcomes, and the even greater variety of available outcome measures, standards are required to guide researchers and clinical services in designing multidimensional approaches to measurement. The Comparing Change Across Outcome Domains study has demonstrated that measures designed to capture similar concepts yield inconsistent results for an important share of service users. In clinical practice, measurement approaches must be harmonised in order to enable fair comparisons between services and systems. In clinical research, lack of standardisation hinders synthesis through meta-analyses, and leads to research waste (Chalmers & Glasziou, 2009; Szatmari et al., 2019).

The findings from the Post-Therapy Outcome Perspectives study suggest that a number of core outcomes can be identified that are salient across stakeholder groups and treatment modalities: depressive symptoms, coping skills and resilience, academic and vocational functioning, and family functioning and relationships. With the exception of academic and vocational functioning, all of these were also prioritised by at least one of the viewpoints identified in the study of Viewpoints on Outcome Priorities. The latter study also highlighted functional impairment as an outcome of particular importance to young people who struggled the most. Based on these findings, a tentative outcome battery can be suggested to replace unidimensional measurement approaches (see Table 8.2):

- Given the general salience of *depressive symptoms*, these should be measured as a core outcome, by drawing on reports from young people, parents, and clinicians. Given the difference in change levels captured by the RCADS and SDQ Emotion in Chapter 7, a detailed and nuanced assessment would be preferable over a brief and broad symptom measure, as it may be more sensitive to change.

- The observed divergence between change in symptoms and functioning suggests that the latter cannot be inferred from the former, and should be measured separately. A measure of *global functioning* should cover different domains and be completed by youth, parents, and clinicians, to reflect their complementary insights into the young person's behaviour in different contexts.

- A self-report measure of *coping skills* should be included to mirror the particular importance of this outcome category for young people in the studies of Post-Therapy Outcome Perspectives and Viewpoints on Outcome Priorities. Ideally, this measure should cover a broad range of coping skills, so that it can be applied across different treatment modalities.

- According to the findings from Chapter 5, parents appear well placed and willing to report on young people's *academic and vocational functioning*, using a suitable nomothetic measure that should cover both attendance and performance.

- The Post-Treatment Outcome Perspectives study suggests that there is some variation in the outcomes prioritised by different treatment modalities. Clinicians should rate change in relation to key treatment or disease mechanisms, which may cut across different thematic domains of the Outcome Taxonomy. To avoid bias, clinician report may need to be triangulated with matching report from young people or parents.

- The study of Viewpoints on Outcome Priorities has demonstrated that youth do not speak with one voice, but have heterogenous viewpoints. This implies that an element of personalisation is needed if routine outcome measurement is to be more than a box-ticking exercise (Batty et al., 2013; S. Merry et al., 2004; Moran et al., 2012; Sharples et al., 2017). Therefore, an idiographic measure should be included alongside nomothetic measures (Ashworth et al., 2019; King et al., 1999; Sales & Alves, 2012; Wolpert, Ford, et al., 2012). A number of relevant idiographic measures are available, including tools that focus on young people's top concerns or problems (Battle et al., 1966; Godfrey et al., 2019; Weisz et al., 2011), and tools focussing on treatment goals (Law, 2006).

Table 8.2. Suggested Outcome Battery

Outcome	Measure	Informants			Degree of standardisation
		Youth	Parents	Clinicians	
Depressive symptoms	nomothetic	✓	✓	✓	Harmonised
Global functioning	nomothetic	✓	✓	✓	Harmonised
Coping skills (broadly defined)	nomothetic	✓	—	—	Harmonised
Family functioning & relationships	nomothetic	—	—	✓	Harmonised
Academic and vocational functioning	nomothetic	—	✓	—	Harmonised
Treatment mechanisms *(theoretically informed)*	nomothetic	as appropriate	as appropriate	✓	To be tailored
Personal goals / target complaints *(individually defined)*	idiographic	✓	✓	—	To be tailored

Note. ✓ indicates that this informant should complete an assessment for the respective outcome; — indicates that this informant should not contribute to the assessment for the respective outcome.

The suggested outcome battery shown in Table 8.2 will need to be tailored, depending on whether it is used for research studies, or for routine outcome monitoring in clinical practice. Personalisation is of particular importance when tracking outcomes in routine settings, to ensure that treatment and evaluation centre around service user needs. In research contexts, personalisation may be less central, as the aim usually is to generate findings that can be compared across individuals and studies. Here, additional emphasis should be placed on specifying treatment mechanisms, and testing relevant hypotheses (Fonagy, 1997; Kazdin, 2009). A taxonomy of treatment outcomes can provide a helpful manual of possible options, and a standard terminology that can facilitate comparisons across studies (Dodd et al., 2018). In addition, reporting standards are required, which hold study authors to account for clearly stating what outcomes were assessed and reported, on what grounds this selection was made, what measures were used for assessment (Butcher et al., 2019).

Analysing and Interpreting Multidimensional Findings

Wherever different dimensions or perspectives are considered, the interpretation of measurement results becomes more difficult – as has been demonstrated by the Comparing Change Across Outcome Domains study (Chapter 7). Frameworks and tools are needed to derive judgements based on conflicting findings (Chorpita, 2001; De Los Reyes & Kazdin, 2008). A key aspect of any such approach is to embrace divergence and exploit the additional information conveyed by conflicting accounts in an attempt to create a more holistic picture of the changes achieved, and the mechanisms through which these have come about (Kazdin, 2009). This study's findings about differential levels of change measured in different outcome domains and by different measures (see Chapter 7) suggest that aggregating results from multiple sources into a single metric can both mask informative nuances in the configuration of change across measures and domains, and lead to false conclusions about differences in treatment efficacy or service effectiveness. A transparent approach is needed that can exploit the additional information conveyed by more complex approaches. To date, no referential model for doing so has been established in child mental health.

One approach that is routinely used for the evaluation of outcomes in international aid programmes is Theory of Change (Brest, 2010; Taplin & Clark, 2012). Theory of Change guides both the conceptualisation and evaluation of interventions by systematically (and visually) mapping the assumed pathways of change from the expected short-, medium-, and long-term outcomes back to the relevant programme inputs, while explicitly articulating the logical relationships between different outcomes, and the assumptions that must hold true for the outlined change process to occur (Taplin & Clark, 2012).

Articulating an intervention logic in this way has several advantages. In clinical research it can help articulate the mechanisms through which treatment is expected to improve higher-level outcomes, and identify intermediate outcome indicators through which the immediate effectiveness of these mechanisms could be tracked. This could help overcome the current knowledge gap in relation to how exactly psychotherapy leads to change (Cuijpers, 2019; Kazdin, 1999a, 2009). In clinical practice, mapping pathways between intermediate and higher-level goals jointly with young people could help identify the desired change process in a personalised way. By explicitly articulating change pathways and their relationships with one another, Theory of Change can help interpret and reconcile conflicting information (e.g., across different outcome domains or perspectives), and provide guidance in triangulating both quantitative and qualitative data to test not only whether change has occurred, but also why this has or has not been the case.

8.4 Next Steps and Areas for Future Research

This doctoral thesis has highlighted a need to develop tools and standards to promote the use of multidimensional outcome measurement in clinical research and practice. It has also highlighted the need for new research into the implications of such approaches.

There is a need to build consensus on the types of core outcomes whose measurement should be recommended by core outcome sets for clinical trials, and/or for routine measurement in clinical practice. This study has identified a number of possible core outcomes, as well as differences in perspectives. Dedicated methodologies (e.g., Delphi consultation methods; Linstone & Turoff, 1975) are needed to confirm or enhance these findings by building consensus amongst key stakeholders about the outcomes that should be included in a core outcome battery. Two such projects are currently underway. One initiative is being led by the International Consortium for Health Outcome Measurement, and focusses on developing a standard set of outcomes for child and youth anxiety, depression, OCD, and PTSD in routine practice worldwide. Its recommendations are currently being finalised (Krause et al., 2020). A second core outcome set is being developed by researchers based at *Toronto Outcomes Research in Child Health* (Hospital for Sick Children Research Institute) and the *Cundill Centre for Child and Youth Depression* (Centre for Addiction and Mental Health) in Toronto, Canada. This core outcome set will be designed more specifically for adolescent depression, and for clinical trials (Monga et al., 2019). The findings from this doctoral research project underscore the importance of including diverse groups of young people, parents, and clinicians in the development of such outcome sets, to ensure that they will be relevant and meaningful to their intended users.

This research has also showcased the need to consider idiographic measures alongside nomothetic measures, as there does not appear to be a one-size-fits-all-solution that meets the needs of all young people, parents, and clinicians; and of different treatment approaches. Future research programmes should inquire about how best to integrate such flexible elements into core outcome sets, and what guidance could be provided to help (a) researchers identify the most suitable measures to track treatment mechanisms; and (b) clinicians and services to identify the most suitable idiographic measures for use in clinical practice.

Alongside efforts to embed and harmonise multidimensional measurement approaches, a further important area for future research relates to furthering our understanding of multidimensional change, and how best to make sense of complex, multidimensional outcome data. This thesis represents a first step towards developing such a research programme. Future inquiries should explore the extent to which change trajectories and magnitudes diverge or converge across different outcome domains, and between idiographic and nomothetic measures; which factors explain these patterns; how best to reconcile conflicting information; and how best to appraise data from multiple measures, domains, and informants without losing informative nuances, and without treating metrics as equal or exchangeable, when they are not directly comparable.

8.5 Strengths and Limitations

8.5.1 General Strengths

This doctoral thesis has a number of strengths. First, adopting a pragmatic mixed-methods design provided the flexibility to explore different aspects of the overarching research question, from different angles and perspectives. Contrasting the systematic review of primarily quantitative treatment efficacy and effectiveness studies, with the qualitative analysis of semi-structured interviews as part of the Post-Therapy Outcome Perspectives study, highlighted contradictions between the unidimensional focus evident in treatment studies, and the more multifaceted narratives provided by key stakeholder groups. The two qualitative studies gave a voice to young people, parents, and clinicians, and produced findings that are grounded in their experiences, while the systematic review and the quantitative analysis of reliable change reflected the types of metrics that dominate research and service management. The triangulation of the different methods and inquiries produced richer and more comprehensive insights, than any individual approach could have done.

Second, the initial taxonomy derived from the Outcome Taxonomy Review, and successively revised through the Outcome Measurement Review and Post-Therapy Outcome Perspectives, provided a consistent (if evolving) conceptual frame of reference throughout the thesis, and facilitated the systematic identification of the types of changes considered important in clinical research, and in the eyes of key stakeholder groups (Kazdin, 1999b).

Third, by focussing specifically on adolescent depression, this thesis embarked on a more focussed and tailored exploration of ‘good outcome’ than previous studies. As such, it reflects suggestions that what is considered to constitute a ‘good outcome’ is likely to vary by presenting problems and developmental stage (Kazdin, 1999b), and that ‘value’ in mental health care delivery can only be defined in relation to specific disorders (Porter, 2010). This enabled the identification of diverse viewpoints amongst young people with lived experience of service use for depression, and showcased that even within the same age group and when focussing on a specific presenting problem, there is still considerable heterogeneity in experiences and priorities. Only when narrowing the focus of inquiry to this point is it possible to draw the conclusion that there is no one-size-fits-all solution to the question of what constitutes a good outcome of treatment, as it may otherwise be suspected that such a solution can be found at the level of specific disorders and age groups.

Fourth, the question of what constitutes a good treatment outcome was approached through a conceptual framework that was agnostic to any specific “theory of therapeutic action” (D. J. Cohen, 1995). This, in combination with the consideration of outcomes across different modalities in the empirical studies (including CBT, psychoanalytic psychotherapy, psychosocial approaches, and treatment as usual in routine care) enabled the identification of a wide range of outcomes – including at an intermediate level of treatment mechanisms. The resulting taxonomy thud provides a meta-frame of reference.

Fifth, each of the five studies placed young peoples' perspectives at the core of the inquiry. The Outcome Taxonomy Review focussed on taxonomies in child mental health; the Outcome Measurement Review explored the extent to which young people had been consulted as informants in the measurement of treatment outcomes; the Post-Therapy Outcome Perspectives study examined experiences and perceptions of outcome as described by young people (as well as parents and clinicians); the Viewpoints on Outcome Priorities study enabled young people to express their outcome priorities with minimal intervention from the researcher through the use of Q-methodology; and the Comparing Change Across Outcome Domains study compared individual-level change across outcome measures and domains self-reported data. As such, the thesis aimed to be itself person-centred. At the same time, the importance of other stakeholders was also acknowledged by considering – to the extent possible – the perspectives of parents and clinicians in the studies of Post-Therapy Outcome Perspectives and Viewpoints on Outcome Priorities.

8.5.2 General Limitations

This thesis' focus on a specific presenting problem and developmental stage (i.e., adolescent depression) is a strength, as well as a limitation. It can only mark a first step in generating a more nuanced understanding of the types of outcomes that are relevant in the treatment of depression, and a case in point for demonstrating the potential of multidimensional approaches. For a more holistic and developmentally informed understanding, complementary research would need to be conducted with younger children to better understand how their perceptions, experiences and priorities, as well as those of their parents and clinicians, differ (if at all) from those of adolescents.

Developing a taxonomy of treatment outcomes is conceptually challenging as some outcome concepts are interrelated or overlapping. For example, changes within the domain of personal growth are closely intertwined: Increased confidence and a stronger sense of identity are likely to link into more assertiveness and greater autonomy. For practical measurement, several of these concepts may be covered by composite scales capturing broader concepts. Measures of global functioning provide one example. Another challenge is the overlap between domains, such as between symptoms and wellbeing, as discussed above.

While this thesis aimed to investigate outcome concepts and priorities amongst different stakeholder groups, parents could not be included in the Viewpoints on Outcome Priorities study due to difficulties with recruitment. In addition, it was beyond the scope of this work to investigate perspectives of additional stakeholder groups such as service managers, commissioners, insurers, policy makers, and funders, who play important roles in defining what constitutes a good outcome, and in acting upon outcome data (Cuijpers, 2019). It has been shown that these groups have specific information needs that are distinct from those of young people or clinicians (Childs et al., 2013). Future studies focussing more specifically on outcome measurement in clinical practice, or on the translation of evidence from clinical studies into clinical practice, may want to explore these differences in more detail. The Q-set used for this study could be enhanced for this purpose, by first widening the concourse through interviews or focus groups with these stakeholder groups, and by expanding the Q-set on this basis.

A final point worth highlighting is that all studies included in this doctoral thesis were inherently limited by their geographical focus on England. What constitutes a good outcome cannot be determined in absolute terms, as such judgements always depend on the constellation of values in a given socio-cultural context (Fonagy, 1997; Hoagwood et al., 1996), and on wider views about what constitutes a good life or a good citizen (Binder et al., 2010). As suggested by Binder and colleagues (2010), this may take the form of fitting into socially ascribed roles or breaking with such roles where necessary, feeling good or having a sense of purpose. As part of this thesis, an attempt was originally made to compare outcome priorities amongst the UK youth and professional samples, with a sample of youth and professionals recruited from mental health services in a different cultural context: Santiago de Chile. Unfortunately, significant delays in ethical approval procedures greatly reduced the time available for recruitment in Santiago and the final samples achieved included only three young people and 15 professionals. While this data has not been included in this thesis, future research should explore the cultural conditioning of outcome concepts in more detail, especially as global initiatives to develop core outcome sets with a worldwide reach are taking shape (Krause et al., 2020)

8.6 General Conclusion

This doctoral thesis explored the question of what constitutes a good outcome for adolescent depression. It took a mixed-methods approach to examining outcome concepts, priorities, and measurement. *Reduction in depressive symptoms* was the most consistently endorsed outcome across young people, parents, and clinicians; but was usually mentioned in the context of other changes, such as an improvement in young people's *coping skills and resilience* ,in *family functioning and relationships*, or in *academic and vocational functioning*. There was considerable diversity in how different outcome domains and categories were prioritised. This suggests that while a few core outcomes can be identified as relevant to most young people, parents, and clinicians, an element of personalisation is required to ensure person-focussed measurement in clinical practice.

While symptom-focussed approaches to outcome measurement have dominated research into treatment effectiveness, these do not grasp the range of additional changes that matter to young people, parents, and clinicians, according to the qualitative research conducted as part of this thesis. In addition, purely symptom-focussed approaches risk conveying an incomplete picture of the change achieved. Young people may achieve personal treatment goals while not showing any reliable symptom change, or they may show important symptom change, but fail to improve their functioning in daily life. Multidimensional approaches to measurement are needed in order to deepen the evidence base about the effectiveness of different treatments, and their constituent treatment mechanisms.

To embed multidimensional approaches to outcome measurement in research and clinical practice, and enable comparisons across studies and settings, conceptual guidance and harmonisation are needed. These can be provided by core outcome sets, of which several are currently under development. At the same time, however, such outcome sets can only present a minimum recommendation. Researchers, clinicians, and service providers must be encouraged to identify additional outcomes that are relevant in their specific contexts. Ideally, such outcomes should be selected with reference to a comprehensive and transparent framework, such as the taxonomy developed in this thesis, on a shared decision-making tool enabling young people to identify their desired outcomes from a range of options. In order to ensure that measurement approaches are person-centred and meaningful in clinical practice, idiographic outcome measures should be considered alongside nomothetic ones. To ensure nuanced and theory-led outcome measurement in clinical research, the selection of outcome measures will need to be explicitly linked to relevant intervention mechanisms.

Two tools have been identified that may help guide both the selection of outcomes to measure, and the interpretation of findings across domains and perspectives. The first is a taxonomy of treatment outcome which showcases the full range of possibly relevant outcomes to avoid bias or omissions in the design of measurement approaches. The value of such a conceptual framework has been demonstrated in this thesis. The second is Theory of Change, a methodology that is currently rarely used in mental health, but could guide the selection of outcomes and the evaluation of findings across domains and perspectives by tying both processes closely to the theoretical assumptions underpinning the intervention logic. There is great potential for multidimensional outcome measurement to not only meet the needs of young people and their families better, but to also generate greater insight into the mechanisms by which treatment is effective. Through the use of a mixed-method approach, this doctoral thesis has thus contributed to advancing current debates and evidence about outcome concept, priorities, and measurement in relation to adolescent depression.

References

Aarons, G. A., Covert, J., Skriner, L. C., Green, A., Marto, D., Garland, A. F., & Landsverk, J. (2010). The eye of the beholder: Youths and parents differ on what matters in mental health services. *Administration and Policy in Mental Health and Mental Health Services Research*, *37*(6), 459–467. https://doi.org/10.1007/s10488-010-0276-1

Abeles, P., Verduyn, C., Robinson, A., Smith, P., Yule, W., & Proudfoot, J. (2009). Computerized CBT for adolescent depression (stressbusters) and its initial evaluation through an extended case series. *Behavioural and Cognitive Psychotherapy*, *37*(2), 151–165. https://doi.org/10.1017/S1352465808005067

Achenbach, T. M. (1991a). *Integrative Guide to the 1991 CBCL/4–18, YSR, and TRF Profiles.* Burlington, VT: University of Vermont, Department of Psychology.

Achenbach, T. M. (1991b). *Manual for the Child Behavior Checklist/4-18 and 1991 profile.* Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. M. (2006). As Others See Us: Clinical and Research Implications of Cross-Informant Correlations for Psychopathology. *Current Directions in Psychological Science*, *15*(2), 94–98.

Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H. C., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: Research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *49*(3), 251–275. https://doi.org/10.1111/j.1469-7610.2007.01867.x

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/Adolescent Behavioral and Emotional Problems: Implications of Cross-Informant Correlations for Situational Specificity. *Psychological Bulletin*, *101*(2), 213–232. https://doi.org/10.1037/0033-2909.101.2.213

Altepeter, T. S., & Breen, M. J. (1989). The Home Situations Questionnaire (HSQ) and the School Situations Questionnaire (SSQ): Normative Data and an Evaluation of Psychometric Properties. *Journal of Psychoeducational Assessment*, Vol. 7, pp. 312–322. https://doi.org/10.1177/073428298900700404

Amaya, M. M., Reinecke, M. A., Silva, S. G., & March, J. S. (2011). Parental marital discord and treatment response in depressed adolescents. *Journal of Abnormal Child Psychology*, *39*(3), 401–411. https://doi.org/10.1007/s10802-010-9466-2

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Arlington, VA: American Psychiatric Association.

American Society for Indexing. (2019). About Taxonomies & Controlled Vocabularies. Retrieved October 15, 2019, from http://www.taxonomies-sig.org/about.htm

Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, *5*(4), 237–249.

Angold, A., Costello, J. E., Farmer, E. M. Z., Burns, B., & Erkanli, A. (1999). Impaired but undiagnosed. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*(2), 129–137. https://doi.org/10.1097/00004583-199902000-00011

Angold, A., Weissman, M. M., John, K., Merikancas, K. R., Prusoff, B. A., Wickramaratne, P., … Warner, V. (1987). Parent and child reports of depressive symptoms in children at low and high risk of depression. *Journal of Child Psychology and Psychiatry*, *28*(6), 901–915.

Angst, F., Stassen, H. H., Clayton, P. J., & Angst, J. (2002). Mortality of patients with mood disorders: follow-up over 34–38 years. *Journal of Affective Disorders*, *68*(2–3), 167–181.

Anthony, W. A. (1993). Recovery from mental illness: the guiding vision of the mental health service system in the 1990s. *Psychosocial Rehabilitation Journal*, *16*(4), 521–537.

APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271–285. https://doi.org/10.1037/0003-066X.61.4.271

Ariga, M., Uehara, T., Takeuchi, K., Ishige, Y., Nakano, R., & Mikuni, M. (2010). Follow-up study of female delinquent adolescents in a detention centre: Effectiveness of psychiatric intervention as a mental health service. *International Journal of Social Psychiatry*, *56*(1), 15–22. https://doi.org/10.1177/0020764008095114

Arrindell, W. A., & van der Ende, J. (1985). Cross-sample invariance of the structure of self-reported distress and difficulty in assertiveness. Experiences with the scale for interpersonal behaviour. *Advances in Behaviour Research and Therapy*, *7*(4), 205–243. https://doi.org/10.1016/0146-6402(85)90013-X

Asarnow, J. R., Emslie, G., Clarke, G., Wagner, K. D., Spirito, A., Vitiello, B., … Brent, D. (2009). Treatment of selective serotonin reuptake inhibitor-resistant depression in adolescents: Predictors and moderators of treatment response. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(3), 330–339. https://doi.org/10.1097/CHI.0b013e3181977476

Asarnow, J. R., Jaycox, L. H., Duan, N., LaBorde, A. P., Rea, M. M., Murray, P., … Wells, K. B. (2005). Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: A randomized controlled trial. *Journal of the American Medical Association*, *293*(3), 311–319. https://doi.org/10.1001/jama.293.3.311

Ascher, B. H., Z. Farmer, E. M., Burns, B. J., & Angold, A. (1996). The Child and Adolescent Services Assessment (CASA). *Journal of Emotional and Behavioral Disorders*, *4*(1), 12–20. https://doi.org/10.1177/106342669600400102

Ashworth, M., Guerra, D., & Kordowicz, M. (2019). Individualised or Standardised Outcome Measures: A Co-habitation? *Administration and Policy in Mental Health and Mental Health Services Research*, *46*(4), 425–428. https://doi.org/10.1007/s10488-019-00928-z

Ashworth, M., Robinson, S., Evans, C., Shepherd, M., Conolly, A., & Rowlands, G. (2007). What does an idiographic measure (PSYCHLOPS) tell us about the spectrum of psychological issues and scores on a nomothetic measure (CORE-OM)? *Primary Care and Community Psychiatry*, *12*(1), 7–16. https://doi.org/10.1080/17468840701560805

Ashworth, M., Robinson, S. I., Godfrey, E., Shepherd, M., Evans, C., Seed, P., … Tylee, A. (2005). Measuring mental health outcomes in primary care: The psychometric properties of a new patient-generated outcome measure, “PSYCHLOPS” ('psychological outcome profiles’). *Primary Care Mental Health*, *3*(4), 261–270.

Ashworth, M., Shepherd, M., Christey, J., Matthews, V., Wright, K., Parmentier, H., … Godfrey, E. (2004). A client-generated psychometric instrument: The development of ‘PSYCHLOPS.’ *Counselling and Psychotherapy Research*, *4*(2), 27–31. https://doi.org/10.1080/14733140412331383913

Atkinson, S. D., Prakash, A., Zhang, Q., Pangallo, B. A., Bangs, M. E., Emslie, G. J., & March, J. S. (2014). A Double-Blind Efficacy and Safety Study of Duloxetine Flexible Dosing in Children and Adolescents with Major Depressive Disorder. *Journal of Child and Adolescent Psychopharmacology*, *24*(4), 180–189. https://doi.org/10.1089/cap.2013.0146

Australian Commission on Safety and Quality in Healthcare. (2010). *Patient-Centred Care: Improving Quality and Safety by Focusing Care on Patients and Consumers Discussion Paper.* Retrieved from http://www.safetyandquality.gov.au/wp-content/uploads/2012/01/PCCC-DiscussPaper.pdf

Australian Institute of Health and Welfare. (2003). The Active Australia Survey: A guide and manual for implementation, analysis and reporting. Retrieved December 7, 2019, from https://www.aihw.gov.au/reports/physical-activity/active-australia-survey/formats

Baker, R., Thompson, C., & Mannion, R. (2006). Q methodology in health economics. *Journal of Health Services Research and Policy*, *11*(1), 38–45. https://doi.org/10.1258/135581906775094217

Ballenger, J. C. (1999). Clinical guidelines for establishing remission in patients with depression and anxiety. *Journal of Clinical Psychiatry*, *60*(Suppl. 22), 29–34.

Bamber, D., Tamplin, A., Park, R. J., Kyte, Z. A., & Goodyer, I. M. (2002). Development of a Short Leyton Obsessional Inventory for Children and Adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*(10), 1246–1252. https://doi.org/10.1097/00004583-200210000-00015

Bambery, M., Porcerelli, J. H., & Ablon, J. S. (2007). Measuring Psychotherapy Process With the Adolescent Psychotherapy Q-Set (Apq): Development and Applications for Training. *Psychotherapy*, *44*(4), 405–422. https://doi.org/10.1037/0033-3204.44.4.405

Battle, C. C., Imber, S. D., Hoehn-Saric, R., Nash, E. R., & Frank, J. D. (1966). Target complaints as criteria of improvement. *American Journal of Psychotherapy*, *20*(1), 184–192.

Batty, M. J., Moldavsky, M., Foroushani, P. S., Pass, S., Marriott, M., Sayal, K., & Hollis, C. (2013). Implementing routine outcome measures in child and adolescent mental health services: From present to future practice. *Child and Adolescent Mental Health*, *18*(2), 82–87. https://doi.org/10.1111/j.1475-3588.2012.00658.x

Bear, H. A., Edbrooke-Childs, J., Norton, S., Krause, K. R., & Wolpert, M. (2020). Systematic Review and Meta-Analysis: Outcomes of Routine Specialist Mental Health Care for Young People With Depression and/or Anxiety. *Journal of the American Academy of Child & Adolescent Psychiatry*. https://doi.org/10.1016/j.jaac.2019.12.002

Beavers, W. R., & Hampson, R. B. (1990). *Successful families: assessment and intervention.* New York, NY: W. W. Norton.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* Oxford, England: Oxford University Press.

Beck, A. T. (1993). Cognitive therapy: past, present, and future. *Journal of Consulting and Clinical Psychology*, *61*(2), 194.

Beck, A. T., Kovacs, M., & Weissman, A. (1979). Assessment of suicidal intention: The Scale for Suicide Ideation. *Journal of Consulting and Clinical Psychology*, *47*(2), 343–352. https://doi.org/10.1037/0022-006X.47.2.343

Beck, A. T., & Steer, R. A. (1993). *Beck Hopelessness Scale manual.* San Antonio, TX: The Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory-ii (bdi-ii).* San Antonio, TX: Psychological Corporation.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*(1), 77–100. https://doi.org/10.1016/0272-7358(88)90050-5

Beck, J. S., Beck, A. T., Jolly, J., & Steer, R. A. (2005). *Beck Youth Inventories^{TM} - Second Edition For Children and Adolescents (BYI-II)* (2nd ed.). San Antonio, TX: Pearson.

Becker-Weidman, E. G., Jacobs, R. H., Reinecke, M. A., Silva, S. G., & March, J. S. (2010). Social problem-solving among adolescents treated for depression. *Behaviour Research and Therapy*, *48*(1), 11–18. https://doi.org/10.1016/j.brat.2009.08.006

Becker, K. D., Brandt, N. E., Stephan, S. H., & Chorpita, B. F. (2014). A review of educational outcomes in the children's mental health treatment literature. *Advances in School Mental Health Promotion*, Vol. 7, pp. 5–23. https://doi.org/10.1080/1754730X.2013.851980

Becker, K. D., Chorpita, B. F., & Daleiden, E. L. (2011). Improvement in symptoms versus functioning: How do our best treatments measure up? *Administration and Policy in Mental Health and Mental Health Services Research*, *38*(6), 440–458. https://doi.org/10.1007/s10488-010-0332-x

Beiser, M. (1990). *Final report submitted in fulfillment of requirements for the grants of the United States National Institute of Mental Health (5-RO1-MH36678–04) and the Canada Health and Welfare National Health Research Directorate Program (NHRDP 6610–132–04).* Toronto, ON, Canada.

Bellack, A. S. (2006). Scientific and consumer models of recovery in schizophrenia: Concordance, contrasts, and implications. *Schizophrenia Bulletin*, *32*(3), 432–442. https://doi.org/10.1093/schbul/sbj044

Benjamin, M. O., Melendez, G., Davis, D. C., & Lunnen, K. M. (2001). The Ohio scales: Practical outcome assessment. *Journal of Child and Family Studies*, *10*(2), 199–212. https://doi.org/10.1023/A:1016651508801

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Beresford, P., & Branfield, F. (2006). Developing inclusive partnerships: User-defined outcomes, networking and knowledge - A case study. *Health and Social Care in the Community*, *14*(5), 436–444. https://doi.org/10.1111/j.1365-2524.2006.00654.x

Bergmans, Y., Langley, J., Links, P., & Lavery, J. V. (2009). The perspectives of young adults on recovery from repeated suicide-related behavior. *Crisis*, *30*(3), 120–127. https://doi.org/10.1027/0227-5910.30.3.120

Bergner, M., Bobbitt, R. A., Carter, W. B., & Gilson, B. S. (1981). The Sickness Impact Profile: Development and Final Revision of a Health Status Measure. *Medical Care*, *19*(8), 787–805. https://doi.org/10.1097/00005650-198108000-00001

Bickman, L., & Rosof-Williams, J. (2000). What information do clinicians value for monitoring adolescent client progress and outcomes? *Professional Psychology: Research and Practice*, *31*(1), 70–74. https://doi.org/10.1037//0735-702S.3U.70

Biernacki, P., & Waldorf, D. (1981). Snowball Sampling: Problems and Techniques of Chain Referral Sampling. *Sociological Methods & Research*, *10*(2), 141–163. https://doi.org/10.1177/004912418101000205

Binder, P. E., Holgersen, H., & Nielsen, G. H. (2010). What is a “good outcome” in psychotherapy? A qualitative exploration of former patients’ point of view. *Psychotherapy Research*, *20*(3), 285–294. https://doi.org/10.1080/10503300903376338

Binder, P., Moltu, C., Hummelsund, D., Henden, S., & Holgersen, H. (2011). Meeting an adult ally on the way out into the world: Adolescent patients’ experiences of useful psychotherapeutic ways of working at an age when independence really matters. *Psychotherapy Research*, *21*(5), 554–566. https://doi.org/10.1080/10503307.2011.587471

Bird, H. R., Gould, M. S., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*(1), 78–85.

Birmaher, B. (2007). Practice Parameter for the Assessment and Treatment of Children and Adolescents With Depressive Disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(11), 1503–1526. https://doi.org/10.1097/chi.0b013e318145ae1c

Black, N. (2013). Patient reported outcome measures could help transform healthcare. *BMJ*, *346*(f167). https://doi.org/10.1136/bmj.f167

Blanton, H., & Jaccard, J. (2006). Arbitrary Metrics in Psychology. *American Psychologist*, *61*(1), 27–41. https://doi.org/10.1037/0003-066X.61.1.27

Blatt, S. J., Schaffer, C. E., Bers, S. A., & Quinlan, D. M. (1992). Psychometric Properties of the Depressive Experiences Questionnaire for Adolescents. *Journal of Personality Assessment*, *59*(1), 82–98. https://doi.org/10.1207/s15327752jpa5901_8

Bloomquist, M. L., Giovanelli, A., Benton, A., Piehler, T. F., Quevedo, K., & Oberstar, J. (2016). Implementation and Evaluation of Evidence-Based Psychotherapeutic Practices for Youth in a Mental Health Organization. *Journal of Child and Family Studies*, *25*(11), 3278–3292. https://doi.org/10.1007/s10826-016-0479-5

Bolger, F., Stranieri, A., & Yearwood, J. (2011). Does the Delphi process lead to increased accuracy in group-based judgmental forecasts or does it simply induce consensus amongst judgmental forecasters? *Technological Forecasting and Social Change*, *78*(9), 1671–1680. https://doi.org/10.1016/j.techfore.2011.06.002

Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, *25*(1), 6–19. https://doi.org/10.1080/10503307.2013.817696

Bothwell, S., & Weissman, M. M. (1977). Social Impairments Four Years After an Acute Depressive Episode. *American Journal of Orthopsychiatry*, *47*(2), 231–237. https://doi.org/10.1111/j.1939-0025.1977.tb00978.x

Bradley, J., Murphy, S., Fugard, A. J. B., Nolas, S.-M., & Law, D. (2013). What kind of goals do children and young people set for themselves in therapy? Developing a goals framework using CORC data. *Child & Family Clinical Psychology Review*, *1*(1), 8–18.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brent, D., Emslie, G., Clarke, G., Wagner, K. D., Asarnow, J. R., Keller, M., … Zelazny, J. (2008). Switching to another SSRI or to venlafaxine with or without cognitive behavioral therapy for adolescents with SSRI-resistant depression: The TORDIA randomized controlled trial. *JAMA*, *299*(8), 901–913. https://doi.org/10.1001/jama.299.8.901

Brest, P. (2010). The power of theories of change. *Stanford Social Innovation Review*, *8*(2), 47–51. Retrieved from https://search-proquest-com.libproxy.ucl.ac.uk/docview/340379561?accountid=14511

Bringhurst, M. D. L., Watson, C. W., Miller, S. D., & Duncan, B. L. (2006). The reliability and validity of the Outcome Rating Scale: A replication study of a brief clinical measure. *Journal of Brief Therapy*, *5*(1), 23–30.

Brookman-Frazee, L., Haine, R. A., & Garland, A. F. (2006). Innovations: Child and Adolescent Psychiatry: Measuring Outcomes of Real-World Youth Psychotherapy: Whom to Ask and What to Ask? *Psychiatric Services*, *57*(10), 1373–1375. https://doi.org/10.1176/ps.2006.57.10.1373

Brown, A., Ford, T., Deighton, J., & Wolpert, M. (2014). Satisfaction in child and adolescent mental health services: Translating users' feedback into measurement. *Administration and Policy in Mental Health and Mental Health Services Research*, *41*(4), 434–446. https://doi.org/10.1007/s10488-012-0433-9

Brown, S. R. (1970). On The Use of Variance Designs in Q Methodology. *The Psychological Record*, *20*(2), 179–189. https://doi.org/10.1007/bf03393928

Brown, S. R. (1980). *Political Subjectivity: Applications of Q Methodology in Political Science*. New Haven, CT: Yale University Press.

Brown, S. R. (1993). A Primer on Q Methodology. *Operant Subjectivity*, *16*(3/4), 91–138.

Brown, S. R. (2002). Q Technique and Questionnaires. *Operant Subjectivity*, *25*(2), 117–126.

Brown, S. R. (2005). Applying Q Methodology to Empowerment. In D. Narayan (Ed.), *Measuring Empowerment: Cross-Disciplinary Perspectives* (pp. 197–215). Retrieved from http://documents.worldbank.org/curated/en/960161468175149824/pdf/344100PAPER0Me101Official0use0only1.pdf

Brown, S. R. (2006). A match made in heaven: A marginalized methodology for studying the marginalized. *Quality and Quantity*, *40*(3), 361–382. https://doi.org/10.1007/s11135-005-8828-2

Brunstein, J. C. (1993). Personal Goals and Subjective Well-Being: A Longitudinal Study. *Journal of Personality and Social Psychology*, *65*(5), 1061–1070. https://doi.org/10.1037/0022-3514.65.5.1061

Bullington, P., Pawola, L., Walker, R., Valenta, A., Briars, L., & John, E. (2007). Identification of medication non-adherence factors in adolescent transplant patients: The patient's viewpoint. *Pediatric Transplantation*, *11*, 914–921. https://doi.org/10.1111/j.1399-3046.2007.00765.x

Busner, J., & Targum, S. D. (2007). The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry (Egmont)*, *4*(7), 28–37.

Butcher, N. J., Monsour, A., Mew, E. J., Szatmari, P., Pierro, A., Kelly, L. E., … Offringa, M. (2019). Improving outcome reporting in clinical trial reports and protocols: Study protocol for the Instrument for reporting Planned Endpoints in Clinical Trials (InsPECT). *Trials*, *20*(1), 1–10. https://doi.org/10.1186/s13063-019-3248-0

Byford, S. (2013). The validity and responsiveness of the EQ-5D measure of health-related quality of life in an adolescent population with persistent major depression. *Journal of Mental Health*, *22*(2), 101–110. https://doi.org/10.3109/09638237.2013.779366

Byford, S., Barrett, B., Roberts, C., Wilkinson, P., Dubicka, B., Kelvin, R. G., … Goodyer, I. (2007). Cost-effectiveness of selective serotonin reuptake inhibitors and routine specialist care with and without cognitive-behavioural therapy in adolescents with major depression. *British Journal of Psychiatry*, *191*, 521–527. https://doi.org/10.1192/bjp.bp.107.038984

Cairns, A. J., Kavanagh, D. J., Dark, F., & McPhail, S. M. (2019). Goal setting improves retention in youth mental health: A cross-sectional analysis. *Child and Adolescent Psychiatry and Mental Health*, *13*(1), 1–8. https://doi.org/10.1186/s13034-019-0288-x

Calderon, A., Schneider, C., Target, M., & Midgley, N. (2017). The Adolescent Psychotherapy Q-Set (APQ): A Validation Study. *Journal of Infant, Child, and Adolescent Psychotherapy*, *16*(1), 106–120. https://doi.org/10.1080/15289168.2016.1255499

Campbell, A. G., & Hemsley, S. (2009). The Validity of the Outcome Rating Scale and Session

Rating Scale in Psychological Practice. *The Clinical Psychologist*, *13*(1), 1–10.

Canino, G., Costello, E. J., & Angold, A. (1999). Assessing Functional Impairment and Social Adaptation for Child Mental Health Services Research: A Review of Measures. *Mental Health Services Research*, *1*(2), 93–108. https://doi.org/10.1023/A:1022334303731

Canino, G., Fisher, P. W., Alegria, M., & Bird, H. R. (2013). Assessing Child Impairment in Functioning in Different Contexts: Implications for Use of Services and the Classification of Psychiatric Disorders. *Open Journal of Medical Psychology*, *2*(01), 29–34. https://doi.org/10.4236/ojmp.2013.21006

Cantwell, D. P., Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1997). Correspondence between adolescent report and parent report of psychiatric diagnostic data. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*(5), 610–619. https://doi.org/10.1097/00004583-199705000-00011

Carandang, C., Robbins, D., Mullany, E., Yazbek, M., & Minot, S. (2007). Lamotrigine in adolescent mood disorders: A retrospective chart review. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *16*(1), 1–8.

Carskadon, M. A. (2004). Sleep difficulties in youg people. *Archives of Pediatrics & Adolescent Medicine*, *158*, 597–598.

Carter, T., Guo, B., Turner, D., Morres, I., Khalil, E., Brighton, E., … Callaghan, P. (2015). Preferred intensity exercise for adolescents receiving treatment for depression: A pragmatic randomised controlled trial. *BMC Psychiatry*, *15*(1), 247. https://doi.org/10.1186/s12888-015-0638-z

Carter, T., Morres, I., Repper, J., & Callaghan, P. (2016). Exercise for adolescents with depression: Valued aspects and perceived change. *Journal of Psychiatric and Mental Health Nursing*, *23*(1), 37–44. https://doi.org/10.1111/jpm.12261

Casey, P., Patalay, P., Deighton, J., Miller, S. D., & Wolpert, M. (2019). The Child Outcome Rating Scale: validating a four-item measure of psychosocial functioning in community and clinic samples of children aged 10–15. *European Child and Adolescent Psychiatry*, (0123456789). https://doi.org/10.1007/s00787-019-01423-4

Castagna, P. J., Calamia, M., & Davis, T. E. (2019). The Discrepancy between Mother and Youth Reported Internalizing Symptoms Predicts Youth’s Negative Self-Esteem. *Current Psychology*, (Advance online publication). https://doi.org/10.1007/s12144-019-00501-9

Chafey, M. I. J., Bernal, G., & Rossello, J. (2009). Clinical case study: CBT for depression in a puerto rican adolescent: challenges and variability in treatment response. *Depression and Anxiety*, *26*(1), 98–103. https://doi.org/10.1002/da.20457

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *Obstetrics and Gynecology*, *114*(6), 1341–1345. https://doi.org/10.1097/AOG.0b013e3181c3020d

Charkhandeh, M., Talib, M. A., & Hunt, C. J. (2016). The clinical effectiveness of cognitive behavior therapy and an alternative medicine approach in reducing symptoms of depression in adolescents. *Psychiatry Research*, *239*, 325–330. https://doi.org/10.1016/j.psychres.2016.03.044

Chee, T. T., Ryan, A. M., Wasfy, J. H., & Borden, W. B. (2016). Current State of Value-Based Purchasing Programs. *Circulation*, *133*(22), 2197–2205. https://doi.org/10.1161/CIRCULATIONAHA.115.010268

Cheung, A. H., Levitt, A. J., & Szalai, J. P. (2003). Impact of Antidepressant Side Effects on Adolescent Quality of Life. *Canadian Journal of Psychiatry*, *48*(11), 727–733. https://doi.org/10.1177/070674370304801104

Cheung, A. H., Mayes, T., Levitt, A., Schaffer, A., Michalak, E., Kiss, A., & Emslie, G. (2010). Anxiety As a Predictor of Treatment Outcome in Children and Adolescents with Depression. *Journal of Child and Adolescent Psychopharmacology*, *20*(3), 211–216. https://doi.org/10.1089/cap.2010.0006

Child Outcomes Research Consortium. (2012). Session Feedback Questionnaire. Retrieved December 8, 2019, from https://www.corc.uk.net/media/1405/sfq_questionnaire.pdf

Childs, J., Deighton, J., & Wolpert, M. (2013). *Defining and measuring mental health and wellbeing: A response mode report requested by the Department of Health for the Policy Research Unit in the Health of Children, Young People and Families*. Retrieved from http://www.ucl.ac.uk/ebpu/docs/publication_files/Defining_and_measuring_mental_health_and_wellbeing_in_children-CPRU_RM_report.pdf

Chorpita, B. F. (2001). Reflections on Clinical Significance: What Do Our Best Treatments Accomplish and How Can We Best Find Out? *Clinical Psychology: Science and Practice*, *8*(4), 451–454. https://doi.org/10.1093/clipsy/8.4.451

Chorpita, B. F., Moffitt, C. E., & Gray, J. (2005). Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. *Behaviour Research and Therapy*, *43*(3), 309–322. https://doi.org/10.1016/j.brat.2004.02.004

Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A revised child anxiety and depression scale. *Behaviour Research and Therapy*, *38*(8), 835–855. https://doi.org/10.1016/S0005-7967(99)00130-8

Christie, G., Marsh, R., Sheridan, J., Wheeler, A., Suaalii-Sauni, T., Black, S., & Butler, R. (2007). The substances and choices scale (SACS)--the development and testing of a new alcohol and other drug screening and outcome measurement instrument for young people. *Addiction*, *102*(9), 1390–1398. https://doi.org/10.1111/j.1360-0443.2007.01916.x

Christmas, D. M. B., & Sweeney, A. (2016). Service user, patient, survivor or client... has the time come to return to “patient”? *British Journal of Psychiatry*, *209*(1), 9–13. https://doi.org/10.1192/bjp.bp.115.167221

Cicchetti, D., & Toth, S. (1995). A Developmental Psychopathology Perspective on Child Abuse and Neglect. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*(5), 541–565. https://doi.org/10.1097/00004583-199505000-00008

Clarke, G. N., & DeBar, L. L. (2010). Group Cognitive-Behavioral Treatment for Adolescent Depression. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents.* (2nd ed., pp. 110–125). New York, NY: The Guilford Press.

Clarke, G. N., DeBar, L. L., Pearson, J. A., Dickerson, J. F., Lynch, F. L., Gullion, C. M., & Leo, M. C. (2016). Cognitive Behavioral Therapy in Primary Care for Youth Declining Antidepressants: A Randomized Trial. *Pediatrics*, *137*(5), e20151851–e20151851. https://doi.org/10.1542/peds.2015-1851

Clarke, G. N., Lewinsohn, P. M., & Hops, H. (1990). *Instructor’s manual for the adolescent coping with depression course*. Portland, OR: Kaiser Permanente Center for Health Research.

Clarke, M., & Williamson, P. (2015). Core outcome sets and trial registries. *Trials*, *16*(1), 15–17. https://doi.org/10.1186/s13063-015-0738-6

Clayborne, Z. M., Varin, M., & Colman, I. (2019). Systematic Review and Meta-Analysis: Adolescent Depression and Long-Term Psychosocial Outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *58*(1), 72–79. https://doi.org/10.1016/j.jaac.2018.07.896

Cleare, A., Pariante, C. M., Young, A. H., Anderson, I. M., Christmas, D., Cowen, P. J., … Uher, R. (2015). Evidence-based guidelines for treating depressive disorders with antidepressants: A revision of the 2008 British Association for Psychopharmacology guidelines. In *Journal of Psychopharmacology* (Vol. 29). https://doi.org/10.1177/0269881115581093

Cohen, D. J. (1995). Psychosocial therapies for children and adolescents: Overview and future directions. *Journal of Abnormal Child Psychology*, *23*(1), 141–156. https://doi.org/10.1007/BF01447049

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

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

Cole, D. A., Hoffman, K., Tram, J. M., & Maxwell, S. E. (2000). Structural differences in parent and child reports of children’s symptoms of depression and anxiety. *Psychological Assessment*, *12*(2), 174–185. https://doi.org/10.1037/1040-3590.12.2.174

Conners, K. C., Sitarenios, G., Parker, J. D. A., & Epstein, J. N. (1998). The revised Conners’ Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology*, *26*(4), 257–268. https://doi.org/10.1023/A:1022602400621

Connolly, M. B., & Strupp, H. H. (1996). Cluster analysis of patient reported psychotherapy outcomes. *Psychotherapy Research*, *6*(1), 30–42. https://doi.org/10.1080/10503309612331331558

Cooper, M., & Duncan, L. (Eds.). (2018). *Working with Goals in Psychotherapy and Counselling*. Oxford, United Kingdom: Oxford University Press.

Cornelius, J. R., Bukstein, O. G., Wood, D. S., Kirisci, L., Douaihy, A., & Clark, D. B. (2009). Double-blind placebo-controlled trial of fluoxetine in adolescents with comorbid major depression and an alcohol use disorder. *Addictive Behaviors*, *34*(10), 905–909. https://doi.org/10.1016/j.addbeh.2009.03.008

Cornelius, J. R., Douaihy, A., Bukstein, O. G., Daley, D. C., Wood, S. D., Kelly, T. M., & Salloum, I. M. (2011). Evaluation of cognitive behavioral therapy/motivational enhancement therapy (CBT/MET) in a treatment trial of comorbid MDD/AUD adolescents. *Addictive Behaviors*, *36*(8), 843–848. https://doi.org/10.1016/j.addbeh.2011.03.016

Cortés, A. M., Fernández, O. M., Capella, C. M., Espinosa, H. D., & Krause, M. (2018). ¿ Qué Es lo que Cambia en la Psicoterapia ? La Perspectiva de Pacientes Adolescentes con Síntomas Depresivos, sus Terapeutas y Observadores Externos? *PSYKHE*, *27*(2), 1–15. https://doi.org/https://doi.org/10.7764/psykhe.27.1.1134

Costa da Silva, L., & Wolpert, M. (2018). *Outcome Measures in Child and Youth Mental Health Services: Results from NHS England Survey [Unpublished manuscript]*. London, United Kingdom: Evidence Based Practice Unit, University College London.

Costello, E. J., Angold, A., & Keeler, G. (2003). Adolescent Outcomes of Childhood Disorders: The Consequences of Severity and Impairment. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*(2), 121–128. https://doi.org/10.1097/00004583-199902000-00010

Costello, E. Jane, Pine, D. S., Hammen, C., March, J. S., Plotsky, P. M., Weissman, M. M., … Leckman, J. F. (2002). Development and natural history of mood disorders. *Biological*

Psychiatry, *52*(6), 529–542. https://doi.org/10.1016/S0006-3223(02)01372-0

Costello, E.J., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and Development of Psychiatric Disorders in Childhood and Adolescence. *Archives of General Psychiatry*, *60*, 837–844.

Cottrell, D. (2002). Practitioner Review: The effectiveness of systemic family therapy for children and adolescents. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*(5), 573–586. https://doi.org/10.1111/1469-7610.00047

Coulter, A. (2017). Measuring what matters to patients. *BMJ (Online)*, *356*, j816. https://doi.org/10.1136/bmj.j816

Cregeen, S., Hughes, C., Midgley, N., Rhode, M., & Rustin, M. (2018). Short-Term Psychoanalytic Psychotherapy for Adolescents with Depression. In *Short-Term Psychoanalytic Psychotherapy for Adolescents with Depression*. https://doi.org/10.4324/9780429480164

Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Creswell, J. W. (2003). *Research design: qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publications.

Crowe, M., Inder, M., & Porter, R. (2015). Conducting qualitative research in mental health: Thematic and content analyses. *Australian and New Zealand Journal of Psychiatry*, *49*(7), 616–623. https://doi.org/10.1177/0004867415582053

Cuijpers, P. (2019). Targets and outcomes of psychotherapies for mental disorders: an overview. *World Psychiatry*, *18*(3), 276–285. https://doi.org/10.1002/wps.20661

Cuijpers, P., Li, J., Hofmann, S. G., & Andersson, G. (2010). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: A meta-analysis. *Clinical Psychology Review*, *30*(6), 768–778. https://doi.org/10.1016/j.cpr.2010.06.001

Cuijpers, P., Stringaris, A., & Wolpert, M. (2020). Treatment outcomes for depression: challenges and opportunities. *The Lancet Psychiatry*, *0366*(2019), 2019–2020. https://doi.org/10.1016/S2215-0366(20)30036-5

Cull, J. G., & Gill, W. S. (1988). *Suicide Probability Scale (SPS) manual.* Los Angels, CA: Western Psychological Services.

Curry, J., Silva, S., Rohde, P., Ginsburg, G., Kratochvil, C., Simons, A., … March, J. (2011). Recovery and recurrence following treatment for adolescent major depression. *Archives of General Psychiatry*, *68*(3), 263–270. https://doi.org/10.1001/archgenpsychiatry.2010.150

Danielson, S. (2009). Q Method and Surveys: Three Ways to Combine Q and R. *Field Methods*, *21*(3), 219–237. https://doi.org/10.1177/1525822X09332082

Davidson, L., & Roe, D. (2007). Recovery from versus recovery in serious mental illness: One strategy for lessening confusion plaguing recovery. *Journal of Mental Health*, *16*(4), 459–470. https://doi.org/10.1080/09638230701482394

Davis-Brown, K., Carter, N., & Miller, B. D. (2012). Youth Advisors Driving Action: Hearing the Youth Voice in Mental Health Systems of Care. *Journal of Psychosocial Nursing and Mental Health Services*, *50*(3), 39–43. https://doi.org/10.3928/02793695-20120207-07

De Hert, M., Correll, C. U., Bobes, J., Cetkovich-Bakmas, M., Cohen, D. A. N., Asai, I., … Leucht, S. (2011). Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. *World Psychiatry*, *10*(1), 52–77. https://doi.org/10.1002/j.2051-5545.2011.tb00014.x

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4), 858–900. https://doi.org/10.1037/a0038498

De Los Reyes, A., Cook, C. R., Gresham, F. M., Makol, B. A., & Wang, M. (2019). Informant discrepancies in assessments of psychosocial functioning in school-based services and research: Review and directions for future research. *Journal of School Psychology*, *74*, 74–89. https://doi.org/10.1016/j.jsp.2019.05.005

De Los Reyes, A., & Kazdin, A. E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment*, *16*(3), 330–334. https://doi.org/10.1037/1040-3590.16.3.330

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, *131*(4), 483–509. https://doi.org/10.1037/0033-2909.131.4.483

De Los Reyes, A., & Kazdin, A. E. (2008). When the Evidence Says, “Yes, No, and Maybe So.” *Current Directions in Psychological Science*, *17*(1), 47–51. https://doi.org/10.1111/j.1467-8721.2008.00546.x

De Los Reyes, A., Kundey, S. M. A., & Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review*, *31*(5), 829–838. https://doi.org/10.1016/j.cpr.2011.03.011

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. A. (2013). Principles Underlying the Use of Multiple Informants’ Reports. *Annual Review of Clinical Psychology*, *9*(1), 123–149. https://doi.org/10.1146/annurev-clinpsy-050212-185617

De Ross, R. L., Gullone, E., & Chorpita, B. F. (2002). The Revised Child Anxiety and Depression Scale: A psychometric investigation with Australian youth. *Behaviour Change*, *19*(2), 90–101. https://doi.org/10.1375/bech.19.2.90

De Smet, M. M., Meganck, R., De Geest, R., Norman, U. A., Truijens, F., & Desmet, M. (2019). What "good outcome" means to patients: Understanding recovery and improvement in psychotherapy for major depression from a mixed-methods perspective. *Journal of Counseling Psychology*. https://doi.org/10.1037/cou0000362

De Smet, M. M., Meganck, R., Van Nieuwenhove, K., Truijens, F. L., & Desmet, M. (2019). No Change? A Grounded Theory Analysis of Depressed Patients' Perspectives on Non-improvement in Psychotherapy. *Frontiers in Psychology*, *10*, 1–17. https://doi.org/10.3389/fpsyg.2019.00588

Dean, R., Siddiqui, S., Beesley, F., Fox, J., & Berry, K. (2018). Staff perceptions of borderline personality disorder and recovery: A Q-sort method approach. *British Journal of Clinical Psychology*, *57*(4), 473–490. https://doi.org/10.1111/bjc.12180

Deegan, P. E. (1988). Recovery: The lived experience of rehabilitation. *Psychosocial Rehabilitation Journal*, *11*(4), 11–19.

DelBello, M. P., Hochadel, T. J., Portland, K. B., Azzaro, A. J., Katic, A., Khan, A., & Emslie, G. (2014). A Double-Blind, Placebo-Controlled Study of Selegiline Transdermal System in Depressed Adolescents. *Journal of Child and Adolescent Psychopharmacology*, *24*(6), 311–317. https://doi.org/10.1089/cap.2013.0138

Delis, D. C. (2012). *Delis rating of executive function (D-REF).* Bloomington, MN: Pearson.

Department of Health. (2010). *Equity and excellence: Liberating the NHS*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/213823/dh_117794.pdf

Department of Health. (2011). *Talking therapies: A four-year plan of action. A supporting document to No health without mental health: A cross- government mental health outcomes strategy for people of all ages*. Retrieved from http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_123759

Department of Health. (2015). *Future in mind: Promoting, protecting and improving our children and young people's mental health and wellbeing*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/414024/Childrens_Mental_Health.pdf

Department of Health and Ageing. (2003). *Mental Health National Outcomes and Casemix Collection: Overview of Clinician-Rated and Consumer Self-Report Measures V1.50*. Retrieved from https://www.amhocn.org/sites/default/files/publication_files/mhnocc_overview_doha.pdf

Diamond, G. S., Wintersteen, M. B., Brown, G. K., Diamond, G. M., Gallop, R., Shelef, K., & Levy, S. (2010). Attachment-Based Family Therapy for Adolescents with Suicidal Ideation: A Randomized Controlled Trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(2), 122–131. https://doi.org/10.1097/00004583-201002000-00006

Dilorenzo, T. A., Bovbjerg, D. H., Montgomery, G. H., Valdimarsdottir, H., & Jacobsen, P. B. (1999). The application of a shortened version of the profile of mood states in a sample of breast cancer chemotherapy patients. *British Journal of Health Psychology*, *4*(4), 315–325. https://doi.org/10.1348/135910799168669

Dirks, M. A., De Los Reyes, A., Briggs-Gowan, M., Cella, D., & Wakschlag, L. S. (2012). Annual research review: Embracing not erasing contextual variability in children's behavior - Theory and utility in the selection and use of methods and informants in developmental psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *53*(5), 558–574. https://doi.org/10.1111/j.1469-7610.2012.02537.x

Dodd, S., Clarke, M., Becker, L., Mavergames, C., Fish, R., & Williamson, P. R. (2018). A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *Journal of Clinical Epidemiology*, *96*, 84–92. https://doi.org/10.1016/j.jclinepi.2017.12.020

Domino, M. E., Burns, B. J., Silva, S. G., Kratochvil, C. J., Vitiello, B., Reinecke, M. A., … March, J. S. (2008). Cost-effectiveness of treatments for adolescent depression: Results from TADS. *American Journal of Psychiatry*, *165*(5), 588–596. https://doi.org/10.1176/appi.ajp.2008.07101610

Domino, M. E., Foster, E. M., Vitiello, B., Kratochvil, C. J., Burns, B. J., Silva, S. G., … March, J. S. (2009). Relative cost-effectiveness of treatments for adolescent depression: 36-week results from the tads randomized trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(7), 711–720. https://doi.org/10.1097/CHI.0b013e3181a2b319

Donner, J. C. (2001). Using Q-sorts in Participatory Processes: An Introduction to the Methodology. In R. A. Krueger, M. A. Casey, J. Donner, S. Kirsch, & J. N. Maack (Eds.), *Social analysis: Selected Tools and Techniques* (pp. 24–49). Washington, DC: The World Bank.

dosReis, S., Camelo Castillo, W., Ross, M., N'Dri, L., & Butler, B. (2018). Treatment Outcome Priorities for Youth with Multi-Morbid Mental Health Conditions: A Latent Class Analysis of a Best-Worst Scaling Experiment. *Value in Health*, *21*, S127. https://doi.org/10.1016/j.jval.2018.04.962

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, *52*(6), 377–384. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9764259

Dudley, A., Ford, S., Melvin, G. A., & Tonge, B. J. (2005). *Outcome of Adolescent Depression Structured Interview [Unpublished manuscript]*. Melbourne, Australia: Centre for Developmental Psychiatry & Psychology, Monash University.

Duncan, B. L., Miller, S. D., Sparks, J. A., Claud, D. A., Reynolds, L. R., Brown, J., … Reynolds, L. R. (2003). The Session Rating Scale: preliminary psychometric properties of a “working” alliance measure. *Journal of Brief Therapy*, *3*(1), 3–12.

Duncan, B. L., Sparks, J. A., Miller, S. D., & Bohanske, R. T. (2006). Giving youth a voice: A preliminary study of the reliability and validity of a brief outcome measure for children, adolescents, and caretakers. *Journal of Brief Therapy*, *5*(2), 71–88.

Dunn, O. J. (1961). Multiple Comparisons among Means. *Journal of the American Statistical Association*, *56*(293), 52–64. https://doi.org/10.1080/01621459.1961.10482090

Dunne, A., Thompson, W., & Leitch, R. (2000). Adolescent males’ experience of the counselling process. *Journal of Adolescence*, *23*, 79–93. https://doi.org/10.1006/jado.1999.0300

Duong, M. T., Lyon, A. R., Ludwig, K., Wasse, J. K., & McCauley, E. (2016). Student perceptions of the acceptability and utility of standardized and idiographic assessment in school mental health. *International Journal of Mental Health Promotion*, *18*(1), 49–63. https://doi.org/10.1080/14623730.2015.1079429

Dworkin, S. L. (2012). Sample size policy for qualitative studies using in-depth interviews. *Archives of Sexual Behavior*, *41*(6), 1319–1320. https://doi.org/10.1007/s10508-012-0016-6

Dyrborg, J., Warborg Larsen, F., Nielsen, S., Byman, J., Buhl Nielsen, B., & Gautrè-Delay, F. (2000). The Children’s Global Assessment Scale (CGAS) and Global Assessment of Psychosocial Disability (GAPD) in clinical practice - Substance and reliability as judged by intraclass correlations. *European Child and Adolescent Psychiatry*, *9*(3), 195–201. https://doi.org/10.1007/s007870070043

Edbrooke-Childs, J., Jacob, J., Law, D., Deighton, J., & Wolpert, M. (2015). Interpreting standardized and idiographic outcome measures in CAMHS: what does change mean and how does it relate to functioning and experience? *Child and Adolescent Mental Health*, *20*(3), 142–148. https://doi.org/10.1111/camh.12107

Edbrooke-Childs, J., Wolpert, M., & Deighton, J. (2016). Introduction to the Special Section on Implementing Feedback from Outcome Measures in Child and Adult Mental Health

Services. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(3), 277–278. https://doi.org/10.1007/s10488-016-0726-5

Edbrooke-Childs, J., Wolpert, M., Zamperoni, V., Napoleone, E., & Bear, H. (2018). Evaluation of reliable improvement rates in depression and anxiety at the end of treatment in adolescents. *BJPsych Open*, *4*(4), 250–255. https://doi.org/10.1192/bjo.2018.31

Elkin, I., Parloff, M. B., Hadley, S. W., & Autry, J. H. (1985). NIMH Treatment of Depression Collaborative Research Program: Background and Research Plan. *Archives of General Psychiatry*, *42*(3), 305–316. https://doi.org/10.1001/archpsyc.1985.01790260103013

Ellingsen, I. T., Shemmings, D., & Størksen, I. (2011). The Concept of ‘Family’ Among Norwegian Adolescents in Long-Term Foster Care. *Child and Adolescent Social Work Journal*, *28*(4), 301–318. https://doi.org/10.1007/s10560-011-0234-0

Ellingsen, I. T., Størksen, I., & Stephens, P. (2010). Q methodology in social work research. *International Journal of Social Research Methodology*, *13*(5), 395–409. https://doi.org/10.1080/13645570903368286

Ellingsen, I. T., Thorsen, A. A., & Størksen, I. (2014). Revealing Children’s Experiences and Emotions through Q Methodology. *Child Development Research*, *2014*, 1–9. https://doi.org/10.1155/2014/910529

Elliott, R., Wagner, J., Sales, C. M. D., Rodgers, B., Alves, P., & Café, M. J. (2016). Psychometrics of the personal questionnaire: A client-generated outcome measure. *Psychological Assessment*, *28*(3), 263–278. https://doi.org/10.1037/pas0000174

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Emslie, G. J., Findling, R. L., Yeung, P. P., Kunz, N. R., & Li, Y. (2007). Venlafaxine ER for the treatment of pediatric subjects with depression: Results of two placebo-controlled trials. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(4), 479–488. https://doi.org/10.1097/chi.0b013e31802f5f03

Emslie, G. J., Heiligenstein, J. H., Wagner, K. D., Hoog, S. L., Ernest, D. E., Brown, E., … Jacobson, J. G. (2002). Fluoxetine for Acute Treatment of Depression in Children and Adolescents: A Placebo-Controlled, Randomized Clinical Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*(10), 1205–1215. https://doi.org/10.1097/00004583-200210000-00010

Emslie, G. J., Kennard, B. D., Mayes, T. L., Nakonezny, P. A., Zhu, L., Tao, R., … Croarkin, P. (2012). Insomnia Moderates Outcome of Serotonin-Selective Reuptake Inhibitor Treatment in Depressed Youth. *Journal of Child and Adolescent Psychopharmacology*, *22*(1), 21–28. https://doi.org/10.1089/cap.2011.0096

Emslie, G. J., Kennard, B. D., Mayes, T. L., Nightingale-Teresi, J., Carmody, T., Hughes, C. W., … Rintelmann, J. W. (2008). Fluoxetine versus placebo in preventing relapse of major depression in children and adolescents. *American Journal of Psychiatry*, *165*(4), 459–467. https://doi.org/10.1176/appi.ajp.2007.07091453

Emslie, G. J., Mayes, T. L., Porta, G., Vitiello, B., Clarke, G. N., Wagner, K. D., … Brent D. (2010). Treatment of Resistant Depression in Adolescents. *American Journal of Psychiatry*, *167*(7), 782–791. https://doi.org/10.1176/appi.ajp.2010.09040552

Emslie, G. J., Prakash, A., Zhang, Q., Pangallo, B. A., Bangs, M. E., & March, J. S. (2014). A Double-Blind Efficacy and Safety Study of Duloxetine Fixed Doses in Children and Adolescents with Major Depressive Disorder. *Journal of Child and Adolescent Psychopharmacology*, *24*(4), 170–179. https://doi.org/10.1089/cap.2013.0096

Emslie, G. J., Rush, A. J., Weinberg, W. A., Kowatch, R. A., Hughes, C. W., Carmody, T., & Rintelmann, J. (1997). A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Archives of General Psychiatry*, *54*(11), 1031–1037. https://doi.org/10.1001/archpsyc.1997.01830230069010

Emslie, G. J., Ventura, D., Korotzer, A., & Tourkodimitris, S. (2009). Escitalopram in the treatment of adolescent depression: A randomized placebo-controlled multisite trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(7), 721–729. https://doi.org/10.1097/CHI.0b013e3181a2b304

Emslie, G. J., Yeung, P. P., & Kunz, N. R. (2007). Long-term, open-label venlafaxine extended-release treatment in children and adolescents with major depressive disorder. *CNS Spectrums: The International Journal of Neuropsychiatric Medicine*, *12*(3), 223–233.

Endicott, J., Nee, J., Yang, R., & Wohlberg, C. (2006). Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (PQ-LES-Q): Reliability and validity. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*(4), 401–407. https://doi.org/10.1097/01.chi.0000198590.38325.81

Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: A Procedure for Measuring Overall Severity of Psychiatric Disturbance. *Archives of General Psychiatry*, *33*(6), 766–771. https://doi.org/10.1001/archpsyc.1976.01770060086012

Epstein, N. B., Baldwin, L. M., & Bishop, D. (1983). The McMaster Family Assessment Device. *Journal of Marital and Family Therapy*, *9*, 171–180.

Esbjørn, B. H., Sømhovd, M. J., Turnstedt, C., & Reinholdt-Dunne, M. L. (2012). Assessing the revised child anxiety and depression scale (RCADS) in a national sample of Danish youth aged 8-16 years. *PLoS ONE*, *7*(5), e37339. https://doi.org/10.1371/journal.pone.0037339

Eskin, M., Ertekin, K., & Demir, H. (2008). Efficacy of a problem-solving therapy for depression and suicide potential in adolescents and young adults. *Cognitive Therapy and Research*, *32*(2), 227–245. https://doi.org/10.1007/s10608-007-9172-8

Eunice Kennedy Shriver National Institute of Child Health and Human Development. (2019). NICHD Pediatric Terminology Files. Retrieved November 6, 2019, from https://evs.nci.nih.gov/ftp1/Pediatric_Terminologies/NICHD/About.html

Evans, C., Margison, F., & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence-Based Mental Health*, *1*(3), 70–72. https://doi.org/10.1136/ebmh.1.3.70

Fabiano, G. A., Pelham, W. E., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M., … Burrows-MacLean, L. (2006). A practical measure of impairment: Psychometric properties of the impairment rating scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. *Journal of Clinical Child and Adolescent Psychology*, *35*(3), 369–385. https://doi.org/10.1207/s15374424jccp3503_3

Feeney, J. A., Noller, P., & Hanrahan, M. (1994). Assessing adult attachment. In M. B. Sperling & W. H. Berman (Eds.), *Attachment in Adults: clinical and developmental perspectives* (pp. 128–152). New York: NY: Guilford Press.

Feeny, N. C., Silva, S. G., Reinecke, M. A., McNulty, S., Findling, R. L., Rohde, P., … March, J. S. (2009). An exploratory analysis of the impact of family functioning on treatment for depression in adolescents. *Journal of Clinical Child and Adolescent Psychology*, *38*(6), 814–825. https://doi.org/10.1080/15374410903297148

Ferdinand, R. F., Van Der Ende, J., & Verhulst, F. C. (2006). Prognostic value of parent-adolescent disagreement in a referred sample. *European Child and Adolescent Psychiatry*, *15*(3), 156–162. https://doi.org/10.1007/s00787-005-0518-z

Ferrier, I. N. (1999). Treatment of major depression: Is improvement enough? *Journal of Clinical Psychiatry*, *60*(Suppl. 6), 10–14.

Findling, R. L., Robb, A., & Bose, A. (2013). Escitalopram in the Treatment of Adolescent Depression: A Randomized, Double-Blind, Placebo-Controlled Extension Trial. *Journal of Child and Adolescent Psychopharmacology*, *23*(7), 468–480. https://doi.org/10.1089/cap.2012.0023

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. . (1997). *Structured Clinical Interview for DSM-IV Axis I Disorders Clinician Version (SCID-CV).* Washington, DC.: American Psyhiatric Press.

Fischer, J., & Corcoran, K. (2007). *Measures for clinical practice and research. A sourcebook.* (4th ed.). Oxford, United Kingdom: Oxford University Press.

Fitzpatrick, R., Davey, C., Buxton, M. J., & Jones, D. R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment*, *2*(14).

Fleiss, J. L. (1981). *Statistical methods for rates and proportions.* New York, NY: John Wiley and Sons.

Fletcher, J. M. (2008). Adolescent depression: diagnosis, treatment, and educational attainment. *Health Economics*, *17*(11), 1215–1235. https://doi.org/10.1002/hec.1319

Fletcher, J. M. (2013). Adolescent Depression and Adult Labor Market Outcomes. *Southern Economic Journal*, *80*(1), 26–49. https://doi.org/10.4284/0038-4038-2011.193

Fombonne, E., Wostear, G., Cooper, V., Harrington, R., & Rutter, M. (2001). The Maudsley long-term follow-up of child and adolescent depression. *British Journal of Psychiatry*, *179*(03), 210–217. https://doi.org/10.1192/bjp.179.3.210

Fonagy, P. (1997). Evaluating the effectiveness of interventions in child psychiatry. *Canadian Journal of Psychiatry*, *42*(6), 584–594. https://doi.org/10.1177/070674379704200604

Fonagy, P. (2010). Psychotherapy research: Do we know what works for whom? *British Journal of Psychiatry*, *197*(2), 83–85. https://doi.org/10.1192/bjp.bp.110.079657

Fonagy, P., & Bateman, A. W. (2006). Mechanisms of change in mentalization-based treatment of BPD. *Journal of Clinical Psychology*, *62*(4), 411–430. https://doi.org/10.1002/jclp.20241

Fonagy, P., Cottrell, D., Phillips, J., Bevington, D., Glaser, D., & Allison, E. (2015). *What Works For Whom? A Critical Review of Psychotherapy Research* (2nd ed.). New York, NY: Guilford Press.

Frank, E., Prien, R. F., Jarrett, R. B., Keller, M. B., Kupfer, D. J., Lavori, P. W., … Weissman, M. M. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. *Archives of General Psychiatry*, *48*(9), 851–855.

Frank, J. D., & Frank, J. B. (1993). *Persuasion and Healing: A Comparative Study of Psychotherapy* (3rd ed.). Baltimore, MD: Johns Hopkins University Press.

Franz, A., Worrell, M., & Vögele, C. (2013). Integrating Mixed Method Data in Psychological Research: Combining Q Methodology and Questionnaires in a Study Investigating Cultural and Psychological Influences on Adolescent Sexual Behavior. *Journal of Mixed Methods Research*, *7*(4), 370–389. https://doi.org/10.1177/1558689813483368

Freake, H., Barley, V., & Kent, G. (2007). Adolescents' views of helping professionals: A review of the literature. *Journal of Adolescence*, *30*, 639–653. https://doi.org/10.1016/j.adolescence.2006.06.001

Fried, E. I. (2016). Are more responsive depression scales really superior depression scales? *Journal of Clinical Epidemiology*, *77*, 4–6. https://doi.org/10.1016/j.jclinepi.2016.05.004

Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, *208*(July 2016), 191–197. https://doi.org/10.1016/j.jad.2016.10.019

Fried, E. I., & Nesse, R. M. (2014). The impact of individual depressive symptoms on impairment of psychosocial functioning. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0090311

Friesen, B. J. (2007). Recovery and resilience in children's mental health: Views from the field. *Psychiatric Rehabilitation Journal*, *31*(1), 38–48. https://doi.org/10.2975/31.1.2007.38.48

Gage, N. L. (1989). The Paradigm Wars and Their Aftermath A "Historical" Sketch of Research on Teaching Since 1989. *Educational Researcher*, *18*(7), 4–10. https://doi.org/10.3102/0013189X018007004

Gallagher, K., & Porock, D. (2010). The Use of Interviews in Q Methodology: Card Content Analysis. *Nursing Research July/August*, *59*(4), 295–300.

Garland, A. F., Lewczyk-Boxmeyer, C. M., Gabayan, E. N., & Hawley, K. M. (2004). Multiple Stakeholder Agreement on Desired Outcomes for Adolescents ' Mental Health Services. *Psychiatric Services*, *55*(6), 671–676. https://doi.org/10.1176/appi.ps.55.6.671

Garner, D. M., Bohr, Y., & Garfinkel, P. E. (1982). The Eating Attitudes Test: Psychometric Features and Clinical Correlates. *Psychological Medicine*, *12*(4), 871–878. https://doi.org/10.1017/S0033291700049163

Garoff, F. F., Heinonen, K., Pesonen, A. K., & Almqvist, F. (2012). Depressed youth: Treatment outcome and changes in family functioning in individual and family therapy. *Journal of Family Therapy*, *34*(1), 4–23. https://doi.org/10.1111/j.1467-6427.2011.00541.x

Garralda, E. M. (2009). Accountability of specialist child and adolescent mental health services. *British Journal of Psychiatry*, *194*(5), 389–391. https://doi.org/10.1192/bjp.bp.108.059477

Geertz, C. (1973). *The Interpretation of Cultures: Selected Essays*. New York: NY: Basic Books.

Gibbs, T., Tan, A., Weis, R. J., & Coffey, B. J. (2012). Stimulant Medication Addition to Antidepressant in an Adolescent with Difficult to Treat Depression. *Journal of Child and Adolescent Psychopharmacology*, *22*(6), 466–469. https://doi.org/10.1089/cap.2012.2263

Gibson, K., & Cartwright, C. (2014). Young clients' narratives of the purpose and outcome of counselling. *British Journal of Guidance and Counselling*, *42*(5), 511–524. https://doi.org/10.1080/03069885.2014.925084

Gibson, K., Cartwright, C., Kerrisk, K., Campbell, J., & Seymour, F. (2016). What Young People Want: A Qualitative Study of Adolescents ' Priorities for Engagement Across Psychological Services. *Journal of Child and Family Studies*, *25*(4), 1057–1065. https://doi.org/10.1007/s10826-015-0292-6

Glasman, D., Finlay, W. M. L., & Brock, D. (2004). Becoming a self-therapist: Using cognitive-behavioural therapy for recurrent depression and/or dysthymia after completing therapy. *Psychology and Psychotherapy: Theory, Research and Practice*, *77*(3), 335–351. https://doi.org/10.1348/1476083041839385

Godfrey, E., Aubrey, M., Crockford, S., Haythorne, D., Kordowicz, M., & Ashworth, M. (2019). The development and testing of PSYCHLOPS Kids: a new child-centred outcome measure. *Child and Adolescent Mental Health*, *24*(1), 54–65. https://doi.org/10.1111/camh.12271

Godin, G. (2011). The Godin-Shephard leisure-time physical activity questionnaire. *Health & Fitness Journal of Canada*, *4*(1), 18–22.

Goldberg, D. P. (1972). *The Detection of Psychiatric Illness by Questionnaire (CHQ-28).* Oxford, United Kingdom: Oxford University Press.

Goldman, H. H., Skodol, A. E., & Lave, T. R. (1992). Revising axis V for DSM-IV: A review of measures of social functioning. *American Journal of Psychiatry*, *149*(9), 1148–1156. https://doi.org/10.1176/ajp.149.9.1148

Goodman, A., & Goodman, R. (2009). Strengths and difficulties questionnaire as a dimensional measure of child mental health. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(4), 400–403. https://doi.org/10.1097/CHI.0b013e3181985068

Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the strengths and difficulties questionnaire (SDQ): Data from british parents, teachers and children. *Journal of Abnormal Child Psychology*, *38*(8), 1179–1191. https://doi.org/10.1007/s10802-010-9434-x

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *J. Child Psychol. Psychiat*, *38*(5), 581–586.

Goodman, R. (1999). The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *40*(5), 791–799.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 1337–1345. https://doi.org/http://dx.doi.org/10.1097/00004583-200111000-00015

Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *European Child and Adolescent Psychiatry*, *7*(3), 125–130.

Goodman, R., Renfrew, D., & Mullick, M. (2000). Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *European Child & Adolescent Psychiatry*, *9*(2), 129–134.

Goodman, R., & Scott, S. (1999). Comparing the Strengths and Difficulties Questionnaire and the child behavior checklist: Is small beautiful? *Journal of Abnormal Child Psychology*, *27*(1), 17–24. https://doi.org/10.1023/A:1022658222914

Goodyer, I. M., Dubicka, B., Wilkinson, P., Kelvin, R., Roberts, C., Byford, S., … Harrington, R. (2007). Selective serotonin reuptake inhibitors (SSRIs) and routine specialist care with and without cognitive behaviour therapy in adolescents with major depression: Randomised controlled trial. *British Medical Journal*, *335*(7611), 142–146. https://doi.org/10.1136/bmj.39224.494340.55

Goodyer, I. M., Reynolds, S., Barrett, B., Byford, S., Dubicka, B., Hill, J., … Fonagy, P. (2017). Cognitive behavioural therapy and short-term psychoanalytical psychotherapy versus a brief psychosocial intervention in adolescents with unipolar major depressive disorder (IMPACT): a multicentre, pragmatic, observer-blind, randomised controlled superiori. *The Lancet Psychiatry*, *4*(2), 109–119. https://doi.org/10.1016/S2215-0366(16)30378-9

Goodyer, I. M., Tsancheva, S., Byford, S., Dubicka, B., Hill, J., Kelvin, R., … Fonagy, P. (2011). Improving mood with psychoanalytic and cognitive therapies (IMPACT): A pragmatic effectiveness superiority trial to investigate whether specialised psychological treatment reduces the risk for relapse in adolescents with moderate to severe unipolar depres. *Trials*, *12*(1), 175. https://doi.org/10.1186/1745-6215-12-175

Gordon, M. S., Tonge, B., & Melvin, G. A. (2011). Outcome of adolescent depression: 6 Months after treatment. *Australian and New Zealand Journal of Psychiatry*, *45*(3), 232–239. https://doi.org/10.3109/00048674.2010.538838

Gould, M. S., King, R., Greenwald, S., Fisher, P., Schwab-Stone, M., Kramer, R., … Shaffer, D. (1998). Psychopathology associated with suicidal ideation and attempts among children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*(9), 915–923. https://doi.org/10.1097/00004583-199809000-00011

Gowers, S. G., Harrington, R. C., Whitton, A., Lelliott, P., Beevor, A., Wing, J., & Jezzard, R. (1999). Brief scale for measuring the outcomes of emotional and behavioural disorders in children. Health of the Nation Outcome Scales for Children and Adolescents (HoNOSCA). *British Journal of Psychiatry*, *174*(MAY), 413–416. https://doi.org/10.1192/bjp.174.5.413

Gray, M. (2017). Value based healthcare. *BMJ (Online)*, *356*, 1–3. https://doi.org/10.1136/bmj.j437

Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation*, *48*(6), 1–6. https://doi.org/10.1111/eci.12931

Grenville, J., & Lyne, P. (1995). Patient-centred evaluation and rehabilitative care. *Journal of Advanced Nursing*, *22*(5), 965–972. https://doi.org/10.1111/j.1365-2648.1995.tb02649.x

Grosse Holtforth, M., & Grawe, K. (2002). BERN INVENTORY OF TREATMENT GOALS: PART 1. Development and First Application of a Taxonomy of Treatment Goal Themes. *Psychotherapy Research*, *12*(1), 79–99. https://doi.org/10.1080/713869618

Grosse Holtforth, M., Reubi, I., Ruckstuhl, L., Berking, M., & Grawe, K. (2004). The value of treatment-goal themes for treatment planning and outcome evaluation of psychiatric inpatients. *The International Journal of Social Psychiatry*, *50*(1), 80–91. https://doi.org/10.1177/0020764004040955

Grossoehme, D. H., & Gerbetz, L. (2004). Adolescent Perceptions of Meaningfulness of Psychiatric Hospitalization. *Clinical Child Psychology and Psychiatry*, *9*(4), 589–596. https://doi.org/10.1177/1359104504046162

Gunlicks-Stoessel, M., Mufson, L., Jekal, A., & Turner, J. B. (2010). The Impact of Perceived Interpersonal Functioning on Treatment for Adolescent Depression: IPT-A Versus Treatment as Usual in School-Based Health Clinics. *Journal of Consulting and Clinical Psychology*, *78*(2), 260–267. https://doi.org/10.1037/a0018935

Hajer, M. A. (1997). *The Politics of Environmental Discourse: Ecological Modernization and the Policy Process*. https://doi.org/10.1093/019829333X.001.000

Hamilton, M. (1986). The Hamilton rating scale for depression. In N. Sartorius & A. Ban, Thomas (Eds.), *Assessment of Depression* (pp. 143–152). Berlin, Germany: Springer.

Hammersley, M. (1992). The Paradigm Wars: Reports from the front. *British Journal of Sociology of Education*, *13*(1), 131–143. https://doi.org/10.1080/0142569920130110

Hammersley, M. (1996). The relationship between qualitative and quantitative research: Paradigm loyalty versus methodological eclecticism. In J. Richardson (Ed.), *Handbook of Qualitative Research Methods for Psychology and the Social Sciences* (pp. 159–174). https://doi.org/10.1038/sj.bdj.2010.1088

Hardeveld, F., Spijker, J., De Graaf, R., Nolen, W. A., & Beekman, A. T. F. (2010). Prevalence and predictors of recurrence of major depressive disorder in the adult population. *Acta Psychiatrica Scandinavica*, *122*(3), 184–191. https://doi.org/10.1111/j.1600-0447.2009.01519.x

Hardeveld, F., Spijker, J., De Graaf, R., Nolen, W. A., & Beekman, A. T. F. (2013). Recurrence of major depressive disorder and its predictors in the general population: Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Psychological Medicine*, *43*(1), 39–48. https://doi.org/10.1017/S0033291712002395

Haver, A., Akerjordet, K., Caputi, P., Magee, C., & Furunes, T. (2015). Measuring mental well-being: A validation of the Short Warwick–Edinburgh Mental Well-Being Scale in Norwegian and Swedish. *Scandinavian Journal of Public Health*, *43*(7), 721–727. https://doi.org/10.1177/1403494815588862

Hawley, K. M., & Weisz, J. R. (2003). Child, Parent, and Therapist (Dis)Agreement on Target Problems in Outpatient Therapy: The Therapist’s Dilemma and Its Implications. *Journal of Consulting and Clinical Psychology*, *71*(1), 62–70. https://doi.org/10.1037/0022-006X.71.1.62

Heath, S., Charles, V., Crow, G., & Wiles, R. (2007). Informed consent, gatekeepers and go-betweens: Negotiating consent in child- and youth-orientated institutions. *British Educational Research Journal*, *33*(3), 403–417. https://doi.org/10.1080/01411920701243651

Heppner, P. P., & Petersen, C. H. (1982). The development and implications of a personal problem-solving inventory. *Journal of Counseling Psychology*, *29*(1), 66–75. https://doi.org/10.1037/0022-0167.29.1.66

Herrman, H., Kieling, C., McGorry, P., Horton, R., Sargent, J., & Patel, V. (2019). Reducing the global burden of depression: a Lancet–World Psychiatric Association Commission. *The Lancet*, *393*(10189), e42–e43. https://doi.org/10.1016/S0140-6736(18)32408-5

Hill, A. B. (1937). *Principles of medical statistics.* London, England: Lancet.

Hill, C. E., Chui, H., & Baumann, E. (2013). Revisiting and reenvisioning the outcome problem in psychotherapy: An argument to include individualized and qualitative measurement. *Psychotherapy*, *50*(1), 68–76. https://doi.org/10.1037/a0030571

Hirschfeld, R., Dunner, D. L., Keitner, G., Klein, D. N., Koran, L. M., Kornstein, S. G., … Keller, M. B. (2002). Does psychosocial functioning improve independent of depressive symptoms? A comparison of nefazodone, psychotherapy, and their combination. *Biological Psychiatry*, *51*(2), 123–133. https://doi.org/10.1016/S0006-3223(01)01291-4

Hirschfeld, R., Montgomery, S. A., Keller, M. B., Kasper, S., Schatzberg, A. F., Möller, H.-J., … Versiani, M. (2000). Social functioning in depression: a review. *The Journal of Clinical Psychiatry*.

Hoagwood, K. E., Jensen, P. S., Acri, M. C., Serene Olin, S., Eric Lewandowski, R., & Herman, R. J. (2012). Outcome domains in child mental health research since 1996: Have they changed and why does it matter? *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(12), 1241–1260. https://doi.org/10.1016/j.jaac.2012.09.004

Hoagwood, K. E., Jensen, P. S., Petti, T., & Burns, B. J. (1996). Outcomes of mental health care for children and adolescents: I. A comprehensive conceptual model. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(8), 1055–1063. https://doi.org/10.1097/00004583-199608000-00017

Hodges, Kay; Doucette-Gates, Ann; Kim, C.-S. (2000). Predicting service utilization with the Child and Adolescent Functional Assessment Scale in a sample of youths with serious emotional disturbance served by center for mental health services-funded demonstrations. *The Journal of Behavioral Health Services & Research*, *27*(1), 47–59.

Hodges, K., Gordon, Y., & Lennon, M. P. (1990). Parent-Child Agreement on Symptoms Assessed via a Clinical Research Interview for Children: The Child Assessment Schedule (CAS). *Journal of Child Psychology and Psychiatry*, *31*(3), 427–436. https://doi.org/10.1111/j.1469-7610.1990.tb01579.x

Holsen, I., & Birkeland, M. S. (2017). Course, Predictors, and Outcomes of Depressed Mood in a Cohort of Norwegian Adolescents Followed From Age 13 Into Adulthood Age 30. *Emerging Adulthood*, *5*(1), 3–15. https://doi.org/10.1177/2167696816635810

Hook, D. (2001). Discourse, Knowledge, Materiality, History: Foucault and Discourse Analysis. *Theory & Psychology*, *11*(4), 521–547. https://doi.org/10.1177/0959354301114006

Horwitz, S. M., Hoagwood, K., Stiffman, A. R., Summerfeld, T., Weisz, J. R., Costello, E. J., … Norquist, G. (2001). Reliability of the Services Assessment for Children and Adolescents. *Psychiatric Services*, *52*(8), 1088–1094. https://doi.org/10.1176/appi.ps.52.8.1088

Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, *17*(8), 10–16.

Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Hunsley, J., & Mash, E. J. (2007). Evidence-Based Assessment. *Annual Review of Clinical Psychology*, *3*(1), 29–51. https://doi.org/10.1146/annurev.clinpsy.3.022806.091419

Hurn, J., Kneebone, I., & Cropley, M. (2006). Goal setting as an outcome measure: A systematic review. *Clinical Rehabilitation*, *20*(9), 756–772. https://doi.org/10.1177/0269215506070793

Iftene, F., Predescu, E., Stefan, S., & David, D. (2015). Rational-emotive and cognitive-behavior therapy (REBT/CBT) versus pharmacotherapy versus REBT/CBT plus pharmacotherapy in the treatment of major depressive disorder in youth; A randomized clinical trial. *Psychiatry Research*, *225*(3), 687–694. https://doi.org/10.1016/j.psychres.2014.11.021

Impact Study CBT Sub-Group. (2010). *Cognitive Behaviour Therapy for Depression in Young People: Manual for Therapists.* Cambridge, United Kingdom.

International Alliance of Patients' Organizations. (2007). *What is pateint-centred health care? A review of definitions and principles*. Retrieved from http://iapo.org.uk/sites/default/files/files/IAPO Patient-Centred Healthcare Review 2nd edition.pdf

Jackson-Blott, K., Hare, D. J., Morgan, S., & Davies, B. E. (2019). Recovery from psychosis in a forensic service: Assessing staff and service users' perspectives using Q methodology. *Journal of Forensic Psychology Research and Practice*, *19*(2), 147–169. https://doi.org/10.1080/24732850.2018.1556516)

Jacob, J. (2019). Moving toward a better understanding of idiographic outcome measurement: A commentary on Lloyd, Duncan, and Cooper (2019). *Clinical Psychology: Science and Practice*, (March), e12287. https://doi.org/10.1111/cpsp.12287

Jacob, J., Edbrooke-Childs, J., Holley, S., Law, D., & Wolpert, M. (2016). Horses for courses? A qualitative exploration of goals formulated in mental health settings by young people, parents, and clinicians. *Clinical Child Psychology and Psychiatry*, *21*(2), 208–223. https://doi.org/10.1177/1359104515577487

Jacob, J., Edbrooke-Childs, J., Law, D., & Wolpert, M. (2017). Measuring what matters to patients: Using goal content to inform measure choice and development. *Clinical Child Psychology and Psychiatry*, *22*(2), 170–186. https://doi.org/10.1177/1359104515615642

Jacob, J., Edbrooke-Childs, J., Lloyd, C., Hayes, D., Whelan, I., Wolpert, M., & Law, D. (2018). Measuring outcomes using goals. In Mick Cooper & D. Law (Eds.), *Working with Goals in Psychotherapy and Counselling* (pp. 111–138). Oxford, United Kingdom: Oxford University Press.

Jacobs, R. H., Becker-Weidman, E. G., Reinecke, M. A., Jordan, N., Silva, S. G., Rohde, P., & March, J. S. (2010). Treating depression and oppositional behavior in adolescents. *Journal of Clinical Child and Adolescent Psychology*, *39*(4), 559–567. https://doi.org/10.1080/15374416.2010.486318

Jacobs, R. H., Silva, S. G., Reinecke, M. A., Curry, J. F., Ginsburg, G. S., Kratochvil, C. J., & March, J. S. (2009). Dysfunctional attitudes scale perfectionism: A predictor and partial mediator of acute treatment outcome among clinically depressed adolescents. *Journal* *of Clinical Child and Adolescent Psychology*, *38*(6), 803–813. https://doi.org/10.1080/15374410903259031

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*(4), 336–352. https://doi.org/10.1016/S0005-7894(84)80002-7

Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects. *Journal of Consulting and Clinical Psychology*, *67*(3), 300–307.

Jacobson, N. S., & Truax, P. (1991). Clinical Significance: A Statistical Approach to Defining Meaningful Change in Psychotherapy Research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Jane Costello, E., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(12), 1263–1271. https://doi.org/10.1111/j.1469-7610.2006.01682.x

Jedeloo, S., van Staa, A. L., Latour, J. M., & van Exel, N. J. A. (2010). Preferences for health care and self-management among Dutch adolescents with chronic conditions: A Q-methodological investigation. *International Journal of Nursing Studies*, *47*(5), 593–603. https://doi.org/10.1016/j.ijnurstu.2009.10.006

Jensen, P. S., Hoagwood, K., & Petti, T. (1996). Outcomes of mental health care for children and adolescents: II. Literature review and application of a comprehensive model. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(8), 1064–1077. https://doi.org/10.1097/00004583-199608000-00018

Jensen, S. A., & Corralejo, S. M. (2017). Measurement Issues: Large effect sizes do not mean most people get better – clinical significance and the importance of individual results. *Child and Adolescent Mental Health*, *22*(3), 163–166. https://doi.org/10.1111/camh.12203

Jia, H., Zack, M. M., Thompson, W. W., Crosby, A. E., & Gottesman, I. I. (2015). Impact of depression on quality-adjusted life expectancy (QALE) directly as well as indirectly through suicide. *Social Psychiatry and Psychiatric Epidemiology*, *50*(6), 939–949. https://doi.org/10.1007/s00127-015-1019-0

Jiménez Chafey, M., Duarté Vélez, Y., & Bernal, G. (2011). Mother-daughter interactions among depressed Puerto Rican adolescents: Two case studies in CBT. *Revista Puertorriqueña de Psicología*, *22*(1), 3.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, *33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Jones, M., Hopkins, K., Kyrke-Smith, R., Davies, R., Vostanis, P., & Wolpert, M. (2013). *Current view tool: Completion guide.* London, United Kingdom: CAMHS Press.

Jonsson, U., Bohman, H., von Knorring, L., Olsson, G., Paaren, A., & von Knorring, A.-L. (2011). Mental health outcome of long-term and episodic adolescent depression: 15-year follow-up of a community sample. *Journal of Affective Disorders*, *130*(3), 395–404. https://doi.org/10.1016/j.jad.2010.10.046

Judd, L. L., Akiskal, H. S., Maser, J. D., Zeller, P. J., Endicott, J., Coryell, W., … Mueller, T. I. (1998). Major depressive disorder: a prospective study of residual subthreshold depressive symptoms as predictor of rapid relapse. *Journal of Affective Disorders*, *50*(2–3), 97–108.

Kampen, J. K., & Tamás, P. (2014). Overly ambitious: contributions and current status of Q methodology. *Quality and Quantity*, *48*(6), 3109–3126. https://doi.org/10.1007/s11135-013-9944-z

Kamphaus, R. W., & Reynolds, C. R. (2006). *PRQ: Parenting relationship questionnaire manual.* Minneapolis, MN: Pearson.

Karlsson, L., Kiviruusu, O., Miettunen, J., Heilä, H., Holi, M., Ruuttu, T., … Marttunen, M. (2008). One-year course and predictors of outcome of adolescent depression: A case-control study in Finland. *Journal of Clinical Psychiatry*, *69*(5), 844–853. https://doi.org/10.1159/000157526

Karlsson, L., Pelkonen, M., Ruuttu, T., Kiviruusu, O., Heilä, H., Holi, M., … Marttunen, M. (2006). Current comorbidity among consecutive adolescent psychiatric outpatients with DSM-IV mood disorders. *European Child and Adolescent Psychiatry*, *15*(4), 220–231. https://doi.org/10.1007/s00787-006-0526-7

Karpenko, V., & Owens, J. S. (2013). Adolescent psychotherapy outcomes in community mental health: How do symptoms align with target complaints and perceived change? *Community Mental Health Journal*, *49*(5), 540–552. https://doi.org/10.1007/s10597-012-9515-0

Kaslow, N. J., Stark, K. D., Printz, B., Livingston, R., & Tsai, S. L. (1992). Cognitive Triad Inventory for Children : Development and Relation to Depression and Anxiety. *Journal of Clinical Child Psychology*, *21*, 339–347. https://doi.org/10.1207/s15374424jccp2104

Kassis, W., Artz, S., & White, J. (2017). Understanding Depression in Adolescents: A Dynamic Psychosocial Web of Risk and Protective Factors. *Child & Youth Care Forum*, *46*(5), 721–743. https://doi.org/10.1007/s10566-017-9404-3

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., … Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American*

Academy of Child and Adolescent Psychiatry, *36*(7), 980–988. https://doi.org/10.1097/00004583-199707000-00021

Kazdin, A. E. (1999a). Current (Lack of) Status of Theory in Child and Adolescent Psychotherapy Research. *Journal of Clinical Child and Adolescent Psychology*, *28*(4), 533–543. https://doi.org/10.1207/S15374424JCCP2804_13

Kazdin, A. E. (1999b). The Meanings and Measurement of Clinical Significance. *Journal of Consulting and Clinical Psychology*, *67*(3), 332–339. https://doi.org/10.1037/0022-006X.67.3.332

Kazdin, A. E. (2006). Arbitrary Metrics: Implications for Identifying Evidence-Based Treatments. *American Psychologist*, *61*(1), 42–49. https://doi.org/10.1037/0003-066X.61.1.42

Kazdin, A. E. (2009). Understanding how and why psychotherapy leads to change. *Psychotherapy Research*, *19*(4–5), 418–428. https://doi.org/10.1080/10503300802448899

Kazdin, A. E., & Wassell, G. (2000). Therapeutic changes in children, parents, and families resulting from treatment of children with conduct problems. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*(4), 414–420. https://doi.org/10.1097/00004583-200004000-00009

Kazdin, A. E., & Weisz, J. R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology*, *66*(1), 19–36. https://doi.org/10.1037/0022-006X.66.1.19

Keller, M. B. (2003). Past, Present, and Future Directions for Defining Optimal Treatment Outcome in Depression: Remission and Beyond. *Journal of the American Medical Association*, *289*(23), 3152–3160. https://doi.org/10.1001/jama.289.23.3152

Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., Mcdonald Scott, P., & Andreasen, N. C. (1987). The Longitudinal Interval Follow-up Evaluation: A Comprehensive Method for Assessing Outcome in Prospective Longitudinal Studies. *Archives of General Psychiatry*, *44*(6), 540–548. https://doi.org/10.1001/archpsyc.1987.01800180050009

Kelvin, R., Dubicka, B., Wilkinson, P., & Goodyer, I. M. (2010). *Brief Psychosocial Intervention (BPI): A specialist clinical care treatment manual for CAMHS use.* Cambridge, United Kingdom: University of Cambridge.

Kennard, B. D., Clarke, G. N., Weersing, V. R., Asarnow, J. R., Shamseddeen, W., Porta, G., … Brent, D. A. (2009). Effective Components of TORDIA Cognitive-Behavioral Therapy for Adolescent Depression: Preliminary Findings. *Journal of Consulting and Clinical Psychology*, *77*(6), 1033–1041. https://doi.org/10.1037/a0017411

Kennard, B. D., Silva, S. G., Mayes, T. L., Rohde, P., Hughes, J. L., Vitiello, B., … March, J. S. (2009). Assessment of safety and long-term outcomes of initial treatment with Placebo in TADS. *American Journal of Psychiatry*, *166*(3), 337–344. https://doi.org/10.1176/appi.ajp.2008.08040487

Kennard, B. D., Silva, S. G., Tonev, S., Rohde, P., Hughes, J. L., Vitiello, B., … March, J. (2009). Remission and recovery in the treatment for adolescents with depression study (TADS): Acute and long-term outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(2), 186–195. https://doi.org/10.1097/CHI.0b013e31819176f9

Kennedy, N., & Foy, K. (2005). The impact of residual symptoms on outcome of major depression. *Current Psychiatry Reports*, *7*(6), 441–446.

Kessler, R. C., Avenevoli, S., & Merikangas, K. R. (2001). Mood disorders in children and adolescents: an epidemiologic perspective. *Biological Psychiatry*, *49*(12), 1002–1014. https://doi.org/10.1016/S0006-3223(01)01129-5

Kessler, R. C., & Bromet, E. J. (2013). The Epidemiology of Depression Across Cultures. *Annual Review of Public Health*, *34*(1), 119–138. https://doi.org/10.1146/annurev-publhealth-031912-114409

King, G. A., McDougall, J., Palisano, R. J., Gritzan, J., & Tucker, M. A. (1999). Goal attainment scaling: Its use in evaluating, pediatric therapy programs. *Physical and Occupational Therapy in Pediatrics*, *19*(2), 31–52. https://doi.org/10.1080/J006v19n02_03

Kiresuk, T. J., & Sherman, R. E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community Mental Health Journal*, *4*(6), 443–453. https://doi.org/10.1007/BF01530764

Kirkwood, B. R., & Sterne, J. A. C. (2003). Essential Medical Statistics. In *Medical statistics* (2nd ed.). Malden, MA: John Wiley & Sons.

Kitson, A., Marshall, A., Bassett, K., & Zeitz, K. (2013). What are the core elements of patient-centred care? A narrative review and synthesis of the literature from health policy, medicine and nursing. *Journal of Advanced Nursing*, *69*(1), 4–15. https://doi.org/10.1111/j.1365-2648.2012.06064.x

Klein, D. N., Dougherty, L. R., & Olino, T. M. (2005). Toward Guidelines for Evidence-Based Assessment of Depression in Children and Adolescents. *Journal of Clinical Child & Adolescent Psychology*, *34*(3), 412–432. https://doi.org/10.1207/s15374424jccp3403_3

Kleinert, S., & Horton, R. (2017). From universal health coverage to right care for health. *The Lancet*, *390*(10090), 101–102. https://doi.org/10.1016/S0140-6736(16)32588-0

Klerman, G. L., & Weissman, M. M. (1994). *Interpersonal Psychotherapy of Depression: A Brief, Focused, Specific Strategy*. Retrieved from https://books.google.com/books?id=ZxuqnPWGcW8C&pgis=1

Kline, P. (1994). *An easy guide to factor analysis*. https://doi.org/10.1016/0191-8869(94)90040-X

Knight, J. R., Sherritt, L., Shrier, L. A., Harris, S. K., & Chang, G. (2002). Validity of the CRAFFT substance abuse screening test among adolescent clinic patients. *Archives of Pediatrics & Adolescent Medicine*, *156*(6), 607–614. https://doi.org/10.1001/archpedi.156.6.607

Knox, S., Schlosser, L. Z., Pruitt, N. T., & Hill, C. E. (2006). A Qualitative Examination of Graduate Advising Relationships: The Advisor Perspective. *The Counseling Psychologist*, *34*(4), 489–518. https://doi.org/10.1177/0011000006290249

Koestner, R., Lekes, N., Powers, T. A., & Chicoine, E. (2002). Attaining personal goals: Self-concordance plus implementation intentions equals success. *Journal of Personality and Social Psychology*, *83*(1), 231–244. https://doi.org/10.1037/0022-3514.83.1.231

Kösters, M. P., Chinapaw, M. J. M., Zwaanswijk, M., van der Wal, M. F., & Koot, H. M. (2015). Structure, reliability, and validity of the revised child anxiety and depression scale (RCADS) in a multi-ethnic urban sample of Dutch children. *BMC Psychiatry*, *15*(1), 132. https://doi.org/10.1186/s12888-015-0509-7

Kovacs, M. (1992). *Children’s Depression Inventory (CDI)*. New York, NY: Multi-Health Systems.

Kovacs, M., Obrosky, S., & George, C. (2016). The course of major depressive disorder from childhood to young adulthood: Recovery and recurrence in a longitudinal observational study. *Journal of Affective Disorders*, Vol. 203, pp. 374–381. https://doi.org/10.1016/j.jad.2016.05.042

Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., & Kupfer, D. J. (2003). A New Approach to Integrating Data From Multiple Informants in Psychiatric Assessment and Research: Mixing and Matching Contexts and Perspectives. *The American Journal of Psychiatry*, *160*(9), 1566–1577. https://doi.org/http://dx.doi.org/10.1176/appi.ajp.160.9.1566

Kramer, T., Phillips, S. D., Hargis, M. B., Miller, T. L., Burns, B. J., & Robbins, J. M. (2004). Disagreement between parent and adolescent reports of functional impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *45*(2), 248–259. https://doi.org/10.1111/j.1469-7610.2004.00217.x

Krause, K., Chung, S., Adewuya, A. O., Albano, A. M., Babins-Wagner, R., Brann, P., … Wolpert, M. (2020). Measuring responses to clinical care in child and adolescent anxiety, depression, OCD and PTSD: an international standard set of outcome measures. *Manuscript in Preparation*.

Krause, K. R., Bear, H. A., Edbrooke-Childs, J., & Wolpert, M. (2019). Review: What Outcomes Count? A Review of Outcomes Measured for Adolescent Depression Between 2007 and 2017. *Journal of the American Academy of Child and Adolescent Psychiatry*, *58*(1), 61–71. https://doi.org/10.1016/j.jaac.2018.07.893

Kruskal, W. H., & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Lacasse, Y., Wong, E., & Guyatt, G. (1999). Individualising questionnaires. In C. R. B. Joyce, C. A. O'Boyle, & H. McGee (Eds.), *Individual quality of life. Approaches to conceptualisation and assessment* (pp. 87 – 103). Amsterdam, Netherlands: Harwood Academic Publishers.

Lachs, M. S. (1993). The more things change.... *Journal of Clinical Epidemiology*, *46*(10), 1091–1092.

Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, *50*(1), 42–51. https://doi.org/10.1037/a0030682

Landeta, J., Barrutia, J., & Lertxundi, A. (2011). Hybrid Delphi: A methodology to facilitate contribution from experts in professional contexts. *Technological Forecasting and Social Change*, *78*(9), 1629–1641. https://doi.org/10.1016/j.techfore.2011.03.009

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lavik, K. O., Veseth, M., Frøysa, H., Binder, P. E., & Moltu, C. (2018). What are "good outcomes" for adolescents in public mental health settings ? *International Journal of Mental Health Systems*, *12*(3), 1–10. https://doi.org/10.1186/s13033-018-0183-5

Law, D. (2006). *Goal Based Outcomes (GBO): some useful information.* London, England: CAMHS Press.

Law, D., & Jacob, J. (2015). *Goals and Goal Based Outcomes (GBOs): Some Useful Information* (3rd ed.). London, England: CAMHS Press.

Law, D., & Wolpert, M. (Eds.). (2014). *Guide to Using Outcomes and Feedback Tools with Children, Young People and Families* (2nd ed.). London, England: CAMHS Press.

Le Noury, J., Nardo, J. M., Healy, D., Jureidini, J., Raven, M., Tufanaru, C., & Abi-Jaoude, E. (2015). Restoring Study 329: Eficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. *BMJ (Online)*, *351*, h4320. https://doi.org/10.1136/bmj.h4320

Leamy, M., Bird, V., Le Boutillier, C., Williams, J., & Slade, M. (2011). Conceptual framework for personal recovery in mental health: Systematic review and narrative synthesis. *British*

Journal of Psychiatry, *199*(6), 445–452. https://doi.org/10.1192/bjp.bp.110.083733

Lewandowski, R. E., Acri, M. C., Hoagwood, K. E., Olfson, M., Clarke, G., Gardner, W., … Horwitz, S. M. (2013). Evidence for the management of adolescent depression. *Pediatrics*, *132*(4). https://doi.org/10.1542/peds.2013-0600

Lewis, C. C., Simons, A. D., Nguyen, L. J., Murakami, J. L., Reid, M. W., Silva, S. G., & March, J. S. (2010). Impact of Childhood Trauma on Treatment Outcome in the Treatment for Adolescents with Depression Study (TADS). *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(2), 132–140. https://doi.org/10.1097/00004583-201002000-00007

Lewis, C. C., Simons, A. D., Silva, S. G., Rohde, P., Small, D. M., Murakami, J. L., … March, J. S. (2009). The Role of Readiness to Change in Response to Treatment of Adolescent Depression. *Journal of Consulting and Clinical Psychology*, *77*(3), 422–428. https://doi.org/10.1037/a0014154

Leyfer, O. T., Ruberg, J. L., & Woodruff-Borden, J. (2006). Examination of the utility of the Beck Anxiety Inventory and its factors as a screener for anxiety disorders. *Journal of Anxiety Disorders*, *20*(4), 444–458. https://doi.org/10.1016/j.janxdis.2005.05.004

Lin, E. H. B., Katon, W. J., VonKorff, M., Russo, J. E., Simon, G. E., Bush, T. M., … Ludman, E. (1998). Relapse of Depression in Primary Care: Rate and Clinical Predictors. *Archives of Family Medicine*, *7*(5), 443–449. https://doi.org/10.1001/archfami.7.5.443

Linstone, H. A., & Turoff, M. (1975). *The Delphi Method: Techniques and Applications.* Reading, MA: Adison Wesley.

Linstone, H. A., & Turoff, M. (2011). Delphi: A brief look backward and forward. *Technological Forecasting and Social Change*, *78*(9), 1712–1719. https://doi.org/10.1016/j.techfore.2010.09.011

Liverpool, S., Webber, H., Matthews, R., Wolpert, M., Edbrooke-Childs, J., & Fonagy, P. (2019). *A Mobile App to Support Parents Making Child Mental Health Decisions: A Feasibility Cluster Randomized Controlled Trial [Unpublished manuscript]*. London, England: University College London.

Loeber, R., Green, S., & Lahey, B. (1990). Mental Health Professionals’ Perception of the Utility of Children, Mothers, and Teachers as Informants on Childhood Psychopathology. *Journal of Clinical Child & Adolescent Psychology*, *19*(2), 136–143. https://doi.org/10.1207/s15374424jccp1902_5

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

Lusk, P., & Melnyk, B. M. (2011). The brief cognitive-behavioral COPE intervention for depressed adolescents: Outcomes and feasibility of delivery in 30-minute outpatient visits. *Journal of the American Psychiatric Nurses Association*, *17*(3), 226–236. https://doi.org/10.1177/1078390311404067

Lynch, F. L., Dickerson, J. F., Clarke, G., Vitiello, B., Porta, G., Wagner, K. D., … Brent, D. (2011). Incremental cost-effectiveness of combined therapy vs medication only for youth with selective serotonin reuptake inhibitor-resistant depression: Treatment of SSRI-resistant depression in adolescents trial findings. *Archives of General Psychiatry*, *68*(3), 253–262. https://doi.org/10.1001/archgenpsychiatry.2011.9

Makol, B. A., De Los Reyes, A., Ostrander, R. S., & Reynolds, E. K. (2019). Parent-Youth Divergence (and Convergence) in Reports of Youth Internalizing Problems in Psychiatric Inpatient Care. *Journal of Abnormal Child Psychology*. https://doi.org/10.1007/s10802-019-00540-7

Makol, B. A., & Polo, A. J. (2018). Parent-child endorsement discrepancies among youth at chronic-risk for depression. *Journal of Abnormal Child Psychology*, *46*(5), 1077–1088. https://doi.org/http://dx.doi.org/10.1007/s10802-017-0360-z

Manglick, M., Rajaratnam, S. M., Taffe, J., Tonge, B., & Melvin, G. (2013). Persistent sleep disturbance is associated with treatment response in adolescents with depression. *Australian and New Zealand Journal of Psychiatry*, *47*(6), 556–563. https://doi.org/10.1177/0004867413481630

Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, *18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Manos, R. C., Kanter, J. W., & Luo, W. (2011). The Behavioral Activation for Depression Scale–Short Form: Development and Validation. *Behavior Therapy*, *42*(4), 726–739. https://doi.org/10.1016/j.beth.2011.04.004

March, J. S., Silva, S., Petrycki, S., Curry, J., Wells, K., Fairbank, J., … Rochon, J. (2007). The Treatment for Adolescents with Depression Study (TADS): Long-term effectiveness and safety outcomes. *Archives of General Psychiatry*, *64*(10), 1132–1144. https://doi.org/10.1001/archpsyc.64.10.1132

March, J., Silva, S., Curry, J., Wells, K., Fairbank, J., Burns, B., … Bartoi, M. (2009). The treatment for adolescents with depression study (TADS): Outcomes over 1 year of naturalistic follow-up. *American Journal of Psychiatry*, *166*(10), 1141–1149. https://doi.org/10.1176/appi.ajp.2009.08111620

March, J., Silva, S., Petrycki, S., Curry, J., Wells, K., Fairbank, J., … Severe, J. (2004). Fluoxetine, Cognitive-Behavioral Therapy, and Their Combination for Adolescents With

Depression: Treatment for Adolescents With Depression Study (TADS) Randomized Controlled Trial. *JAMA*, *292*(7), 807–820. https://doi.org/10.1001/jama.292.7.807

Mason, M. (2010). Sample size and saturation in PhD studies using qualitative interviews. *Forum Qualitative Sozialforschung*, *11*(3). https://doi.org/http://dx.doi.org/10.17169/fqs-11.3.1428

Massey, E. K., Gebhardt, W. A., & Garnefski, N. (2008). Adolescent goal content and pursuit: A review of the literature from the past 16 years. *Developmental Review*, *28*(4), 421–460. https://doi.org/10.1016/j.dr.2008.03.002

Massey, E. K., Gebhardt, W. A., & Garnefski, N. (2009). Self-generated goals and goal process appraisals : Relationships with sociodemographic factors and well-being. *Journal of Adolescence*, *32*(3), 501–518. https://doi.org/10.1016/j.adolescence.2008.07.003

Maxcy, S. J. (2003). Pragmatic Threads in Mixed Methods Research in the Social Sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social & Behavioral Research* (pp. 51–89). Thousand Oaks, CA: SAGE Publications.

Mayes, T. L., Tao, R., Rintelmann, J. W., Carmody, T., Hughes, C. W., Kennard, B. D., … Emslie, G. J. (2007). Do children and adolescents have differential response rates in placebo-controlled trials of fluoxetine?. *CNS Spectrums*, *12*(2), 147–154.

McArthur, K., Cooper, M., & Berdondini, L. (2013). School-based humanistic counseling for psychological distress in young people: Pilot randomized controlled trial. *Psychotherapy Research*, *23*(3), 355–365. https://doi.org/10.1080/10503307.2012.726750

McGlinchey, J. B., Zimmerman, M., & Atkins, D. C. (2008). Clinical significance and remission in treating major depressive disorder: Parallels between related outcome constructs. *Harvard Review of Psychiatry*, *16*(1), 25–34. https://doi.org/10.1080/10673220701885815

McHugh, M. L. (2012). Lessons in biostatistics Interrater reliability : the kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

McKnight, P. E., & Kashdan, T. B. (2009). The importance of functional impairment to mental health outcomes: A case for reassessing our goals in depression treatment research. *Clinical Psychology Review*, *29*(3), 243–259. https://doi.org/10.1016/j.cpr.2009.01.005

McLeod, J. (2001). Introduction: Research into the client’s experience of therapy. *Counselling and Psychotherapy Research*, *1*(1), 41–41. https://doi.org/10.1080/14733140112331385238

McLeod, J. (2011). *Qualitative Research in Counselling and Psychotherapy* (2nd ed.).

https://doi.org/10.4135/9781849209663

McMakin, D. L., Olino, T. M., Porta, G., Dietz, L. J., Emslie, G., Clarke, G., … Brent, D. A. (2012). Anhedonia predicts poorer recovery among youth with selective serotonin reuptake inhibitor treatmentresistant depression. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(4), 404–411. https://doi.org/10.1016/j.jaac.2012.01.011

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, *12*(2), 153–157.

Measure Applications Partnership. (2017). *Strengthening the Core Set of Healthcare Quality Measures for Adults Enrolled in Medicaid, 2017*. Retrieved from file:///C:/Users/KKrause/Downloads/map_medicaid_child_draft_report.pdf

Melnyk, B. M., Jacobson, D., Kelly, S., O'Haver, J., Small, L., & Mays, M. Z. (2009). Improving the Mental Health, Healthy Lifestyle Choices, and Physical Health of Hispanic Adolescents: A Randomized Controlled Pilot Study. *Journal of School Health*, *79*(12), 575–584. https://doi.org/10.1111/j.1746-1561.2009.00451.x

Melvin, G. A., Dudley, A. L., Gordon, M. S., Ford, S., Taffe, J., & Tonge, B. J. (2013). What happens to depressed adolescents? A follow-up study into early adulthood. *Journal of Affective Disorders*, *151*(1), 298–305. https://doi.org/10.1016/j.jad.2013.06.012

Melvin, G. A., Tonge, B. J., King, N. J., Heyne, D., Gordon, M. S., & Klimkeit, E. (2006). A comparison of cognitive-behavioral therapy, sertraline, and their combination for adolescent depression. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*(10), 1151–1161. https://doi.org/10.1097/01.chi.0000233157.21925.71

Merry, S. N., Stasiak, K., Shepherd, M., Frampton, C., Fleming, T., & Lucassen, M. F. G. (2012). The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: Randomised controlled non-inferiority trial. *BMJ (Online)*, *344*(7857), e2598–e2598. https://doi.org/10.1136/bmj.e2598

Merry, S., Stasiak, K., Parkin, A., Seymour, F., Lambie, I., Crengle, S., & Pasene-Mizziebo, E. (2004). *Child and Youth Outcome Measures: Examining current use and acceptability of measures in mental health services and recommending future directions*. Retrieved from https://www.tepou.co.nz/uploads/files/resource-assets/Child-and-Youth-Outcome-Measures.pdf

Mertens, D. M. (2007). Transformative Paradigm. *Journal of Mixed Methods Research*, *1*(3), 212–225. https://doi.org/10.1177/1558689807302811

Mertens, D. M. (2010). Transformative mixed methods research. *Qualitative Inquiry*, *16*(6), 469–474. https://doi.org/10.1177/1077800410364612

Micco, J. A., Henin, A., & Hirshfeld-Becker, D. R. (2014). Efficacy of interpretation bias modification in depressed adolescents and young adults. *Cognitive Therapy and Research*, *38*(2), 89–102. https://doi.org/10.1007/s10608-013-9578-4

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., … Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, *46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(1). https://doi.org/10.1186/1748-5908-6-42

Midgley, N., Ansaldo, F., Parkinson, S., Holmes, J., Stapley, E., & Target, M. (2011). *Experience of Therapy Interview (Young Person, Parent and Therapist Versions) [Unpublished manuscript]*. London: Anna Freud Centre.

Midgley, N., Ansaldo, F., & Target, M. (2014). The meaningful assessment of therapy outcomes: Incorporating a qualitative study into a randomized controlled trial evaluating the treatment of adolescent depression. *Psychotherapy*, *51*(1), 128–137.

Midgley, N., O'Keeffe, S., French, L., & Kennedy, E. (2017). Psychodynamic psychotherapy for children and adolescents: an updated narrative review of the evidence base. *Journal of Child Psychotherapy*, *43*(3), 307–329. https://doi.org/10.1080/0075417X.2017.1323945

Midgley, N., Parkinson, S., Holmes, J., Stapley, E., Eatough, V., & Target, M. (2015). Beyond a diagnosis: The experience of depression among clinically-referred adolescents. *Journal of Adolescence*, *44*, 269–279. https://doi.org/10.1016/j.adolescence.2015.08.007

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The Outcome Rating Scale: A Preliminary Study of the Reliability, Validity, and Feasibility of a Brief Visual Analog Measure. *Journal of Brief Therapy*, *2*(2), 91–100.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and MetaAnalyses: The PRISMA Statement. *PLOS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed1000097

Mollon, J., David, A. S., Morgan, C., Frissa, S., Glahn, D., Pilecka, I., … Reichenberg, A. (2016). Psychotic experiences and neuropsychological functioning in a population-based sample. *JAMA Psychiatry*, *73*(2), 129–138. https://doi.org/10.1001/jamapsychiatry.2015.2551

Monga, S., Butcher, N., Monsour, A., Mew, E., Szatmari, P., & Offringa, M. (2019). Core set of outcomes for adolescents with major depressive disorder: A tool of standardized outcomes for clinical research. Retrieved December 10, 2019, from http://www.comet-initiative.org/studies/details/1122

Montgomery, S. A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry*, *134*(4), 382–389. https://doi.org/10.1192/bjp.134.4.382

Moran, P., Kelesidi, K., Guglani, S., Davidson, S., & Ford, T. (2012). What do parents and carers think about routine outcome measures and their use? A focus group study of CAMHS attenders. *Clinical Child Psychology and Psychiatry*, *17*(1), 65–79. https://doi.org/10.1177/1359104510391859

Morgan, D. L. (2007). Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research*, *1*(1), 48–76. https://doi.org/10.1177/2345678906292462

Morin, C. M. (1993). *Insomnia: Psychological Assessment and Management*. New York, NY: Guilford Press.

Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *Journal of Counseling Psychology*, *52*(2), 250–260. https://doi.org/10.1037/0022-0167.52.2.250

Morse, J. M. (1995). The significance of saturation. *Qualitative Health Research*, Vol. 5, pp. 147–149. https://doi.org/10.1177/104973239500500201

Mufson, L., Yanes-Lukin, P., Gunlicks-Stoessel, M., & Wickramaratne, P. (2014). Cultural Competency and Its Effect on Treatment Outcome of IPT-A in School-Based Health Clinics. *American Journal of Psychotherapy*, *68*(4), 417–442.

Mulley, A., Coulter, A., Wolpert, M., Richards, T., & Abbasi, K. (2017). New approaches to measurement and management for high integrity health systems. *BMJ*, *356*, j1401. https://doi.org/10.1136/bmj.j1401

Mullick, M. S. I., & Goodman, R. (2001). Questionnaire screening for mental health problems in Bangladeshi children: A preliminary study. *Social Psychiatry and Psychiatric Epidemiology*, *36*(2), 94–99. https://doi.org/10.1007/s001270050295

Muris, P., Meesters, C., & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ): Further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child and Adolescent Psychiatry & Adolescent Psychiatry*, *12*(1), 1–8.

Napoleone, E., Evans, C., Patalay, P., Edbrooke-Childs, J., & Wolpert, M. (2019). Trajectories of change of youth depressive symptoms in routine care: shape, predictors, and service-use implications. *European Child and Adolescent Psychiatry*, (0123456789). https://doi.org/10.1007/s00787-019-01317-5

National Committee for Quality Assurance (NCQA). (2018). National Collaborative for Innovation in Quality Measurement: Implementing and Improving. Retrieved July 3, 2018, from http://www.ncqa.org/hedis-quality-measurement/research/ncinq

National Health Service. (2018). Governance arrangements for research ethics committees: 2018 edition. Retrieved June 11, 2019, from https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/governance-arrangement-research-ethics-committees/

National Institute for Health and Care Excellence. (2011). *Service user experience in adult mental health: Improving the experience of care for people using adult NHS Mental Health Services*. Retrieved from https://www.nice.org.uk/guidance/cg136/resources/service-user-experience-in-adult-mental-health-improving-the-experience-of-care-for-people-using-adult-nhs-mental-health-services-35109513728197

National Quality Forum. (2017). *Shared decision making: A standard of care for all patients. A National Quality Partners ™ Action Brief*. Retrieved from http://www.qualityforum.org/Publications/2017/10/NQP_Shared_Decision_Making_Action_Brief.aspx

Newman, I., Ridenour, C., Newman, C., & DeMarco, G. M. P. J. (2003). A typology of research purposes and its relationship to mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social & Behavioral Research* (pp. 167–188). Thousand Oaks, CA: Sage.

Ngo, V. K., Asarnow, J. R., Lange, J., Jaycox, L. H., Rea, M. M., Landon, C., … Miranda, J. (2009). Outcomes for Youths From Racial-Ethnic Minority Groups in a Quality Improvement Intervention for Depression Treatment. *Psychiatric Services*, *60*(10), 1357–1364. https://doi.org/10.1176/ps.2009.60.10.1357

Niederhofer, H., & von Klitzing, K. (2011). Therapy augmentation by bright light treatment for non-seasonal depression of adolescents. *Clinical Neuropsychiatry*, *8*(3), 225–227.

Niederhofer, H., & Von Klitzing, K. (2012). Bright light treatment as mono-therapy of non-seasonal depression for 28 adolescents. *International Journal of Psychiatry in Clinical Practice*, *16*(3), 233–237. https://doi.org/10.3109/13651501.2011.625123

Nierenberg, A. A., & Wright, E. C. (1999). Evolution of remission as the new standard in the treatment of depression. *Journal of Clinical Psychiatry*, *60*(Suppl. 22), 7–11.

NIMH. (2008). National Institute of Mental Health Strategic Plan. Retrieved February 27, 2020, from https://www.naminys.org/images/uploads/pdfs/Strategic Plan Mailer.pdf

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depres- sion and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta earthquake. *Journal of Personality and Social Psychology*, *61*, 115–121.

Norton, C. L. (2010). Into the wilderness-a case study: The psychodynamics of adolescent depression and the need for a holistic intervention. *Clinical Social Work Journal*, *38*(2), 226–235. https://doi.org/10.1007/s10615-009-0205-5

O'Keeffe, S., Martin, P., Target, M., & Midgley, N. (2019). "I just stopped going": A Mixed Methods Investigation Into Types of Therapy Dropout in Adolescents With Depression. *Frontiers in Psychology*, *10*(FEB), 1–14. https://doi.org/10.3389/fpsyg.2019.00075

O'Shea, G., Spence, S. H., & Donovan, C. L. (2015). Group versus Individual Interpersonal Psychotherapy for Depressed Adolescents. *Behavioural and Cognitive Psychotherapy*, *43*(01), 1–19. https://doi.org/10.1017/S1352465814000216

Oakley, A. (1999). Paradigm wars: Some thoughts on a personal and public trajectory. *International Journal of Social Research Methodology*, *2*(3), 247–254. https://doi.org/10.1080/136455799295041

Odhammar, F., & Carlberg, G. (2015). Parents' and psychotherapists' goals prior to psychodynamic child psychotherapy. *European Journal of Psychotherapy and Counselling*, *17*(3), 277–295. https://doi.org/10.1080/13642537.2015.1059865

Office for National Statistics. (2006). *Review of the Dissemination of Health Statistics: Confidentiality Guidance*. Retrieved from https://www.ons.gov.uk/file?uri=/methodology/methodologytopicsandstatisticalconcepts/disclosurecontrol/healthstatistics/confidentialityguidanctcm77181864.pdf

Owens, L. (2016). The use of Q sort methodology in research with teenagers. In J. Prior & J. Van Herwegen (Eds.), *Practical Research with Children* (pp. 228–248). London, England: Routledge.

Pan, A., Sun, Q., Okereke, O. I., Rexrode, K. M., & Hu, F. B. (2011). Depression and Risk of Stroke Morbidity and Mortality: A Meta-analysis and Systematic Review. *JAMA*, *306*(11), 1241–1249. https://doi.org/10.1001/jama.2011.1282

Park, R. J., & Goodyer, I. M. (2000). Clinical guidelines for depressive disorders in childhood and adolescence. *European Child and Adolescent Psychiatry*, *9*(3), 147–161.

Parker, A. G., Hetrick, S. E., Jorm, A. F., Mackinnon, A. J., McGorry, P. D., Yung, A. R., … Purcell, R. (2016). The effectiveness of simple psychological and physical activity interventions for high prevalence mental health problems in young people: A factorial randomised controlled trial. *Journal of Affective Disorders*, *196*, 200–209. https://doi.org/10.1016/j.jad.2016.02.043

Pass, L., Brisco, G., & Reynolds, S. (2015). Adapting brief Behavioural Activation (BA) for adolescent depression: a case example. *The Cognitive Behaviour Therapist*, *8*, e17. https://doi.org/10.1017/S1754470X15000446

Patalay, P., & Fitzsimons, E. (2016). Correlates of Mental Illness and Wellbeing in Children: Are They the Same? Results From the UK Millennium Cohort Study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *55*(9), 771–783. https://doi.org/10.1016/j.jaac.2016.05.019

Patalay, P., Fonagy, P., Deighton, J., Belsky, J., Vostanis, P., & Wolpert, M. (2015). A general psychopathology factor in early adolescence. *British Journal of Psychiatry*, *207*(1), 15–22. https://doi.org/10.1192/bjp.bp.114.149591

Paterson, C. (1996). Measuring outcomes in primary care: A patient generated measure, MYMOP, compared with the SF-36 health survey. *British Medical Journal*, *312*(7037), 1016–1020. https://doi.org/10.1136/bmj.312.7037.1016

Patrick, D. L., & Deyo, R. A. (1989). Generic and Disease-Specific Measures in Assessing Health Status and Quality of Life. *Medical Care*, *27*(3 Suppl), S217–S232.

Patton, G. C., Coffey, C., Romaniuk, H., Mackinnon, A., Carlin, J. B., Degenhardt, L., … Moran, P. (2014). The prognosis of common mental disorders in adolescents: A 14-year prospective cohort study. *The Lancet*, *383*(9926), 1404–1411. https://doi.org/10.1016/S0140-6736(13)62116-9

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768–774. https://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

Peters, A. T., Jacobs, R. H., Feldhaus, C., Henry, D. B., Albano, A. M., Langenecker, S. A., … Curry, J. F. (2016). Trajectories of Functioning into Emerging Adulthood Following Treatment for Adolescent Depression. *Journal of Adolescent Health*, *58*(3), 253–259. https://doi.org/10.1016/j.jadohealth.2015.09.022

Picker Institute. (2004). *Patient-Centered Care 2015: Scenarios, Vision, Goals & Next Steps*. Retrieved from Picker Institute website: http://www.altfutures.org/pubs/health/Picker Final Report May 14 2004.pdf

Pickles, A., Rowe, R., Simonoff, E., Foley, D., Rutter, M., & Silberg, J. (2001). Child psychiatric symptoms and psychosocial impairment: Relationship and prognostic significance. *British Journal of Psychiatry*, *179*(SEPT.), 230–235. https://doi.org/10.1192/bjp.179.3.230

Piers, E. V., Harris, D. B., & Herzberg, D. S. (1969). *Piers-Harris 2. Piers-Harris Children’s Self- Concept Scale (2nd ed.).* Los Angels, CA: Western Psychological Services.

Pilgrim, D., & McCranie, A. (2013). *Recovery and Mental Health: A Critical Sociological Account*. New York, NY: Palgrave Macmillan.

Pollard, E. L., & Lee, P. D. (2003). Child well-being: A systematic review of the literature. *Social Indicators Research*, *61*(1), 59–78.

Poole, L. A., Knight, T., Toumbourou, J. W., Lubman, D. I., Bertino, M. D., & Lewis, A. J. (2017). A Randomized Controlled Trial of the Impact of a Family-Based Adolescent Depression Intervention on both Youth and Parent Mental Health Outcomes. *Journal of Abnormal Child Psychology*, *46*(1), 1–13. https://doi.org/10.1007/s10802-017-0292-7

Pope, C., & Mays, N. (1995). Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ: British Medical Journal*, *311*(July), 42–45.

Porter, M. E. (2010). What Is Value in Health Care? *New England Journal of Medicine*, *363*(26), 2477–2481. https://doi.org/10.1056/NEJMp1011024

Posner, K., Brown, G. K., Stanley, B., Brent, D. A., Yershova, K. V., Oquendo, M. A., … Mann, J. J. (2011). The Columbia-suicide severity rating scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. *American Journal of Psychiatry*, *168*(12), 1266–1277. https://doi.org/10.1176/appi.ajp.2011.10111704

Potter, J., & Wetherell, M. (1987). *Discourse and social psychology: Beyond attitudes and behaviour*. London, England: Sage.

Powell, C. (2003). The Delphi Technique: myths and realities. *Methodological Issues in Nursing Research*, *41*(4), 376–382. https://doi.org/10.1046/j.1365-2648.2003.02537.x

Poznanski, E., & Mokros, H. (1996). *Children’s Depression Rating Scale–Revised (CDRS-R)*. Los Angeles, CA: Western Psychological Services.

Pybis, J., Cooper, M., Hill, A., Cromarty, K., Levesley, R., Murdoch, J., & Turner, N. (2015). Pilot randomised controlled trial of school-based humanistic counselling for psychological distress in young people: Outcomes and methodological reflections. *Counselling and Psychotherapy Research*, *15*(4), 241–250. https://doi.org/10.1002/capr.12009

Radloff, L. S. (1977). The CES-D Scale. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Rengasamy, M., Mansoor, B. M., Hilton, R., Porta, G., He, J., Emslie, G. J., … Brent, D. A. (2013). The bi-directional relationship between parent-child conflict and treatment outcome in treatment-resistant adolescent depression. *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*(4), 370–377.

https://doi.org/10.1016/j.jaac.2013.01.012

Reynolds, C. R. (2010). Behavior Assessment System for Children. In I. B. Weiner & C. W. Edward (Eds.), *The Corsini Encyclopedia of Psychology*. https://doi.org/10.1002/9780470479216.corpsy0114

Reynolds, C. R., & Richmond, B. O. (1978). What I think and feel: A revised measure of children's manifest anxiety. *Journal of Abnormal Child Psychology*, *6*(2), 271–280. https://doi.org/10.1023/A:1025751206600

Reynolds, W. M. (1987). *Professional Manual for the Reynolds Adolescent Depression Scale (RADS)*. Odessa, FL: Psychological Assessment Resources.

Reynolds, W. M. (1988). *Suicidal Ideation Questionnaire. Professional Manual.* Odessa, FL: Psychological Assessment Resources.

Riley, K. (2011). Effects of a Collaborative Outpatient Therapy Program on Self-Concept of Adolescents with Depression. *Therapeutic Recreation Journal*, *45*(1), 32–46. https://doi.org/10.1192/bjp.188.3.216

Robin, A. L., & Foster, S. L. (1989). *Negotiating parent–adolescent conflict: A behavioral-family systems approach.* New York, NY: Guilford Press.

Robinson, O. C. (2014). Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide. *Qualitative Research in Psychology*, *11*(1), 25–41. https://doi.org/10.1080/14780887.2013.801543

Rockwood, K., Joyce, B., & Stolee, P. (1997). Use of goal attainment scaling in measuring clinically important change in cognitive rehabilitation patients. *Journal of Clinical Epidemiology*, *50*(5), 581–588. https://doi.org/10.1016/S0895-4356(97)00014-0

Rogers, G. M., Park, J.-H., Essex, M. J., Klein, M. H., Silva, S. G., Hoyle, R. H., … March, J. S. (2009). The Dysfunctional Attitudes scale: Psychometric properties in depressed adolescents. *Journal of Clinical Child and Adolescent Psychology*, *38*(6), 781–789. https://doi.org/10.1080/15374410903259007

Rohde, P., Silva, S. G., Tonev, S. T., Kennard, B. D., Vitiello, B., Kratochvil, C. J., … March, J. S. (2008). Achievement and maintenance of sustained response during the treatment for adolescents with depression study continuation and maintenance therapy. *Archives of General Psychiatry*, *65*(4), 447–455. https://doi.org/10.1001/archpsyc.65.4.447

Rohde, P., Waldron, H. B., Turner, C. W., Brody, J., & Jorgensen, J. (2014). Sequenced versus coordinated treatment for adolescents with comorbid depressive and substance use disorders. *Journal of Consulting and Clinical Psychology*, *82*(2), 342–348. https://doi.org/10.1037/a0035808

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton

University Press.

Rosselló, J., Bernal, G., & Rivera-Medina, C. (2008). Individual and Group CBT and IPT for Puerto Rican Adolescents With Depressive Symptoms. *Cultural Diversity and Ethnic Minority Psychology*, *14*(3), 234–245. https://doi.org/10.1037/1099-9809.14.3.234

Rowe, G., & Wright, G. (2011). The Delphi technique: Past, present, and future prospects - Introduction to the special issue. *Technological Forecasting and Social Change*, *78*(9), 1487–1490. https://doi.org/10.1016/j.techfore.2011.09.002

Rudolph, K. D., & Klein, D. N. (2009). Exploring depressive personality traits in youth: Origins, correlates, and developmental consequences. *Development and Psychopathology*, *21*(4), 1155–1180. https://doi.org/10.1017/S0954579409990095

Rupani, P., Cooper, M., McArthur, K., Pybis, J., Cromarty, K., Hill, A., … Turner, N. (2014). The goals of young people in school-based counselling and their achievement of these goals. *Counselling and Psychotherapy Research*, *14*(4), 306–314. https://doi.org/10.1080/14733145.2013.816758

Rush, A. J., Crismon, M. L., Toprac, M. G., Trivedi, M. H., Rago, W. V., Shon, S., & Altshuler, K. Z. (1998). Consensus guidelines in the treatment of major depressive disorder. *Journal of Clinical Psychiatry*, *59*(SUPPL. 20), 73–84.

Rush, A. J., First, M. B., & Blacker, D. (2008). *Handbook of psychiatric measures* (2nd ed.). Washington, DC: American Psychiatric Association Publishing.

Rutledge, T., Reis, V. A., Linke, S. E., Greenberg, B. H., & Mills, P. J. (2006). Depression in heart failure: a meta-analytic review of prevalence, intervention effects, and associations with clinical outcomes. *Journal of the American College of Cardiology*, *48*(8), 1527–1537.

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, *312*(7023), 71–72. https://doi.org/10.1136/bmj.312.7023.71

Sakolsky, D. J., Perel, J. M., Emslie, G. J., Clarke, G. N., Wagner, K. D., Vitiello, B., … Brent, D. A. (2011). Antidepressant exposure as a predictor of clinical outcomes in the Treatment of Resistant Depression in Adolescents (TORDIA) study. *Journal of Clinical Psychopharmacology*, *31*(1), 92–97. https://doi.org/10.1097/JCP.0b013e318204b117

Salbach-Andrae, H., Klinkowski, N., Lenz, K., & Lehmkuhl, U. (2009). Agreement between youth-reported and parent-reported psychopathology in a referred sample. *European Child & Adolescent Psychiatry*, *18*(3), 136–143. https://doi.org/10.1007/s00787-008-0710-z

Sales, C. M. D. (2017). Seeing the person in the patient: Making the case for individualized PROMs in mental health care. *Current Psychiatry Reviews*, *13*(3), 184–187.

https://doi.org/http://dx.doi.org/10.2174/1573400513666170505111736

Sales, C. M. D., & Alves, P. C. G. (2012). Individualized patient-progress systems: Why we need to move towards a personalized evaluation of psychological treatments. *Canadian Psychology*, *53*(2), 115–121. https://doi.org/10.1037/a0028053

Sales, C. M. D., & Alves, P. C. G. (2016). Patient-Centered Assessment in Psychotherapy: A Review of Individualized Tools. *Clinical Psychology: Science and Practice*, *23*(3), 265–283. https://doi.org/10.1111/cpsp.12162

Sales, C. M. D., Alves, P. C. G., Evans, C., Elliott, R., Almeida, S., Ashworth, M., … Tavares, S. (2014). The Individualised Patient-Progress System: A decade of international collaborative networking. *Counselling and Psychotherapy Research*, *14*(3), 181–191. https://doi.org/10.1080/14733145.2014.929417

Sales, C. M. D., Neves, I. T. D., Alves, P. G., & Ashworth, M. (2018). Capturing and missing the patient’s story through outcome measures: A thematic comparison of patient-generated items in PSYCHLOPS with CORE-OM and PHQ-9. *Health Expectations*, *21*(3), 615–619. https://doi.org/10.1111/hex.12652

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption. *Addiction*, *88*(6), 791–804.

Schirman, S., Kronenberg, S., Apter, A., Brent, D., Melhem, N., Pick, N., … Gothelf, D. (2010). Effectiveness and tolerability of citalopram for the treatment of depression and anxiety disorders in children and adolescents: An open-label study. *Journal of Neural Transmission*, *117*(1), 139–145. https://doi.org/10.1007/s00702-009-0330-x

Schmolck, P. (2008, October). *Common and specific approaches in the analysis of Q-sort data with PQMethod. Keynote speech presented at the 24th Annual Q Conference*. Hamilton, Ontario.

Schmolck, P. (2014). *PQMethod (Version 2.35) [Computer Software]*. Retrieved from http://schmolck.org/qmethod/

Schrank, B., & Slade, M. (2007). Recovery in psychiatry. *Psychiatric Bulletin*, *31*, 321–325. https://doi.org/10.1192/pb.bp.10

Scott, J., Teasdale, J. D., Paykel, E. S., Johnson, A. L., Abbott, R., Hayhurst, H., … Garland, A. (2000). Effects of cognitive therapy on psychological symptoms and social functioning in residual. *British Journal of Psychiatry*, *177*(NOV.), 440–446. https://doi.org/10.1192/bjp.177.5.440

Sechrest, L., McKnight, P. E., & McKnight, K. (1996). Calibration of measures for psychotherapy outcome studies. *American Psychologist*, *51*, 1065–1071. https://doi.org/10.1007/978-1-4614-6435-8_452-3

Secker, J., Membrey, H., Grove, B., & Seebohm, P. (2002). Recovering from Illness or Recovering your Life? Implications of Clinical Versus Social Models of Recovery from Mental Health Problems for Employment Support Services. *Disability & Society*, *17*(4), 403–418. https://doi.org/10.1080/09687590220140340

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC- IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*(1), 28–38. https://doi.org/10.1097/00004583-200001000-00014

Shamseddeen, W., Asarnow, J. R., Clarke, G., Vitiello, B., Wagner, K. D., Birmaher, B., … Brent, D. A. (2011). Impact of physical and sexual abuse on treatment response in the treatment of resistant depression in adolescent study (TORDIA). *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(3), 293–301. https://doi.org/10.1016/j.jaac.2010.11.019

Shamseddeen, W., Clarke, G., Keller, M. B., Wagner, K. D., Birmaher, B., Emslie, G. J., … Brent, D. A. (2012). Adjunctive Sleep Medications and Depression Outcome in the Treatment of Serotonin-Selective Reuptake Inhibitor Resistant Depression in Adolescents Study. *Journal of Child and Adolescent Psychopharmacology*, *22*(1), 29–36. https://doi.org/10.1089/cap.2011.0027

Shamseddeen, W., Clarke, G., Wagner, K. D., Ryan, N. D., Birmaher, B., Emslie, G., … Brent, D. A. (2011). Treatment-resistant depressed youth show a higher response rate if treatment ends during summer school break. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(11), 1140–1148. https://doi.org/10.1016/j.jaac.2011.07.022

Shapiro, J. P., Welker, C. J., & Jacobson, B. J. (1997). The Youth Client Satisfaction Questionnaire: development, construct validation, and factor structure. *Journal of Clinical Child Psychology*, *26*(1), 87–98. https://doi.org/10.1207/s15374424jccp2601_9

Sharples, E., Qin, C., Goveas, V., Gondek, D., Deighton, J., Wolpert, M., & Edbrooke-Childs, J. (2017). A qualitative exploration of attitudes towards the use of outcome measures in child and adolescent mental health services. *Clinical Child Psychology and Psychiatry*, *22*(2), 219–228. https://doi.org/10.1177/1359104516652929

Shields, C. G., Franks, P., Harp, J. J., McDaniel, S. H., & Campbell, T. L. (1992). Development of the Family Emotional Involvement and Criticism Scale (Feics) - a Self-Report Scale to Measure Expressed Emotion. *Journal of Marital and Family Therapy*, *18*(4), 395–407. https://doi.org/10.1111/j.1752-0606.1992.tb00953.x

Shirk, S. R., Gudmundsen, G., Kaplinski, H. C., & McMakin, D. L. (2008). Alliance and outcome in cognitive-behavioral therapy for adolescent depression. *Journal of Clinical Child and Adolescent Psychology*, *37*(3), 631–639. https://doi.org/10.1080/15374410802148061

Shirk, S. R., & Saiz, C. C. (1992). Clinical, empirical, and developmental perspectives on the therapeutic relationship in child psychotherapy. *Development and Psychopathology*, *4*(4), 713–728. https://doi.org/10.1017/S0954579400004946

Shon, S.-H., Joo, Y., Lee, J.-S., & Kim, H.-W. (2014). Lamotrigine Treatment of Adolescents with Unipolar and Bipolar Depression: A Retrospective Chart Review. *Journal of Child and Adolescent Psychopharmacology*, *24*(5), 285–287. https://doi.org/10.1089/cap.2013.0121

Sigafoos, A. D., Feinstein, C. B., Damond, M., & Reiss, D. (1988). The measurement of behavioral autonomy in adolescence: the Autonomous Functioning Checklist. *Adolescent Psychiatry*, *15*, 432–462.

Simonoff, E., Pickles, A., Meyer, J. M., Silberg, J. L., Maes, H. H., Loeber, R., … Eaves, L. J. (1997). The Virginia Twin Study of Adolescent Behavioral Development: Influences of Age, Sex, and Impairment on Rates of Disorder. *JAMA Psychiatry*, *54*(9), 801–808. https://doi.org/10.1001/archpsyc.1997.01830210039004

Simons, A. D., Marti, C. N., Rohde, P., Lewis, C. C., Curry, J., & March, J. (2012). Does Homework “Matter” in Cognitive Behavioral Therapy for Adolescent Depression? *Journal of Cognitive Psychotherapy*, *26*(4), 390–404. https://doi.org/10.1891/0889-8391.26.4.390

Skinner, H., Steinhauer, P., & Santa-Barbara, J. (1995). *Family assessment measure version III*. North Tonawanda, NY: Multi-Health Systems.

Smith, J. A., & Osborn, M. (2008). Interpretative Phenomenological Analysis. In J. A. S. G. Breakwell, C. Fife- Schaw, S. Hammond (Ed.), *Doing Social Psychology Research* (3rd ed., pp. 229–254). https://doi.org/10.1002/9780470776278.ch10

Snaith, P. (1993). What do Depression Rating Scales Measure? *British Journal of Psychiatry*, *163*, 293–298.

Sobell, L. C., Sobell, M. B., Leo, G. I., & Cancilla, A. (1988). Reliability of a Timeline Method: assessing normal drinkers’ reports of recent drinking and a comparative evaluation across several populations. *British Journal of Addiction*, *83*(4), 393–402. https://doi.org/10.1111/j.1360-0443.1988.tb00485.x

Spearman, C. (1987). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, *100*(3/4), 441–471. https://doi.org/10.2307/1422689

Spence, S. H. (1995). *Social Skills Training: enhancing social competence with children and* *adolescents. Research and technical supplement.Spence, S. H. (1995). Social Skills Training: enhancing social competence with children and adolescents. Research and technical supplement*. Windsor, England: NFER-Nelson.

Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy*, *36*(5), 545–566. https://doi.org/10.1016/S0005-7967(98)00034-5

Spence, S. H., O'Shea, G., & Donovan, C. L. (2016). Improvements in Interpersonal Functioning Following Interpersonal Psychotherapy (IPT) with Adolescents and their Association with Change in Depression. *Behavioural and Cognitive Psychotherapy*, *44*(3), 257–272. https://doi.org/10.1017/S1352465815000442

Sperry, L., Brill, P. L., Howard, K. I., & Grissom, G. R. (1996). Treatment outcomes in psychotherapy and psychiatric interventions. In *Treatment outcomes in psychotherapy and psychiatric interventions.* Philadelphia, PA: Brunner/Mazel.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory (STAI).* Palo Alto, CA: Consulting Psychologists Press.

Spirito, A., Abebe, K. Z., Iyengar, S., Brent, D., Vitiello, B., Clarke, G., … Keller, M. (2009). Sources of Site Differences in the Efficacy of a Multisite Clinical Trial: The Treatment of SSRI-Resistant Depression in Adolescents. *Journal of Consulting and Clinical Psychology*, *77*(3), 439–450. https://doi.org/10.1037/a0014834

Spitzer, R. L., Kroenke, K., & Williams, J. B. W. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ Primary Care Study. *Journal of the American Medical Association*, *282*(18), 1737–1744. https://doi.org/10.1001/jama.282.18.1737

Stasiak, K., Parkin, A., Seymour, F., Lambie, I., Crengle, S., Pasene-Mizziebo, E., & Merry, S. (2013). Measuring outcome in child and adolescent mental health services: Consumers' views of measures. *Clinical Child Psychology and Psychiatry*, *18*(4), 519–535. https://doi.org/10.1177/1359104512460860

Stenner, P. H. D., Cooper, D., & Skevington, S. M. (2003). Putting the Q into quality of life; the identification of subjective constructions of health-related quality of life using Q methodology. *Social Science and Medicine*, *57*(11), 2161–2172. https://doi.org/10.1016/S0277-9536(03)00070-4

Stephenson, W. (1935). Correlating persons in stead of tests. *Character and Personality*, *4*, 17–24.

Stotland, N. L., Mattson, M. G., & Bergeson, S. (2008). The Recovery Concept: Clinician and Consumer Perspectives. *Journal of Psychiatric Practice*, *14*, 45–54.

Strandholm, T., Karlsson, L., Kiviruusu, O., Pelkonen, M., & Marttunen, M. (2014). Treatment Characteristics and Outcome of Depression Among Depressed Adolescent Outpatients

With and Without Comorbid Axis II Disorders. *Journal of Personality Disorders*, *28*(6), 853–863. https://doi.org/10.1521/pedi_2012_26_073

Stratton, P., Bland, J., Janes, E., & Lask, J. (2010). Developing an indicator of family function and a practicable outcome measure for systemic family and couple therapy: The SCORE. *Journal of Family Therapy*, *32*(3), 232–258. https://doi.org/10.1111/j.1467-6427.2010.00507.x

Striley, C. W., Stiffman, A. R., & Spitznagel, E. L. (2003). Functioning Mediates Between Symptoms and Provider Assessment. *Mental Health Services Research*, *5*(3).

Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcomes. With special reference to negative effects in psychotherapy. *The American Psychologist*, *32*(3), 187–196.

Stulz, N., & Lutz, W. (2007). Multidimensional patterns of change in outpatient psychotherapy: The phase model revisited. *Journal of Clinical Psychology*, *63*(9), 817–833. https://doi.org/10.1002/jclp.20397

Szatmari, P., Offringa, M., Butcher, N. J., & Monga, S. (2019). Counting What Counts: The Case for Harmonized Outcomes in Child and Youth Mental Health Research. *Journal of the American Academy of Child and Adolescent Psychiatry*, *58*(7), 656–658. https://doi.org/10.1016/j.jaac.2019.02.016

Tao, R., Emslie, G. J., Mayes, T. L., Nakonezny, P. A., & Kennard, B. D. (2010). Symptom Improvement and Residual Symptoms during Acute Antidepressant Treatment in Pediatric Major Depressive Disorder. *Journal of Child and Adolescent Psychopharmacology*, *20*(5), 423–430. https://doi.org/10.1089/cap.2009.0116

Tao, R., Emslie, G., Mayes, T., Nakonezny, P., Kennard, B., & Hughes, C. (2009). Early prediction of acute antidepressant treatment response and remission in pediatric major depressive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(1), 71–78. https://doi.org/10.1097/CHI.0b013e318190043e

Taplin, D. H., & Clark, H. (2012). *Theory of change basics: A primer on theory of change*. New York, NY.

Tashakkori, A., & Creswell, J. W. (2007). Editorial: Exploring the Nature of Research Questions in Mixed Methods Research. *Journal of Mixed Methods Research*, *1*(3), 207–211. https://doi.org/10.1177/1558689807302814

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.

Tashakkori, A., & Teddlie, C. (2003). *Handbook of Mixed Methods in Social & Behavioral Research*. Thousand Oaks, CA: Sage Publications.

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., … Stewart-Brown, S. (2007). The Warwick-Edinburgh mental well-being scale (WEMWBS): development and UK validation. *Health and Quality of Life Outcomes*, *5*(1), 63.

Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *Lancet*, *379*(9820), 1056–1067. https://doi.org/10.1016/S0140-673660871-4

The Health Foundation. (2016). *Person-centred care made simple: What everyone should know about person-centred care*. Retrieved from https://www.health.org.uk/sites/default/files/PersonCentredCareMadeSimple_0.pdf

The King's Fund. (2012). *From vision to action: Making patient-centred care a reality*. Retrieved from https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/Richmond-group-from-vision-to-action-april-2012-1.pdf

The Lancet Psychiatry. (2020). Measuring success: the problem with primary outcomes. *The Lancet Psychiatry*, *7*(1), 1. https://doi.org/10.1016/S2215-0366(19)30483-3

The Scottish Government. (2009). *Changing Lives: Personalisation: A Shared Understanding*. Retrieved from https://www2.gov.scot/resource/doc/269193/0080033.pdf

Thompson, M., Kaslow, N. J., & Nolen-hoeksema, S. (1998). Children's Attributional Style Questionnaire — Revised : Psychometric examination. *Psychological Assessment*, *10*(2), 166–170. https://doi.org/10.1037/1040-3590.10.2.166

Thurstone, L. L. (1947). *Multiple-factor analysis.* Chicago, IL: University of Chicago Press.

Tielen, M., Van Staa, A. L., Jedeloo, S., Van Exel, N. J. A., & Weimar, W. (2008). Q-methodology to identify young adult renal transplant recipients at risk for nonadherence. *Transplantation*, *85*(5), 700–706. https://doi.org/10.1097/TP.0b013e318166163e

Timulak, L., & Creaner, M. (2010). Qualitative meta-analysis of outcomes of personcentred/experiential therapies. In M. Cooper, J. C. Watson, & D. Holledampf (Eds.), *Person-Centred and Experiential Psychotherapies Work*. Ross-on-Wye, England: PCCS Books.

Todd, Z., Nerlich, B., McKeown, S., & Clarke, D. D. (Eds.). (2004). *Mixing Methods in Psychology: The Integration of Qualitative and Quantitative Methods in Theory and Practice*. Hove, England: Psychology Press.

Torpey, D. C., & Klein, D. N. (2008). Chronic depression: Update on classification and treatment. *Current Psychiatry Reports*, *10*(6), 458–464. https://doi.org/10.1007/s11920-008-0074-6

Torraco, R. J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review*, *4*(3), 356–367.

https://doi.org/10.1177/1534484305278283

Trauer, T. (1998). Issues in the assessment of outcome in mental health. *Australian and New Zealand Journal of Psychiatry*, *32*(3), 337–343. https://doi.org/10.3109/00048679809065525

Treatment for Adolescents With Depression Study (TADS) Team. (2004). Fluoxetine, Cognitive-Behavioral Therapy, and Their Combination for Adolescents With Depression. *JAMA*, *292*(7), 807. https://doi.org/10.1001/jama.292.7.807

Troupp, C. (2013). Distinguishing patient-generated outcome measures in CAMHS from a bespoke sandwich. *Child and Family Clinical Psychology Review*, *1*(1), 19–28.

Trowell, J., Joffe, I., Campbell, J., Clemente, C., Almqvist, F., Soininen, M., … Tsiantis, J. (2007). Childhood depression: A place for psychotherapy - An outcome study comparing individual psychodynamic psychotherapy and family therapy. *European Child and Adolescent Psychiatry*, *16*(3), 157–167. https://doi.org/10.1007/s00787-006-0584-x

Turoff, M. (1971). An alternative approach to cross impact analysis. *Technological Forecasting and Social Change*, *3*(c), 309–339. https://doi.org/10.1016/S0040-1625(71)80021-5

Tweed, D. L. (1993). Depression-related impairment estimating concurrent and lingering effects. *Psychological Medicine*, *23*(2), 373–386. https://doi.org/10.1017/S0033291700028476

Urwin, C. (2007). Revisiting “What works for whom?”: A qualitative framework for evaluating clinical effectiveness in child psychotherapy. *Journal of Child Psychotherapy*, *33*(2), 134–160. https://doi.org/10.1080/00754170701431370

US Department of Health and Human Services. (2008). *Personalized Health Care: Pioneers, Partnerships, Progress*. Washington, DC.

Valderas, J. M., Kotzeva, A., Espallargues, M., Guyatt, G., Ferrans, C. E., Halyard, M. Y., … Alonso, J. (2008). The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Quality of Life Research*, *17*, 179–193. https://doi.org/10.1007/s11136-007-9295-0

Values-Based Child and Adolescent Mental Health System Commission. (2016). *What Really Matters in Children and Young People’s Mental Health.* Retrieved from https://valuesbasedpractice.org/wp-content/uploads/2015/04/Values-based-full-report.pdf

Van der Kooy, K., van Hout, H., Marwijk, H., Marten, H., Stehouwer, C., & Beekman, A. (2007). Depression and the risk for cardiovascular diseases: Systematic review and meta analysis. *International Journal of Geriatric Psychiatry*, *22*(7), 613–626. https://doi.org/10.1002/gps.1723 van Exel, J. A., & de Graaf, G. (2005). *Q methodology: A sneak preview*. Retrieved from http://qmethod.org/articles/vanExel.pdf

van Exel, J. A., de Graaf, G., & Brouwer, W. B. F. F. (2006). “Everyone dies, so you might as well have fun!” Attitudes of Dutch youths about their health lifestyle. *Social Science and Medicine*, *63*(10), 2628–2639. https://doi.org/10.1016/j.socscimed.2006.06.028

Varni, J. W., Seid, M., & Kurtin, P. S. (2001). PedsQL™ 4.0: Reliability and Validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in Healthy and Patient Populations. *Medical Care*, *39*(8), 800–812. https://doi.org/10.1097/00005650-200108000-00006

Vitiello, B., Emslie, G., Clarke, G., Wagner, K. D., Asarnow, J. R., Keller, M. B., … Brent, D. A. (2011). Long-term outcome of adolescent depression initially resistant to selective serotonin reuptake inhibitor treatment: A follow-up study of the TORDIA sample. *Journal of Clinical Psychiatry*, *72*(3), 388–396. https://doi.org/10.4088/JCP.09m05885blu

Vrouva, I., Fonagy, P., Fearon, P. R. M., & Roussow, T. (2010). The risk-taking and self-harm inventory for adolescents: Development and psychometric evaluation. *Psychological Assessment*, *22*(4), 852–865. https://doi.org/10.1037/a0020583

Ware, J. E., Kosinski, M., & Keller, S. D. (1996). A 12 item short form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*, *34*(3), 220–233. https://doi.org/10.2307/3766749

Watts, S., & Stenner, P. (2005). Doing Q methodology: Theory, method and interpretation. *Qualitative Research in Psychology*, *2*(1), 67–91. https://doi.org/10.1191/1478088705qp022oa

Watts, S., & Stenner, P. (2012). *Doing Q Methodological Research: Theory, Method and Interpretation*. https://doi.org/10.4135/9781446251911

Weber, R. (1990). *Basic Content Analysis*. Thousand Oaks, CA: SAGE.

Weeks, M., Ploubidis, G. B., Cairney, J., Wild, T. C., Naicker, K., & Colman, I. (2016). Developmental pathways linking childhood and adolescent internalizing, externalizing, academic competence, and adolescent depression. *Journal of Adolescence*, *51*, 30–40. https://doi.org/10.1016/j.adolescence.2016.05.009

Weissman, M. M., & Bothwell, S. (1976). Assessment of Social Adjustment by Patient Self-Report. *Archives of General Psychiatry*, *33*(9), 1111–1115. https://doi.org/10.1001/archpsyc.1976.01770090101010

Weisz, J. R., Chorpita, B. F., Frye, A., Ng, M. Y., Lau, N., Bearman, S. K., … Hoagwood, K. E. (2011). Youth Top Problems: using idiographic, consumer-guided assessment to identify treatment needs and to track change during psychotherapy. *Journal of*

Consulting and Clinical Psychology, *79*(3), 369–380. https://doi.org/10.1037/a0023307

Weisz, J. R., Doss, A. J., & Hawley, K. M. (2005). Youth Psychotherapy Outcome Research: A Review and Critique of the Evidence Base. *Annual Review of Psychology*, *56*(1), 337–363. https://doi.org/10.1146/annurev.psych.55.090902.141449

Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., … Fordwood, S. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *American Psychologist*, Vol. 72, pp. 79–117. American Psychological Association.

Weisz, J. R., Southam-Gerow, M. A., Gordis, E. B., Connor-Smith, J. K., Chu, B. C., Langer, D. A., … Weiss, B. (2009). Cognitive-Behavioral Therapy Versus Usual Clinical Care for Youth Depression: An Initial Test of Transportability to Community Clinics and Clinicians. *Journal of Consulting and Clinical Psychology*, *77*(3), 383–396. https://doi.org/10.1037/a0013877

Weisz, J. R., & Weiss, B. (1991). Studying the “referability” of child clinical problems. *Journal of Consulting and Clinical Psychology*, *59*(2), 266–273. https://doi.org/10.1037/0022-006x.59.2.266

Wells, K. B., Tang, L., Carlson, G. A., & Asarnow, J. R. (2012). Treatment of Youth Depression in Primary Care Under Usual Practice Conditions: Observational Findings from Youth Partners in Care. *Journal of Child and Adolescent Psychopharmacology*, *22*(1), 80–90. https://doi.org/10.1089/cap.2011.0074

Westen, D., & Bradley, R. (2005). Empirically Supported Complexity: Rethinking Evidence-Based Practice in Psychotherapy. *Current Directions in Psychological Science*, *14*(5), 266–271.

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, *52*(5), 546–553. https://doi.org/10.1111/j.1365-2648.2005.03621.x

Whooley, M. A., De Jonge, P., Vittinghoff, E., Otte, C., Moos, R., Carney, R. M., … Browner, W. S. (2008). Depressive symptoms, health behaviors, and risk of cardiovascular events in patients with coronary heart disease. *JAMA*, *300*(20), 2379–2388. https://doi.org/10.1001/jama.2008.711

Wichstrom, L. (1995). Harter’s Self-Perception Profile for Adolescents: Reliability, Validity, and Evaluation of the Question Format. *Journal of Personality Assessment*, *65*(1), 100–116. https://doi.org/10.1207/s15327752jpa6501_8

Wiese, B. S. (2007). Successful pursuit of personal goals and subjective well-being. In B. R. Little, K. Salmela-Aro, & S. D. Phillips (Eds.), *Personal project pursuit: Goals, action, and human flourishing* (Vol. 14, pp. 301–328). Mahwah, NJ: Lawrence Erlbaum Associates

Publishers.

Wiggins, A., Oakley Browne, M., Bearsley-Smith, C., & Villanueva, E. (2010). Depressive disorders among adolescents managed in a child and adolescent mental health service. *Australasian Psychiatry*, *18*(2), 134–141. https://doi.org/10.3109/10398560903296657

Wilkinson, P., Dubicka, B., Kelvin, R., Roberts, C., & Goodyer, I. (2009). Treated depression in adolescents: Predictors of outcome at 28 weeks. *British Journal of Psychiatry*, *194*(4), 334–341. https://doi.org/10.1192/bjp.bp.108.052381

Wilkinson, P., & Goodyer, I. M. (2008). The effects of cognitive-behavioural therapy on mood-related ruminative response style in depressed adolescents. *Child and Adolescent Psychiatry and Mental Health*, *2*(1), 3. https://doi.org/10.1186/1753-2000-2-3

Williams, K., Thomson, D., Seto, I., Contopoulos-Ioannidis, D. G., Ioannidis, J. P. A., Curtis, S., … Offringa, M. (2012). Standard 6: Age groups for pediatric trials. *Pediatrics*, *129*(SUPPL. 3), S153–S160. https://doi.org/10.1542/peds.2012-0055I

Wolfe, A. (2001). Institute of Medicine Report: Crossing the Quality Chasm: A New Health Care System for the 21st Century. *Policy, Politics, & Nursing Practice*, *2*(3), 233–235. https://doi.org/10.1177/152715440100200312

Wolpert, M. (2017). Commentary: Why measuring clinical change at the individual level is challenging but crucial – commentary on Jensen and Corralejo (2017). *Child and Adolescent Mental Health*, *22*(3), 167–169. https://doi.org/10.1111/camh.12228

Wolpert, M., Cheng, H., & Deighton, J. (2015). Measurement issues: Review of four patient reported outcome measures: SDQ, RCADS, C/ORS and GBO - their strengths and limitations for clinical use and service evaluation. *Child and Adolescent Mental Health*, *20*(1), 63–70. https://doi.org/10.1111/camh.12065

Wolpert, M., Ford, T., Trustam, E., Law, D., Deighton, J., Flannery, H., & Fugard, R. J. B. (2012). Patient-reported outcomes in child and adolescent mental health services (CAMHS): Use of idiographic and standardized measures. *Journal of Mental Health*, *21*(2), 165–173. https://doi.org/10.3109/09638237.2012.664304

Wolpert, M., Fugard, A. J. B., Deighton, J., & Görzig, A. (2012). Routine outcomes monitoring as part of children and young people’s Improving Access to Psychological Therapies (CYP IAPT) - improving care or unhelpful burden? *Child and Adolescent Mental Health*, *17*(3), 129–130. https://doi.org/10.1111/j.1475-3588.2012.00676.x

Wolpert, M., Görzig, A., Deighton, J., Fugard, A. J. B., Newman, R., & Ford, T. (2015). Comparison of indices of clinically meaningful change in child and adolescent mental health services: Difference scores, reliable change, crossing clinical thresholds and “added value” - an exploration using parent rated scores on the SDQ. *Child and Adolescent Mental Health*, *20*(2), 94–101. https://doi.org/10.1111/camh.12080

Wolpert, M., Jacob, J., Napoleone, E., Whale, A., Calderon, A., & Edbrooke-Childs, J. (2016). *Child- and Parent-reported Outcomes and Experience from Child and Young People's Mental Health Services 2011–2015*. London, England: CAMHS Press.

Wolpert, M., & Rutter, H. (2018). Using flawed, uncertain, proximate and sparse (FUPS) data in the context of complexity: Learning from the case of child mental health. *BMC Medicine*, *16*(1), 1–11. https://doi.org/10.1186/s12916-018-1079-6

World Health Organization. (2007). *People-Centred Health Care: A Policy Framework*. Retrieved from http://www.wpro.who.int/health_services/people_at_the_centre_of_care/documents/ENG-PCIPolicyFramework.pdf

World Health Organization. (2008). International classification of functioning, disability and health: ICF. In *World Report on Child Injury Prevention.* Retrieved from https://apps.who.int/iris/bitstream/handle/10665/43851/9789241563574_eng.pdf;jsessionid=39FF8A78C7D7F436BCE40AE9AC872A51?sequence=1

World Health Organization. (2012). *Depression: A global public health concern.* Retrieved from http://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf

World Health Organization. (2017a). *Depression and Other Common Mental Disorders: Global Health Estimates*. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf

World Health Organization. (2017b). *Recognizing adolescence*. Retrieved from http://apps.who.int/adolescent/second-decade/section2/page1/recognizing-adolescence.html

World Health Organization. (2018). International Classification of Diseases for Mortality and Morbidity Statistics (11th Revision). Retrieved October 15, 2019, from https://icd.who.int/browse11/l-m/en

World Health Organization. (2019). Adolescent mental health. Retrieved December 3, 2019, from https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health

Yang, W., Zhang, J. X., Ding, Z., & Xiao, L. (2016). Attention Bias Modification Treatment for Adolescents with Major Depression: A Randomized Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *55*(3), 208–218. https://doi.org/10.1016/j.jaac.2015.12.005

Yao, S., Zhang, C., Zhu, X., Jing, X., McWhinnie, C. M., & Abela, J. R. Z. (2009). Measuring adolescent psychopathology: psychometric properties of the self-report strengths and difficulties questionnaire in a sample of Chinese adolescents. *The Journal of Adolescent*

Health : Official Publication of the Society for Adolescent Medicine, *45*(1), 55–62. https://doi.org/10.1016/j.jadohealth.2008.11.006

Yardley, L. (2000). Dilemmas in qualitative health research. *Psychology and Health*, *15*(2), 215–228. https://doi.org/10.1080/08870440008400302

Yeh, M., & Weisz, J. R. (2001). Why are we here at the clinic? Parent-child (dis)agreement on referral problems at outpatient treatment entry. *Journal of Consulting and Clinical Psychology*, *69*(6), 1018–1025. https://doi.org/10.1037/0022-006X.69.6.1018

Zhand, N., Courtney, D. B., & Flament, M. F. (2015). Use of electroconvulsive therapy in adolescents with treatment-resistant depressive disorders a case series. *Journal of ECT*, *31*(4), 238–245. https://doi.org/10.1097/YCT.0000000000000236

Zimmerman, M., Martinez, J. H., Friedman, M., Boerescu, D. A., Attiullah, N., & Toba, C. (2012). How Can We Use Depression Severity to Guide Treatment Selection When Measures of Depression Categorize Patients Differently? *The Journal of Clinical Psychiatry*, *73*(10), 1287–1291. https://doi.org/10.4088/JCP.12m07775

Zimmerman, M., McGlinchey, J. B., Posternak, M. A., Friedman, M., Boerescu, D., & Attiullah, N. (2008). Remission in depressed outpatients: More than just symptom resolution? *Journal of Psychiatric Research*, *42*(10), 797–801. https://doi.org/10.1016/j.jpsychires.2007.09.004

Appendices

Appendix A. Supplementary Materials – Chapter 3

Table A.1. Models and Taxonomies of Treatment Outcome in Child Mental Health

Domains	**Subdomains / descriptions / Examples**
Hoagwood et al. (1996, 2012)	
Symptoms and diagnoses	E.g. impulsivity, depression, or anxiety.
Functioning	Capacity to adapt to expectation of the family, school, or community.
Consumer satisfaction	Quality of life, satisfaction with the care provided, family strain or burden.
Environment	Stability of the child's immediate environment (e.g. family relationships; dynamics at school or in the community).
Service	Characteristics of and changes in the use, organisation, resourcing and planning of services.
Parental symptoms and diagnosis	Symptoms experienced by the caregiver.
Health	Physical or biological changes that can be observed in the child's body following treatment, excluding physical expressions of a metal health disorder.
Fonagy (1997)	
Symptomatic / diagnostic level	Full breadth of an individual's symptoms.
Adaptation	Functioning / adaptation to the psychosocial environment.
Mechanisms	Cognitive and emotional capacities assumed to underpin symptomatology and adaptation (e.g. affect regulation, understanding emotions, and attributional biases).
Transactional level	"Transactional interactions between the mental state and behavioural predispositions of the child and the reactions of the environment to it across time" (1997, p. 587)
Service utilisation	Service utilisation; quality and integration of health services.
Grosse-Holtforth & Grawe (2002)	
Coping with specific problems and symptoms	• Depressive symptoms • Suicidality and self-injury • Fears or anxiety • Obsessive thoughts and compulsive behaviours • Coping with trauma • Substance use and addiction • Eating behaviours • Sleep • Sexuality • Coping with somatic problems • Difficulties in specific life domains/stress
Interpersonal goals	• Current relationship • Parenthood • Other specific relationships • Loneliness and grief • Assertiveness • Connectedness and intimacy

Domains	**Subdomains / descriptions / Examples**
Wellbeing and functioning	• Exercise and activity • Relaxation and composure • Well-being
Existential issues	• Past, present, and future • Meaning of life
Personal growth	• Attitude toward self • Desires and wishes • Responsibility and self-control • Emotion regulation
Childs et al. (2013)	
Individual level	• Symptoms • Achievement of symptom-related goals • Achievement of self-defined goals • Functioning • Empowerment and self-efficacy • Appropriate discharge
Interpersonal level	• Interpersonal relationships • Functioning of the social context • Systemic nature of difficulties • Caregiver’s stress and management of YP • Support through wider context
Service / societal level	• Service engagement • Service user satisfaction • Degree of linkage between CYP's network • Whether best practice was followed
Bradley et al. (2013)	
Relationship/interpersonal: listening and understanding	• Make more friends / get along with others better • Make family (or living situation) feel better • Be able to communicate with people more • Talk about feelings and thoughts • Getting on better with mum or dad (or both) • Letting people know what help I need
Coping with specific problems and symptoms	• Managing mood, negative thoughts and feelings • Controlling and managing my anger • Stop feeling anxious, stressed, worried • Be good / help with behaviour • Better sleep • Feel happier • Stop harming myself • Sleep on my own • Understanding anger
Personal growth & functioning: Understanding and improving self	• Feeling more confident / better within myself • Goals with personal meaning/ related to hobby • To be responsible for myself / more independent • Doing better at school (includes behaviour) • Feeling comfortable at school and able to attend • Thinking about and understanding me and my past • Relationships at school including bullying • Enjoying life

Domains	**Subdomains / descriptions / Examples**
Rupani et al. (2014)	
Emotional goals	• Controlling or reducing anger • Controlling or managing emotions • Increasing happiness/reducing • feelings of upset • Reducing anxiety or worry/increasing calmness
Interpersonal Goals	• Dealing with bullying or being criticized • Dealing with bereavement • Improving relationship(s) with family member(s) • Improving friendships and relationships (exc. bullying) • Talking more about feelings and experiences
Goals targeting specific issues	• Addressing self-harm or suicidal thoughts and feelings • Improving sleep • Improving things at school/academic work • Addressing physical health issues • Other/goals not otherwise specified
Personal growth	• Making sense of personal history • Increasing self-confidence or self- acceptance

Appendix B. Supplementary Materials – Chapter 4

Table B.1. Search Syntax

Element	Line	Syntax
Identification of trials	1	((trial adj2 (clinical or treatment or control*)).ab,ti.
	2	((experimental adj (design or treatment)).ab,ti.
	3	(randomi#ed or randomly or random allocation).ab,ti.
	4	(placebo or waitlist).ab,ti.
	5	((group* or patients or children) adj (control or comparison or treatment)).ab,ti.
	6	((singl* OR doubl* OR trebl* OR tripl*) adj3(blind* OR mask*)).ab,ti.
	7	(1 or 2 or 3 or 4 or 5 or 6).ab,ti.
	8	(exp animals/ not humans.sh.).ab,ti.
	9	(7 not 8).ab,ti.
Identification of other study designs	10	((routine adj3 care) or (routine adj3 practice) or (routine adj3 monitoring) or (care adj2 system) or service use*)).ab,ti.
	11	(treatment as usual).ab,ti.
	12	((mental health adj3 clinic*) or (mental health adj3 hospital*) or (mental health adj3 patient*) or (mental health adj3 service*) or (mental health adj3 setting*) or (mental health adj3 system*)).ab,ti.
	13	((community adj3 clinic*) or (community* adj3 hospital*) or (community adj3 service*) or (community* adj3 setting*) or (community* adj3 care)).ab,ti.
	14	((psych* adj3 clinic) or (psych* adj3 hospital) or (psych* adj3 patient*) or (psych* adj3 service*)).ab,ti.
	15	((school* adj3 settings*) or (school* adj3 intervention*) or (school* adj3 treat*)).ab,ti.
	16	((clinical adj3 care) or (clinical adj3 setting*)).ab,ti.
	17	((service adj3 data) or (service adj3 report*) or (service adj3 evaluat*) or (service adj3 record*) or (service adj3 feedback)).ab,ti.
	18	((clinic*2 adj3 data) or (clinic*2 adj3 report*) or (clinic*2 adj3 feedback) or (clinic*2 adj3 records*) or (clinic*2 adj3 sample)).ab,ti.
	19	((program* adj3 data) or (program* adj3 report*) or (program* adj3 evaluat*) or (program* adj3 assess*)).ab,ti.
	20	((hospital adj3 record*) or (hospital adj3 data) or (hospital adj3 sample)).ab,ti.
	21	((feedback adj3 patient*) or (feedback adj3 service user*) or (feedback adj3 client*)).ab,ti.
	22	(outpatient* or inpatient*).ab,ti.

Element	Line	Syntax
Specification: treatment effectiveness studies	23	((intervention adj4 effect*) or (intervention adj4 efficacy) or (intervention adj4 outcome*) or (intervention adj4 success) or (intervention adj4 impact) or (intervention adj4 result*)).ab,ti.
	24	((treat* adj4 effect*) or (treat* adj4 efficacy) or (treat* adj4 outcome*) or (treat* adj4 success) or (treat* adj4 impact) or (treat* adj4 result*)).ab,ti.
	25	((program*3 adj4 effect*) or (program*3 adj4 efficacy) or (program*3 adj4 outcome*) or (program*3 adj4 success) or (program*3 adj4 impact) or (program*3 adj4 result*)).ab,ti.
	26	(outcom* or impact or effective* or efficacy).ab,ti.
Population	27	(child*3).ab,ti.
	28	(adolescen*).ab,ti.
	29	(((Young adj3 person) or (young people))).ab,ti.
	30	(teenag*).ab,ti.
	31	(pupil*).ab,ti.
	32	(student*).ab,ti.
	33	(youth*).ab,ti.
	34	(young adj3 adults).ab,ti.
	35	(juvenile*).ab,ti.
	36	(school-age*).ab,ti.
	37	(age*1 adj4 ("5" or "6" or "7" or "8" or "9" or "10" or "11" or "12" or "13" or "14" or "15" or "16" or "17")).ab,ti.
Diagnosis	38	(mental health).ab,ti.
	39	(Psychopatholog*).ab,ti.
	40	((Psycholog* adj2 difficult*) or (psycholog* adj2 disorder*) or (psycholog* adj2 distress*) or (psycholog* adj2 illness*) or (psycholog* adj2 needs) or (psycholog* adj2 problem*)).ab,ti.
	41	(Psychiatr*).ab,ti.
	42	(Emotional adj3 (difficult* or disorder* or functioning or health or issue* or problem* or recovery or well#being)).ab,ti.
	43	(Depress* or bipolar).ab,ti.
	44	(Internalising adj3 (difficult* or disorder* or issue* or problem*)).ab,ti.
	45	(Mood adj3 (disorder* or dysregulation or low*)).ab,ti.
	46	(affective disorder*).ab,ti.
	47	(Anxiety or panic or phobia*).ab,ti.
Combinations	48	9
	49	10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22
	50	23 or 24 or 25 or 26
	51	27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37
	52	38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47
	53	48 or 49
	54	50 and 51 and 52 and 53
	55	54 and 2007:2017.(sa_year).
	56	55 and "Peer Reviewed Journal".sa_pubt.

Element	Line	Syntax
Exclusion terms	57	(toddler* or preschool* or infant*1 or baby or babies).ti.
	58	(foetal or foetus or embryo* or pre#natal or post#natal or peri#natal or pre#birth or post birth).ti.
	59	(dyslexia or dyslexic or dyspraxia or dyspraxia or ((learning adj3 (difficulties or disabilit* or disorder)) or intellectual disabilit*)).ti.
	60	(ADHD or attention-deficit hyperactivity disorder).ti.
	61	(Autism or autist*).ti.
	62	(down-syndrome or developmental disorder or cerebral palsy).ti.
	63	(asthma or eczema or cancer or diabetes).ti.
	64	(geriatric or dementia or Alzheimer*).ti.
	65	(57 or 58 or 59 or 60 or 61 or 62 or 63 or 64).ti.
Final line	66	56 not 65

Note. The final search was completed on 25 July 2017.

Table B.2. Adapted Downs and Black (1998) Checklist for Data Quality

Area	No.	Criterion
Reporting	1	Hypothesis / aim / objective clear?
	2	Main outcomes described in intro / method?
	3	Sample characteristics clear?
	4	Interventions clearly described?
	5	Distribution of main confounders described
	6	Main findings clearly described?
	7	Random variability estimates for main outcome?
	9	Participants lost to follow-up described?
	10	Actual probability values for main outcome?
External validity	11	Subjects approached representative?
	12	Subjects consented representative?
	13	Treatment context representative?
Internal validity (bias)	14	Participants blinded?
	15	Assessors blinded?
	17	Adjust for different times of follow-up?
	18	Appropriate statistical tests used?
	20	Main outcome measures accurate?
Internal validity (confounding)	21	Patients in different groups recruited from same population?
	22	Patients in different groups recruited over same period?
	23	Randomised?
	25	Adequate adjustments for confounding?
	26	Losses of patients to follow up taken into account?
Power	27	Power sufficient?

Note. The original Downs & Black (1998) check list was adapted by removing four of the original criteria, thus reducing the total attainable quality score from 27 to 23. The removed items were: 8. *Have all important adverse events that may be a consequence of the intervention been reported?* ; 16. *If any of the results of the study were based on "data dredging", was this made clear?*; 19. *Was compliance with the intervention/s reliable?* 24.*Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?* These were removed because the large majority of studies did not report sufficient information for the doctoral candidate to assess these criteria.

Table B.3. Results of the Data Quality Assessment

Study	Criteria																						
	1	2	3	4	5	6	7	9	10	11	12	13	14	15	17	18	20	21	22	23	25	26	27	Quality Score
Abeles et al. (2009)	0	1	1	1	0	1	1	0	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	10
Amaya et al. (2011)	1	1	1	1	0	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	16
Ariga et al. (2010)	0	1	1	0	1	1	0	1	1	0	0	1	0	0	0	1	1	1	1	0	0	1	0	12
Asarnow et al. (2009)	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	0	19
Atkinson et al. (2014)	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	19
Becker-Weidman et al. (2010)	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	19
Bloomquist et al. (2016)	1	1	1	1	1	0	1	1	0	0	0	1	0	0	0	0	1	1	1	0	0	0	0	11
Brent et al. (2008)	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	21
Byford et al. (2007)	1	1	1	1	1	1	1	0	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	18
Carandang et al. (2007)	1	1	1	0	1	0	1	1	0	1	1	1	0	0	0	1	1	1	1	0	0	1	0	14
Carter et al. (2015)	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	20
Carter et al. (2016)												*Not assessed*												
Chafey et al. (2009)												*Not assessed*												
Charkhandeh et al. (2016)	1	1	1	1	1	1	1	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	15
Cheung et al. (2010)	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1	0	0	18
Clarke et al. (2016)	0	1	1	1	0	1	0	0	1	1	0	1	0	1	1	1	1	1	1	1	0	1	0	15
Cornelius et al. (2009)	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	0	18
Cornelius et al. (2011)	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	0	0	15
Curry et al. (2011)	1	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	17
DelBello et al. (2014)	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	0	1	1	1	1	0	1	18
Diamond et al. (2010)	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	0	18
Domino et al. (2008)	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	22
Domino et al. (2009)	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	22
Emslie et al. (2007)	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	19
Emslie et al. (2007)	1	1	1	1	1	0	1	1	0	0	0	1	0	0	1	0	1	1	1	0	1	0	0	13
Emslie et al. (2009)	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	20

Study	Criteria																							
	1	2	3	4	5	6	7	9	10	11	12	13	14	15	17	18	20	21	22	23	25	26	27	Quality Score
Emslie et al. (2010)	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	20
Emslie et al. (2012)	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	1	20
Emslie et al. (2014)	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	0	1	19
Eskin et al. (2008)	1	1	1	1	1	0	1	1	0	0	0	0	0	0	1	1	1	1	1	1	0	1	0	14
Feeny et al. (2009)	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	21
Findling et al. (2013)	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	0	0	18
Garoff et al. (2012)	1	1	1	1	1	0	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	0	0	16
Gibbs et al. (2012)													*Not assessed*											
Goodyer et al. (2017)	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	20
Goodyer et al. (2007)	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	20
Gordon et al. (2011)	1	1	1	1	0	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	0	0	16
Gunlicks-Stoessel et al. (2010)	1	1	1	1	1	0	1	0	0	0	0	1	0	1	1	1	1	1	1	1	1	0	0	15
Iftene et al. (2015)	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	0	0	0	16
Jacobs et al. (2009)	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	0	1	19
Jacobs et al. (2010)	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	20
Jiménez Chafey et al. (2011)	1	1	1	1	0	1	0	1	0	0	0	0	0	0	1	1	1	1	1	0	0	1	0	12
Karlsson et al. (2008)	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	1	0	0	0	1	1	0	15
Kennard et al. (2009)	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	21
Kennard et al. (2009)	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	20
Kennard et al. (2009)	1	1	1	1	1	0	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	1	18
Le Noury et al. (2015)	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	0	0	1	17
Lewis et al. (2009)	1	1	1	1	0	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	0	17
Lewis et al. (2010)	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	21
Lusk & Melnyk (2011)	1	1	1	1	1	1	0	1	1	0	0	1	0	0	1	0	1	1	1	0	0	1	0	14
Lynch et al. (2011)	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0	19
Manglick et al. (2013)	1	1	1	1	1	0	1	0	1	0	0	1	0	0	1	1	1	1	1	1	1	0	0	15
Mayes et al. (2007)	1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	17
McMakin et al. (2012)	1	1	1	1	0	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	1	18

Study	Criteria																							
	1	2	3	4	5	6	7	9	10	11	12	13	14	15	17	18	20	21	22	23	25	26	27	Quality Score
Melvin et al. (2013)	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	0	0	18
Merry et al. (2012)	1	1	1	0	0	1	1	1	0	0	0	1	0	1	1	0	1	1	1	1	0	0	0	13
Mufson et al. (2014)	1	1	1	1	1	1	1	0	1	0	0	1	0	1	1	1	1	1	1	1	1	0	0	17
Ngo et al. (2009)	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0	19
Niederhofer & Klitzing (2011)	1	1	1	1	1	0	0	1	0	0	0	0	0	0	1	0	1	1	1	0	0	1	0	11
Niederhofer & Klitzing (2012)	1	1	1	1	0	0	0	0	1	0	0	0	1	0	1	0	1	1	1	1	0	0	0	11
Norton (2010)												*Not assessed*												
O'Shea et al. (2015)	1	1	1	1	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0	18
Parker et al. (2016)	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	0	1	19
Pass et al. (2015)	1	1	1	1	1	1	0	1	0	0	0	1	0	0	1	1	1	1	1	0	0	1	0	14
Peters et al. (2016)	1	1	0	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	0	18
Poole et al. (2017)	1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	18
Rengasamy et al. (2013)	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	19
Riley (2011)	1	1	0	1	0	0	0	1	1	0	0	1	0	0	1	1	1	1	1	0	0	0	0	11
Rohde et al. (2008)	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	20
Rohde et al. (2014)	1	1	1	1	1	0	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	0	0	15
Rossello et al. (2008)	1	1	1	1	0	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	0	16
Sakolsky et al. (2011)	1	1	1	1	0	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	1	18
Schirman et al. (2010)	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0	14
Shamseddeen et al. (2011)	0	1	0	1	0	0	1	0	1	0	0	1	1	1	1	0	1	1	1	1	1	0	0	13
Shamseddeen et al. (2011)	0	1	1	1	1	0	1	0	1	0	0	1	1	1	1	0	1	1	1	1	1	0	0	15
Shamseddeen et al. (2012)	1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1	1	1	1	0	0	16
Shirk et al. (2008)	1	1	1	1	0	1	1	1	0	0	0	1	0	0	1	1	1	1	1	0	0	0	0	13
Shon et al. (2014)	1	1	1	0	0	0	1	0	1	1	1	1	0	0	1	1	1	1	1	0	1	0	0	14
Simons et al. (2012)	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	19
Spence et al. (2016)	1	1	1	1	0	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	0	0	16
Spirito et al. (2009)	1	1	1	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	1	19
Strandholm et al. (2014)	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	1	1	1	0	1	0	0	16

Study	Criteria																							
	1	2	3	4	5	6	7	9	10	11	12	13	14	15	17	18	20	21	22	23	25	26	27	Quality Score
TADS (2007)	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	1	1	1	0	1	0	0	16
TADS (2009)	1	1	1	1	0	1	1	1	1	0	0	1	0	0	1	1	1	1	1	0	1	1	0	16
Tao et al. (2009)	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	21
Tao et al. (2010)	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	20
Trowell et al. (2007)	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	1	1	19
Vitiello et al. (2011)	1	1	1	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	19
Weisz et al. (2009)	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0	19
Wells et al. (2012)	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	18
Wiggins et al. (2010)	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	0	0	0	0	16
Wilkinson & Goodyer (2008)	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	1	0	18
Wilkinson et al. (2009)	1	1	1	0	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	0	1	0	17
Yang et al. (2016)	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	20
Zhand et al. (2015)	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	1	0	0	1	0	16

Note. The criteria that these numbers refer to are listed in Table B.2 (above). Four studies were not assessed because they were either purely qualitative studies (*n* = 3) or a single case study (*n* = 1) so that the assessment criteria could not be meaningfully applied. Zero indicates that the relevant criterion was not met; one indicates that it was met.

Table B.4. Characteristics of Included Studies

Short title	Year	Country	Study Type	N	Age Range	Treatment	Anchor study
Abeles et al. (2009)	2009	UK	OBS	23	12–16	PsyTh	—
Amaya et al. (2011)	2011	USA	RCT	260	12–17	PsyTh + Med	TADS[1]
Ariga et al. (2010)	2010	Japan	OBS	64	16–19	PsyTh + Med	—
Asarnow et al. (2009)	2009	USA	RCT	334	12–18	PsyTh + Med	TORDIA[2]
Atkinson et al. (2014)	2014	USA, Europe, South Africa	RCT	337	7–17	Med	—
Becker-Weidman et al. (2010)	2010	USA	RCT	439	12–17	PsyTh + Med	TADS[1]
Bloomquist et al. (2016)	2016	USA	OBS	28	12–17	PsyTh	—
Brent et al. (2008)	2008	USA	RCT	334	12–18	PsyTh + Med	TORDIA[2]
Byford et al. (2007)	2007	UK	RCT	208	11–17	PsyTh + Med	ADAPT[3]
Carandang et al. (2007)	2007	Canada & USA	RCR	42	13–17	Med	—
Carter et al. (2015)	2015	UK	RCT	87	14–17	PsyTh	—
Carter et al. (2016)	2016	UK	QUAL	26	14–17	PsyTh	—
Chafey et al. (2009)	2009	Puerto Rico	SCS	1	15	PsyTh	—
Charkhandeh et al. (2016)	2016	Iran	RCT	188	12–17	PsyTh	—
Cheung et al. (2010)	2010	USA	RCT	309	7–17	Med	Emslie et al. (2002, 1997)
Clarke et al. (2016)	2016	USA	RCT	212	12–18	PsyTh + Med	—
Cornelius et al. (2009)	2009	USA	RCT	50	15–20	PsyTh + Med	—
Cornelius et al. (2011)	2011	USA	RCT	75	15–20	PsyTh + Med	Cornelius et al. (2009)
Curry et al. (2011)	2011	USA	NFUP	196	14–22	PsyTh + Med	TADS[1]
DelBello et al. (2014)	2014	USA	RCT	308	12–17	Med	—
Diamond et al. (2010)	2010	USA	RCT	66	12–17	PsyTh	—
Domino et al. (2008)	2008	USA	RCT	351	12–17	PsyTh + Med	TADS[1]
Domino et al. (2009)	2009	USA	RCT	327	12–17	PsyTh + Med	TADS[1]
Emslie et al. (2007)	2007	USA	RCT	334	7–17	Med	—
Emslie et al. (2007)	2007	USA	OBS	86	7–17	Med	—
Emslie et al. (2009)	2009	USA	RCT	312	12–17	Med	—
Emslie et al. (2010)	2010	USA	RCT	334	12–18	PsyTh + Med	TORDIA[2]
Emslie et al. (2012)	2012	USA	RCT	309	7–17	Med	Emslie et al. (2002, 1997)

Short title	Year	Country	Study Type	N	Age Range	Treatment	Anchor study
Emslie et al. (2014)	2014	USA, Canada, Mexico, Argentina	RCT	463	7–17	Med	—
Eskin et al. (2008)	2008	Iran	RCT	46	—	PsyTh	—
Feeny et al. (2009)	2009	USA	RCT	439	12–17	PsyTh + Med	TADS[1]
Findling et al. (2013)	2013	USA	RCT	165	12–17	Med	Emslie et al. (2009)
Garoff et al. (2012)	2012	UK, Greece, Finland	RCT	72	9–15	PsyTh	Trowell et al. (2007)
Gibbs et al. (2012)	2012	USA	SCS	1	15	PsyTh + Med	—
Goodyer et al. (2017)	2017	UK	RCT	470	11–17	PsyTh	IMPACT[4]
Goodyer et al. (2007)	2007	UK	RCT	208	11–17	PsyTh + Med	ADAPT[3]
Gordon et al. (2011)	2011	Australia	RCT	130	12–18	PsyTh + Med	Melvin et al. (2006)
Gunlicks-Stoessel et al. (2010)	2010	USA	RCT	63	12–18	PsyTh	—
Iftene et al. (2015)	2015	Romania	RCT	88	11–17	PsyTh + Med	—
Jacobs et al. (2009)	2009	USA	RCT	439	12–17	PsyTh + Med	TADS[1]
Jacobs et al. (2010)	2010	USA	RCT	439	12–17	PsyTh + Med	TADS[1]
Jiménez Chafey et al. (2011)	2011	Puerto Rico	SCS	2	14	PsyTh	—
Karlsson et al. (2008)	2008	Finland	NFUP	174	13–19	TAU (only)	ADS[5]
Kennard et al. (2009)	2009	USA	RCT	334	12–18	PsyTh + Med	TORDIA[2]
Kennard et al. (2009)	2009	USA	RCT	439	12–17	PsyTh + Med	TADS[1]
Kennard et al. (2009)	2009	USA	RCT	439	12–17	PsyTh + Med	TADS[1]
Le Noury et al. (2015)	2015	Canada & USA	RCT	275	12–18	Med	—
Lewis et al. (2009)	2009	USA	RCT	332	12–17	PsyTh + Med	TADS[1]
Lewis et al. (2010)	2010	USA	RCT	427	12–17	PsyTh + Med	TADS[1]
Lusk & Melnyk (2011)	2011	USA	OBS	15	12–17	PsyTh	—
Lynch et al. (2011)	2011	USA	RCT	334	12–18	PsyTh + Med	TORDIA[2]
Manglick et al. (2013)	2013	Australia	RCT	166	12–18	PsyTh + Med	—
Mayes et al. (2007)	2007	USA	RCT	315	—	Med	Emslie et al. (2002, 1997)
McMakin et al. (2012)	2012	USA	RCT	334	12–18	PsyTh + Med	TORDIA[2]
Melvin et al. (2013)	2013	Australia	NFUP	140	17–24	PsyTh + Med	TFF/BH(Melvin et al., 2006)
Merry et al. (2012)	2012	New Zealand	RCT	187	12–19	PsyTh	—
Mufson et al. (2014)	2014	USA	RCT	63	12–18	PsyTh	—
Ngo et al. (2009)	2009	USA	RCT	325	13–21	QI	YPIC[6]

Short title	Year	Country	Study Type	N	Age Range	Treatment	Anchor study
Niederhofer & Klitzing (2011)	2011	Unclear	SCS	2	14–17	PsyTh	Niederhofer & Klitzing (2012)
Niederhofer & Klitzing (2012)	2012	Unclear	RCT	28	14–17	PsyTh	—
Norton (2010)	2010	USA	SCS	1	16	PsyTh	—
O'Shea et al. (2015)	2015	Australia	RCT	39	13–19	PsyTh	Spence et al. (2016)
Parker et al. (2016)	2016	Australia	RCT	176	15–25	PsyTh	—
Pass et al. (2015)	2015	UK	SCS	1	15	PsyTh	—
Peters et al. (2016)	2016	USA	NFU	196	14–22	PsyTh + Med	TADS
Poole et al. (2017)	2017	Australia	RCT	64	12–18	PsyTh	—
Rengasamy et al. (2013)	2013	USA	RCT	334	12–18	PsyTh + Med	TORDIA
Riley (2011)	2011	USA	OBS	7	12–16	PsyTh	—
Rohde et al. (2008)	2008	USA	RCT	242	12–17	PsyTh + Med	TADS
Rohde et al. (2014)	2014	USA	RCT	170	13–18	PsyTh	—
Rossello et al. (2008)	2008	Puerto Rico	RCT	112	12–18	PsyTh	—
Sakolsky et al. (2011)	2011	USA	RCT	334	12–18	PsyTh + Med	TORDIA
Schirman et al. (2010)	2010	Israel	OBS	78	7–18	Med	—
Shamseddeen et al. (2011)	2011	USA	RCT	334	12–18	PsyTh + Med	TORDIA
Shamseddeen et al. (2011)	2011	USA	RCT	334	12–18	PsyTh + Med	TORDIA
Shamseddeen et al. (2012)	2012	USA	RCT	334	12–18	PsyTh + Med	TORDIA
Shirk et al. (2008)	2008	USA	OBS	54	14–18	PsyTh	—
Shon et al. (2014)	2014	Korea	RCR	37	6–18	Med	—
Simons et al. (2012)	2012	USA	RCT	218	12–17	PsyTh	TADS
Spence et al. (2016)	2016	Australia	RCT	39	13–19	PsyTh	—
Spirito et al. (2009)	2009	USA	RCT	334	12–18	PsyTh + Med	TORDIA
Strandholm et al. (2014)	2014	Finland	OBS	151	13–19	TAU (only)	ADS
TADS (2007)	2007	USA	RCT	439	12–17	PsyTh + Med	TADS
TADS (2009)	2009	USA	NFUP	439	12–17	PsyTh + Med	TADS
Tao et al. (2009)	2009	USA	OBS	168	7–18	Med	Emslie et al. 2008
Tao et al. (2010)	2010	USA	OBS	168	7–18	Med	Emslie et al. 2008
Trowell et al. (2007)	2007	UK, Greece, Finland	RCT	72	9–15	PsyTh	—
Vitiello et al. (2011)	2011	USA	NFUP	334	12–18	PsyTh + Med	TORDIA

Short title	Year	Country	Study Type	N	Age Range	Treatment	Anchor study
Weisz et al. (2009)	2009	USA	RCT	57	8–15	PsyTh	—
Wells et al. (2012)	2012	USA	RCT	344	13–21	QI	YPIC
Wiggins et al. (2010)	2010	Australia	OBS	76	12–18	TAU (only)	—
Wilkinson & Goodyer (2008)	2008	UK	RCT	23	11–17	PsyTh + Med	ADAPT
Wilkinson et al. (2009)	2009	UK	RCT	192	11–17	PsyTh + Med	ADAPT
Yang et al. (2016)	2016	China	RCT	45	12–18	PsyTh	—
Zhand et al. (2015)	2015	Canada	RCR	13	15–18	PsyTh	—

Note. ADAPT = The Adolescent Depression Antidepressant and Psychotherapy Trial (Goodyer et al., 2007). ADS = Adolescent Depression Study (Karlsson et al., 2006). IMPACT = Improving Mood with Psychoanalytic and Cognitive Therapies (Goodyer et al., 2017). NFUP = naturalistic follow-up. Med = medication. OBS = observational study (i.e., studies without control), including open label trials and case series. PsyTh = psychotherapy (and other active treatments). PsyTh + Med = treatment combining psychotherapeutic modalities and medication. QUAL = qualitative study. QI = quality improvement intervention; RCT = randomized control trial. RCR = retrospective chart review. SCS = single case study. TADS = Treatment for Adolescents with Depression Study (March et al., 2004). TORDIA = Treatment of Resistant Depression in Adolescents Study (Brent et al., 2008). YPIC = Youth Partners in Care (Asarnow et al., 2005).

Table B.5. Characteristics of Outcome Measurement in the Reviewed Studies

Short title	Study objective	Number of outcome domains covered	Number of outcome measures used	Informant(s) consulted	Primary outcome measure (if defined)
Abeles et al. (2009)	Assess single intervention	5	13	CL, YP, PA	—
Amaya et al. (2011)	Study of predictors / moderators	1	1	CL	CGI-I
Ariga et al. (2010)	Assess single intervention	2	3	YP	—
Asarnow et al. (2009)	Study of predictors / moderators	2	2	CL	CGI-I; CDRS-R
Atkinson et al. (2014)	Assess single intervention	2	2	CL	CDRS-R
Becker-Weidman et al. (2010)	Study of predictors / moderators	1	2	CL, YP	CDRS-R; SIQ-Jr / SIQ
Bloomquist et al. (2016)	Assess transportability of an intervention	4	7	YP & PA	—
Brent et al. (2008)	Compare interventions	2	6	CL, YP	CGI-I; CDRS-R
Byford et al. (2007)	Cost-effectiveness of an intervention	2	2	CL, YP	HoNOSCA
Carandang et al. (2007)	Assess single intervention	1	2	CL	CGI-S; CGI-I
Carter et al. (2015)	Assess single intervention	3	3	YP	CDI
Carter et al. (2016)	Assess treatment experience (qualitative)	6	—	—	—
Chafey et al. (2009)	Assess single intervention	3	6	CL, YP	—
Charkhandeh et al. (2016)	Compare interventions	1	1	YP	CDI
Cheung et al. (2010)	Study of predictors / moderators	2	3	CL	CGI-I; CDRS-R
Clarke et al. (2016)	Assess single intervention	6	14	CL, YP, PA	A-LIFE
Cornelius et al. (2009)	Assess single intervention	2	7	CL, YP	—
Cornelius et al. (2011)	Examine long-term outcomes	2	7	CL, YP	—
Curry et al. (2011)	Examine long-term outcomes	3	3	CL, YP	K-SADS
DelBello et al. (2014)	Compare different intensities of an intervention	2	3	CL	CDRS-R
Diamond et al. (2010)	Assess single intervention	2	4	CL, YP	SIQ-Jr / SIQ; SSI; BDI;
Domino et al. (2008)	Cost-effectiveness of an intervention	1	1	CL	CDRS-R
Domino et al. (2009)	Cost-effectiveness of an intervention	3	3	CL, YP	CDRS-R
Emslie et al. (2007)	Assess single intervention	2	5	*not stated*	CDRS-R

Short title	Study objective	Number of outcome domains covered	Number of outcome measures used	Informant(s) consulted	Primary outcome measure (if defined)
Emslie et al. (2007)	Assess single intervention	2	3	CL	CDRS-R
Emslie et al. (2009)	Assess single intervention	2	4	CL	CDRS-R
Emslie et al. (2010)	Compare interventions	2	3	CL	CGI-I; CDRS-R; A-LIFE;
Emslie et al. (2012)	Study of predictors / moderators	2	3	CL	CDRS-R; CGI-I
Emslie et al. (2014)	Assess single intervention	2	2	CL	CDRS-R; CGI-S
Eskin et al. (2008)	Assess single intervention	3	7	CL, YP	—
Feeny et al. (2009)	Study of predictors / moderators	2	3	CL, YP	CDRS-R; RADS; CGAS;
Findling et al. (2013)	Assess intervention extension	2	4	CL	CDRS-R
Garoff et al. (2012)	Compare interventions	2	4	CL, YP, PA	—
Gibbs et al. (2012)	Assess single intervention	2	—	—	—
Goodyer et al. (2017)	Compare interventions	4	9	CL, YP	MFQ
Goodyer et al. (2007)	Assess combination treatment	2	6	CL, YP	HoNOSCA
Gordon et al. (2011)	Study of predictors / moderators	1	1	YP	RADS
Gunlicks-Stoessel et al. (2010)	Study of predictors / moderators	1	1	CL	HAM-D
Iftene et al. (2015)	Compare interventions	3	5	CL, YP	—
Jacobs et al. (2009)	Study of predictors / moderators	1	2	CL, YP	CDRS-R; SIQ-Jr / SIQ
Jacobs et al. (2010)	Study of predictors / moderators	2	2	CL & PA	CPRS-R
Jiménez Chafey et al. (2011)	Study of predictors / moderators	4	7	CL, YP	—
Karlsson et al. (2008)	Study of predictors / moderators	1	1	CL	K-SADS
Kennard et al. (2009)	Study of predictors / moderators	2	3	CL	CGI-I; CDRS-R
Kennard et al. (2009)	Assess single intervention	3	4	*not stated*	CDRS-R; CGI-I
Kennard et al. (2009)	Assess single intervention	1	2	CL	CDRS-R
Le Noury et al. (2015)	Compare interventions	4	7	CL, YP, PA	HAM-D
Lewis et al. (2009)	Study of predictors / moderators	1	1	CL	CDRS-R
Lewis et al. (2010)	Study of predictors / moderators	1	2	CL, YP	CDRS-R

Short title	Study objective	Number of outcome domains covered	Number of outcome measures used	Informant(s) consulted	Primary outcome measure (if defined)
Lusk & Melnyk (2011)	Assess single intervention	4	8	YP & PA	—
Lynch et al. (2011)	Cost-effectiveness of an intervention	2	2	CL	CGI-I; CDRS-R
Manglick et al. (2013)	Study of predictors / moderators	2	2	CL	K-SADS
Mayes et al. (2007)	Study of predictors / moderators	2	4	CL	—
McMakin et al. (2012)	Study of predictors / moderators	1	2	CL	CDRS-R
Melvin et al. (2013)	Examine long-term outcomes	2	3	CL	—
Merry et al. (2012)	Assess single intervention	1	1	CL	CDRS-R
Mufson et al. (2014)	Study of predictors / moderators	2	6	CL, YP	—
Ngo et al. (2009)	Study of predictors / moderators	3	4	YP	—
Niederhofer & Klitzing (2011)	Assess single intervention	2	2	CL, YP	BDI; Salivary Melatonin
Niederhofer & Klitzing (2012)	Assess single intervention	2	3	CL, YP	BDI; Salivary Melatonin
Norton (2010)	Assess single interventions	1	—	—	—
O'Shea et al. (2015)	Compare interventions	3	7	CL, YP, PA	—
Parker et al. (2016)	Compare interventions	4	6	CL, YP	BDI; BAI; MADRS;
Pass et al. (2015)	Assess single intervention	4	5	YP & PA	—
Peters et al. (2016)	Examine long-term outcomes	2	5	CL, YP	CGAS; GAF
Poole et al. (2017)	Assess single intervention	4	6	*not stated*	—
Rengasamy et al. (2013)	Study of predictors / moderators	1	2	CL	A-LIFE
Riley (2011)	Assess single intervention	1	1	YP	TSCS-2 (short)
Rohde et al. (2008)	Assess intervention extension	2	2	CL	CGI-I
Rohde et al. (2014)	Assess combination treatment	2	2	CL	CDRS-R; TLFB
Rossello et al. (2008)	Compare interventions	3	4	YP & PA	CDI
Sakolsky et al. (2011)	Study of predictors / moderators	2	2	CL	CGI-I; CDRS-R
Schirman et al. (2010)	Assess single intervention	2	6	CL, YP	CGI-I
Shamseddeen et al. (2011)	Study of predictors / moderators	2	2	CL	CDRS-R; CGI-I

Short title	Study objective	Number of outcome domains covered	Number of outcome measures used	Informant(s) consulted	Primary outcome measure (if defined)
Shamseddeen et al. (2011)	Study of predictors / moderators	2	2	CL	CDRS-R; CGI-I
Shamseddeen et al. (2012)	Study of predictors / moderators	2	2	CL	CDRS-R; CGI-I
Shirk et al. (2008)	Study of predictors / moderators	1	1	CL, YP	—
Shon et al. (2014)	Assess single intervention	1	2	CL	CGI-I; CGI-S
Simons et al. (2012)	Study of predictors / moderators	2	5	CL, YP	CDRS-R
Spence et al. (2016)	Assess single intervention	3	6	CL, YP, PA	—
Spirito et al. (2009)	Study of predictors / moderators	2	2	CL	CDRS-R; CGI-I
Strandholm et al. (2014)	Study of predictors / moderators	1	1	CL	HAM-D
TADS (2007)	Assess single intervention	2	4	CL, YP	CDRS-R; CGI-I
TADS (2009)	Examine long-term outcomes	2	4	CL, YP	CDRS-R; CGI-I
Tao et al. (2009)	Study of predictors / moderators	1	1	CL	CDRS-R
Tao et al. (2010)	Assess single intervention	1	1	CL	CDRS-R
Trowell et al. (2007)	Compare interventions	3	5	CL, YP	—
Vitiello et al. (2011)	Examine long-term outcomes	2	6	CL, YP	CGI-I; CDRS-R; A-LIFE; CGAS
Weisz et al. (2009)	Assess transportability of an intervention	2	5	CL, YP, PA	—
Wells et al. (2012)	Assess single intervention	1	1	YP	CES-D
Wiggins et al. (2010)	Assess routine service performance	1	2	CL	—
Wilkinson & Goodyer (2008)	Assess combination treatment	2	2	YP	RDQ
Wilkinson et al. (2009)	Study of predictors / moderators	1	1	CL	—
Yang et al. (2016)	Assess single intervention	3	6	CL, YP	Attentional bias score; HAM-D
Zhand et al. (2015)	Assess single intervention	2	2	CL, YP	BDI; CGI-I

Note. A-LIFE = Adolescent Longitudinal Interval Follow-up Evaluation (Keller et al., 1987); BAI = Beck Anxiety Inventory (Leyfer et al, 2006); BDI = Beck Depression Inventory, 1st or 2nd version (A. T. Beck et al., 1988); CDI = Children's Depression Inventory, 1st or 2nd version (Kovacs, 1992); CDRS-R = Children's Depression Rating Scale Revised (Poznanski & Mokros, 1996); CES-D = Centre for Epidemiologic Studies Depression Scale (Radloff, 1977); CGAS = Children's Global Assessment Scale (Dyrborg et al., 2000); CGI-I = Clinical Global Impressions Scale IMPROVEMENT (Busner & Targum, 2007); CGI-S = Clinical Global Impressions Scale SEVERITY (Busner & Targum, 2007); CL = clinician; GAF = Global Assessment of Functioning (Endicott et al., 1976); CPRS-R = Conners'

Parent Rating Scale–Revised (Conners et al., 1998); HAM-D = Hamilton Rating Scale for Depression (Hamilton, 1986); HoNOSCA = Health of the Nation Outcome Scale (Gowers et al., 1999); K-SADS = Kiddie-SADS-Present and Lifetime Version (Kaufman et al., 1997); MADRS = Montgomery-Asberg Depression Rating Scale (Montgomery & Asberg, 1979); MFQ = Mood and Feelings Questionnaire short or long version (Angold et al., 1995); PA = parent; YP = young person; RADS = Reynolds Adolescent Depression Scale (W. M. Reynolds, 1987); RDQ = Responses to Depression Questionnaire (Nolen-Hoeksema & Morrow, 1991); SIQ-Jr / SIQ = Suicidal Ideation Questionnaire Junior (W. M. Reynolds, 1988); TFBM = Timeline Follow Back Method (Sobell et al., 1988); TSCS-2 (short) = Tennessee Self-Concept Scale version 2 short form (Horwitz et al., 2001).

Table B.6. Overview of Outcome Measures Reported by Outcome Domain

Outcome measure		Reporter	Number of studies using this outcome measure
Symptoms			
Depression symptoms and diagnosis			
A-LIFE	Adolescent Longitudinal Interval Follow-up Evaluation (Keller et al., 1987)	CL	5
BDI (any)	Beck Depression Inventory, 1st and 2nd version (A. T. Beck et al., 1988)	YP	16
CDI	Children's Depression Inventory, 1st and 2nd version (Kovacs, 1992)	YP or PA	10
C-DISC Subscale[a]	Computerized Diagnostic Interview Scale for Children (Shaffer et al., 2000)	CL	1
CDRS-R	Children's Depression Rating Scale (Poznanski & Mokros, 1996)	CL	45
CES-D	Centre for Epidemiologic Studies Depression Scale (Radloff, 1977)	YP	4
DISC_IV	Diagnostic Interview Schedule for Children (Shaffer et al., 2000)	CL	3
DSD	DSM Scale for Depression (American Psychiatric Association, 1994)	YP	1
HAM-D	Hamilton Rating Scale for Depression (Hamilton, 1986)	CL	10
K-SADS	Kiddie-SADS-Present and Lifetime Version(Kaufman et al., 1997)	CL	14
MADRS	Montgomery-Asberg Depression Rating Scale (Montgomery & Asberg, 1979)	CL (YP available)	2
MFQ	Mood and Feelings Questionnaire, short or long version (Angold et al., 1995)	YP	6
POMS-SF	Profile of Mood States (Dilorenzo et al., 1999)	YP	1
RADS	Reynolds Adolescent Depression Scale (W. M. Reynolds, 1987)	YP	6
RCADS	Revised Child Anxiety and Depression Scale (Chorpita et al., 2000)	YP or PA	1
SCID	Structured Clinical Interview for DSM-IV Axis I Disorders (First et al., 1997)	CL	3
SDQ emotion	Strengths and Difficulties Questionnaire (R. Goodman, 2001)	YP or PA	2
YSR Internalizing	Youth Self Report - Subscale for internalizing problems (Achenbach, 1991a)	YP	1
General psychopathology or emotional problems			
CBCL	Child Behaviour Checklist (Achenbach, 1991b)	YP and/or PA	4
GHQ-28	General Health Questionnaire (Goldberg, 1972)	YP	1
SDQ	Strengths and Difficulties Questionnaire (R. Goodman, 2001)	YP and/or PA	2

Outcome measure		Reporter	Number of studies using this outcome measure
Self-harm			
RTSHIA	Risk-Taking and Self-Harm Inventory for Adolescents (Vrouva et al., 2010)	YP	1
Suicidality			
C-SSRS	Columbia–Suicide Severity Rating Scale (Posner et al., 2011)	CL[a]	1
CDRS-R suicidality	Children's Depression Rating Scale Subscale (Poznanski & Mokros, 1996)	CL	2
K-SADS suicidality	K-SADS-PL items related to suicidality and self-harm (Kaufman et al., 1997)	CL	2
SIQ-Jr / SIQ	Suicidal Ideation Questionnaire (W. M. Reynolds, 1988)	YP	10
SPS	Suicide Probability Scale (Cull & Gill, 1988)	YP	1
SSI	Scale for Suicidal Ideation (A. T. Beck et al., 1979)	YP	1
Anger, conduct, impulsiveness			
BANI	Beck Youth Anger Inventory (J. S. Beck et al., 2005)	YP	2
ABQ	Anti-Social Behaviour Questionnaire (Goodyer et al., 2017)	YP	1
BDBI	Beck Youth Destructive Behaviour Inventory (J. S. Beck et al., 2005)	YP	1
BIS-11	Barratt Impulsiveness Scale 11th version (J. H. Patton et al., 1995)	YP	1
CPRS-R	Conners' Parent Rating Scale–Revised (Conners et al., 1998)	PA	1
HSQ	Home Situations Questionnaire (Altepeter & Breen, 1989)	PA	1
SDQ conduct	SDQ subscale for conduct / behaviour problems (R. Goodman, 2001)	YP and/or PA	1
Anxiety			
BAI	Beck Anxiety Inventory (Leyfer et al., 2006)	YP	2
RCMAS	Revised Children's Manifest Anxiety Scale (C. R. Reynolds & Richmond, 1978)	YP	2
SCAS	Spence Children's Anxiety Scale (Spence, 1998)	YP	1
STAI	State Trait Anxiety Inventory (Spielberger, 1983)	YP and/or PA	2
Eating disorder			
EAT-26	Eating Attitudes Test 26 (Garner et al., 1982)	YP	1

Outcome measure		Reporter	Number of studies using this outcome measure
Sleeping problems			
Carskadon Sleep Survey	Carskadon Sleep Survey (Carskadon, 2004)	YP	1
ISI	Insomnia Severity Index (Morin, 1993)	YP	1
K-SADS sleep	K-SADS-PL items related to insomnia and hypersomnia (Kaufman et al., 1997)	CL	1
Alcohol and drug use			
AUDIT Subscale[a]	Alcohol Use Disorders Identification Test (Saunders et a., 1993)	Mainly CL[a]	1
CRAFFT	CRAFFT Screening interview (Knight et al., 2002)	Mainly CL[a]	1
DSM alcohol count	DSM alcohol count (abuse and dependency)	Derived	2
SACS	Substance and Choices Scale (Christie et al., 2007)	YP	1
TLFB	Timeline Follow Back Method (Sobell et al., 1988)	CL	3
Other comorbidities			
K-SADS comorbidities	K-SADS-PL to assess comorbidities (Kaufman et al., 1997)	CL	2
LOI	Leyton Obsessional Inventory (Bamber et al., 2002)	YP	1
Self-management			
Cognition and behaviour			
Attentional bias score	Attentional bias score (Yang et al., 2016)	CL	1
ATQ	Automatic Thoughts Questionnaire (Fischer & Corcoran, 2007)	CL	2
BHS	Beck Hopelessness Scale (A. T. Beck & Steer, 1993)	YP	1
CASQ-R	Children’s Attributional Style Questionnaire – Revised (Thompson et al., 1998)	CL	1
CTI-C	The Cognitive Triad Inventory for Children (Kaslow et al., 1992)	CL	1
DAS	Dysfunctional Attitudes Scale (Rogers et al., 2009)	YP	3
DEQ-A	The Depressive Experiences Questionnaire for Adolescents, short (Blatt et al., 1992)	CL	1
RDQ	Responses to Depression Questionnaire (Nolen-Hoeksema & Morrow, 1991)	CL	1
RRS	Rumination Response Style (Nolen-Hoeksema & Morrow, 1991)	CL	1

Outcome measure		Reporter	Number of studies using this outcome measure
Behavioural and physical activation			
AAS	Active Australia Survey (Australian Institute of Health and Welfare, 2003)	YP	1
BADS-SF	Behaviour Activation for Depression Scale – Short Form (Manos et al., 2011)	YP	1
LTEQ	Leisure Time Exercice Questionnaire (Godin, 2011)	YP	1
PES	Pleasant Events Schedule (Clarke et al., 1990)	YP	1
Coping			
PBS	Personal Beliefs Scale – Teens (Melnyk et al., 2009)	YP	1
PSI	Problem Solving Inventory (Heppner & Petersen, 1982)	YP	1
Functioning			
Global functioning			
BASC-2	The Behavioural Assessment System for Children-2 (C. R. Reynolds, 2010)	YP and/or PA	1
CGAS	Children's Global Assessment Scale (Dyrborg et al., 2000)	CL	13
CGI-I	Clinical Global Impressions Scale IMPROVEMENT (Busner & Targum, 2007)	CL	32
CGI-S	Clinical Global Impressions Scale SEVERITY (Busner & Targum, 2007)	CL	16
DMs	Developmental Milestones (Peters et al., 2016)	YP	1
GAF	Global Assessment of Functioning (Endicott et al., 1976)	CL	2
HoNOSCA	Health of the Nation Outcome Scale (Gowers et al., 1999)	CL	6
IRS	Impairment Rating Scale (Fabiano et al., 2006)	PA	1
OADSI	Outcome of Adolescent Depression Structured Interview (OADSI) (Dudley et al., 2005)	CL	1
ORS	Outcome Rating Scale (Miller et al., 2003)	YP and/or PA	1
SOFAS	Social and Occupational Functional Assessment Scale (Goldman et al., 1992)	CL	1
Social functioning			
SSQ	Social Skills Questionnaire (Spence, 1995)	YP and/or PA	1
SAS-SR	Social Adjustment Scale (Self-Report) (Weissman & Bothwell, 1976)	YP	1
SASCA	Social Adjustment Scale for Children and Adolescents (Beiser, 1990)	YP	1

Outcome measure		Reporter	Number of studies using this outcome measure
Executive functioning			
D-REF	Delis Rating of Executive Functioning(Delis, 2012)	PA	1
Relationships			
Family functioning			
BFAM	Brief Family Assessment Measure (Skinner et al., 1995)	YP	1
BISC	Beavers Interactional Competence Scale (Beavers & Hampson, 1990)	CL	1
BISS	Beavers Interactional Style Scale (Beavers & Hampson, 1990)	CL	1
CBQ-20	Conflict Behaviour Questionnaire (Robin & Foster, 1989)	YP and/or PA	1
FAD	Family Assessment Device (Epstein et al., 1983)	YP and/or PA	1
FEICS	Family Emotional Involvement and Criticism Scale (Shields et al., 1992)	YP and/or PA	1
PRQ	Parenting Relationship Questionnaire (Kamphaus & Reynolds, 2006)	PA	1
Attachment style			
ASQ	Attachment Style Questionnaire (Feeney et al., 1994)	YP	1
Peer relationships			
SCPQ	Social Competence with Peers Questionnaire (Spence, 1995)	YP and/or PA	1
Personal Growth			
Assertiveness			
SIB	Scale for Interpersonal Behaviour (Arrindell & van der Ende, 1985)	YP	1
Autonomy			
AFC	Autonomous Functioning Checklist (Sigafoos et al., 1988)	YP	1

Outcome measure		Reporter	Number of studies using this outcome measure
Self-Concept			
BSCI	Beck Self-Concept Inventory for Youth (J. S. Beck et al., 2005)	YP	1
PHCSC	Piers Harris Children's Self-concept scale (Piers et al., 1969)	YP	3
TSCS-2 (short)	Tennessee Self-Concept Scale version 2 – short form (Horwitz et al., 2001)	YP	1
Self-esteem			
SES	Rosenberg Self-esteem scale (Rosenberg, 1965)	YP	1
SPPA	Self-Perception Profile Adolescents (Wichstrom, 1995)	YP	1
Service quality, and satisfaction			
Use of other services			
SACA	Service Assessment for Children and Adolescents (Horwitz et al., 2001)	PA	1
CASA	Child and Adolescent Services Assessment (Ascher et al., 1996)	YP and/or PA	2
Additional Service use	Use of external / addition mental health services	YP	2
Therapeutic alliance and process			
TAS	Therapeutic Alliance Scale (Eskin et al., 2008)	YP	1
TASC	Therapeutic Alliance Scale for Children(Shirk & Saiz, C. S., 1992)	YP and/or PA	1
Treatment retention			
Treatment retention	*Non-standardised measure of treatment retention*	N/A	1
Treatment Content			
Quiz	*Non-standardised quiz about treatment content*	YP and/or PA	1
Client satisfaction			
SRS	Session Rating Scale (Duncan et al., 2003)	YP	1
Evaluation questionnaire	*Non-standardised evaluation questionnaire*	YP and/or PA	1
Satisfaction	*Satisfaction with mental health care received (non-identified scale)*	YP	1

Outcome measure		Reporter	Number of studies using this outcome measure
Wellbeing			
Health-related quality of life			
EQ-5D	The EuroQol group (Byford, 2013)	YP	3
MCS-12	Mental Component Summary Scale (Ware et al., 1996)	YP	1
PEDS-QL	Pediatric Quality of Life Inventory (Varni et al., 2001)	YP and/or PA	1
PQ-LES-Q	Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (Endicott et al., 2006)	YP	1
SIP	Sickness Impact Profile (Bergner et al., 1981)	YP	1
Physical health			
Biomarkers			
Cortisol	Cortisol	CL	1
Norepinephrine	Norepinephrine	CL	1
Salivary Melatonin	Salivary melatonin	RS / CL	2
Serotonin	Serotonin	CL	1
Parental support and wellbeing			
Parental wellbeing			
BDI (any)	Beck Depression Inventory, 1st and 2nd version (A. T. Beck et al., 1996)	PA	1
DASS-21	Depression Anxiety Stress Scale(Lovibond & Lovibond, 1995)	PA	1
Idiographic measures			
Self-defined goals	*Idiographic self-defined goals*	YP	1
TCC	Target Complaint Checklist (TCC; Elkin et al., 1985)	YP and/or PA	1

Note. CL = Clinician; PA = parent; RS = researcher; YP = young person.

[a] YP self-report available.

Appendix C. Supplementary Materials – Chapter 5

Table C.1. Code Descriptions, Definitions, and Illustrative Quotes

Outcome domain	Outcome category	Description	Example quote
Symptom change	Mood and affect	YP are less low and depressed, happier and more cheerful, less prone to mood swings, less withdrawn.	“She seemed to be happy, you seemed to be able to have a joke with her and a laugh with her and ... she was just generally happier in herself.” (PA #21)
		YP return to be the person they used to be; or appear to be a new person.	“If I hadn’t gone there, I think I’d be a completely different person, like I’d still be really low now.” (Natalie, 15 yrs, STPP)
		Low mood and negative affect are more fleeting, less overwhelming (often linked to YP coping)	“A lot is different with (name). She’s sort of made a full recovery if you like.” (Therapist of Louise, 17 yrs), CBT)
	Anger and aggression	YP are less angry, irritable or aggressive, less prone to outburst, and better able to manager their temper.	“I would BREAK things, and I would like break my stuff that I really like […] but I don’t do that now at all, I just cry, that’s it.” (Jenny, 17 yrs, STPP)
	Appetite	YP have a healthier appetite and weight.	“I started eating properly.” (Gemma, 16 yrs, BPI)
	Sleeping and energy	YP have healthier sleep patterns and energy levels.	“working on the sleep was a bit of a quick win.” (Therapist of Connor, 18 yrs, CBT)
	Self-harm	YP engage less in self-harm (e.g., cutting, trichotillomania).	“Well, certainly all the self-harm and thoughts of self-harm… cleared up within the first… 5, 6 weeks” (Therapist of Jenny, 17 yrs, STPP)
	Suicidality	Reduced suicidal ideation and behaviour.	“First of all, I don’t feel suicidal anymore um and like to me I think that is the greatest improvement.” (Natalie, 15 yrs, STPP)
	Anxiety	YP experience fewer fears, worries, panic attacks, or social anxiety; they engage in activities they previously avoided.	“I don’t feel as anxious like I haven’t had a panic attack since November or something.”(Poppy, 18 yrs, BPI)
	Other comorbid issues	Improvements in less frequently reported comorbid problems such as substance use or obsessive-compulsive symptoms.	“The Class A’s and stuff that’s very rare, that’s sort of a party thing now.” (Dylan, 16 yrs, STPP)
Cognition, coping and self-management	Behavioural activation	YP become more active, return to hobbies or engage in new activities, including volunteering or work experience; YP gain a sense of purpose, routine and structure.	“The activity planning […] I think that’s the thing that helped the most […] just not sitting around doing nothing all the time but instead say actually being more sociable and... kind of getting back to... to the person I was before.” (Charlotte, 14 yrs, CBT)

Outcome domain	Outcome category	Description	Example quote
	Coping and resilience	YP learn specific coping strategies (e.g., breathing or counting exercises). YP have a better understanding of their feelings and thoughts and how these link to behaviours; YP can anticipate and manage challenges. YP are more resilient, with greater self-efficacy, and sense of control.	"If you're nervous before an exam she says take deep breaths, do the 10 seconds breath exercise that we did […] It actually helps." (Madeline, 15 yrs, CBT). "It did wake me up to how my o- sort of how it all works and like how my brain works and how to sort of react to certain things […] The fact that if you can understand something you can fix something, that's my motto. So, if I can understand like in a computer game, if I can understand why it's not working, I can fix the problem." (Dylan, 16 yrs, STPP).
	Cognition and behaviour	YP can challenge negative automatic thoughts, imagine alternatives, and approach situations differently, with more flexible thinking styles.	"I'm now able to think it through logically and think about different peoples' point of view and stick it together and be more diplomatic about things and yeah just react to situations differently." (Athena, 19 yrs, STPP)
Functioning	Global functioning	YP 'function' better across a range of life domains and are able to engage in activities considered typical for adolescence.	"I guess I can sort of just do a bit more than I used to be able to… slightly more focused and erm just feeling a bit better as a whole." (Stuart, 15 yrs, STPP)
	Executive functioning	YP are better able to get things done, due to improved concentration, motivation, planning and organisation.	"I remember her giving me a sheet where, cause I have a very disorganised personality and during this time, school wise, I need that, I need that so my grades don't fail completely. So, she tried to help me have a more organised personality. That didn't really work, but it got me more thinking of how organised I should be." (Adrian, 16 yrs, CBT)
	Academic and vocational functioning	YP attend school more regularly after frequent absence or a period of leave caused by the depression. YP work more effectively in school and achieve better results.	"If she'd have carried on being like she was last year at school there's no way she would get the grades that they are predicting." (Parent #14, STPP). "She's staying in school, so she's actually attending classes which she wasn't, she was coming home too much, or sitting in the office not doing work, so she actually started doing lessons, she was paranoid about her GCSEs that she wouldn't get any and she's done the first year and she's been predicted very high grades." (Mother of Ella, 15 yrs, BPI)
	Social functioning	YP are more outgoing and talkative, more present within friendship groups and more socially connected. YP find it easier to make conversation, relate to others, are more approachable.	"Linking up with friends, I mean this was something that we worked on quite a lot: Could she bear to actually link up with people that she may not know that well just for the sake of having somebody to go in the lunch queue with." (Therapist of Jenny, 17 yrs, STPP)

Outcome domain	Outcome category	Description	Example quote
Personal growth	Assertiveness	YP are better able to stand up for their needs and opinions, overcome the urge to please, stand by their preferences and desires, and can express disagreement or disapproval where appropriate.	"I guess it's almost thinking more about myself instead of trying to please other people and try to be like that, diplomatic to remember that my health is the most important thing." (Athena, 19 yrs, STPP)
	Autonomy and responsibility	YP are more independent and able to take responsibility for their lives and actions.	"When she speaks to her psychologist, maybe she gets advised that she is a grown young woman now and that she can make her own choices." (Mother of Mikayla, 15 yrs, STPP)
	Identity	YP find out who they are and how to be themselves around other people; less idealised self-images that can accommodate both positive and challenging personality traits; positive and negative feelings.	"I think that-that strengthened her enormously it actually gave her a much more rounded experience of herself... erm that she could acknowledge that she had erm... the bad and the good the sort of the kind parts and the... selfish parts of her and that-that they all went to make up (name of adolescent) the girl that she was." (Therapist of Jenny, 17 yrs, STPP).
	Processing past and present	YP are able to make sense of challenging past or ongoing experiences such as bereavement, parental divorce, or family conflict.	"I mean obviously, you know, the loss and the pain and all of that are still there but it became much more manageable for (name) and it wasn't really-, I think the thing that had sort of troubled her the most was how alienated it had left her feeling and the conversations kind of healed that to an extent." (Therapist of Louise, 17 yrs, CBT)
	Confidence and self-esteem	YP feel more confident within themselves, less insecure and vulnerable to the judgement of others, have higher self-regard.	"I began feeling more confident about myself and I realise I'm not really as much of an idiot as I thought it was." (Priya, 14 yrs, BPI)
Relationships	Ability to talk	YP feel more able to talk about feelings and thoughts, which helps deepen relationships; having a stronger support network also facilitates opening up.	"I think cause like I've talked about it s-so much that I'm still trying to find a way to um uh express myself correctly because before I had a problem I couldn't e-express myself or tell people how I felt. But now I think like that's something I've really improved on but I'm still improving on." (Poppy, 18 yrs, BPI)
	Family functioning and relationships	YP get on better with their family: less conflict as YP cope better and families understand them better; easing of entrenched tensions between family members; families communicate more openly; YP clarify their role within the family system.	"[Mother and daughter] come across like they get on much more and there is sort of better attunement between the two, a lighter, a lighter emotional quality to, to their relationship. Whereas before it was very sort of anxious, suspicious sort of just almost just not knowing each other." (Therapist of Louise, 17 yrs, CBT)
	Friendships	Reactivation or deepening of existing friendships, expanding friendship groups or changing friends by turning towards more supportive friendships.	"I think I've just changed the people that are around me and that's helped a lot. I think erm…yeah, it's a lot more grown-up people and more people that are a lot better at supporting me than previously I was friends with, which I think is very helpful." (Athena, 19 yrs, STPP).

Outcome domain	Outcome category	Description	Example quote
Relationships *(continued)*	Peer relationships	Getting on better with peers in school (e.g., less bullying) or a romantic partner.	"I'm just living life really apart from the fact that I'm not bullied anymore." (Priya, 14 yrs, BPI).
Therapeutic process	Therapeutic space	YP could speak about feelings and thoughts without feeling judged, and without having to worry about repercussions in their daily lives.	"I feel she unders-, understands me and she […] listens carefully to me you know. The things like that really help you know, just sometimes it's nice just to be heard. (Poppy, 18 yrs, BPI)
		Clinician offered new perspectives and encouraged YP to think about things differently.	"back then I felt like nobody cared about me and I don't I think it made me feel good within myself because it was just it's kinda what I needed like to feel like someone (breathes out) does care and that like they are there for me" (Natalie, 15 yrs STPP).
		YP was able to unload feelings and thoughts rather than bottling them up.	
		YP felt listened to and cared for.	
	Practical or systemic progress	TPs helped solve practical problems by engaging with different parts of the system, including parents, school, general practitioner, etc.	"... we were also having link up with the doctor because the doctor thought that (name) might have some hormonal problems... so we picked up on that … and liaison with school college… working with the family... so... sort of looking at all the different systems…" (Therapist of Gemma, 16 yrs, BPI)
Wellbeing	Peace of mind	YP feel calmer, more balanced, relaxed, and carefree; YP fee as if a weight had been lifted off their shoulders; YP are more accepting of things they cannot change.	"She looked younger than 14 and it was almost like she was a little child carrying the world on her shoulders but as erm she seemed less burdened she was able to be in a way just be herself more." (Therapist of Priya, 14 yrs, BPI)
	Optimism	YP have a more positive and optimistic outlook into their lives and the future.	"I guess people probably notice that I've been a bit more optimistic about things." (Adrian, 16 yrs, CBT)
	Future orientation	YP can make plans for the future and have goals (e.g., applying for college or university, saving money to travel abroad)	"I made new friends started planning summer… yeah just getting excited for things, I dunno, just making plans I think was like… coz when you're really down you don't make plans." (Ada, 18 yrs, STPP)
Parental support and wellbeing	Parental support	PAs are better able to understand their child's difficulties and more aware of how their parenting practices may contribute to these difficulties. Parents learn to support and parent their child more effectively.	"I think we're a bit more self-aware. I'm not saying we don't make the same mistakes I think we still make mistakes perhaps about what we do... but I think we're more aware when we do it now... and then we try-we try and stop." (Mother of Jenny, 17 yrs, STPP).
	Parental wellbeing	PA feel less guilty, isolated, stressed and worried; PA feel reassured, supported, and able to express their own frustrations and issues.	"As a mother… not having to be totally responsible for fixing and sorting something out… that is too big for you… urm so just... yes entering the system and... being able to relinquish some of that responsibility is positive in itself." (Mother of Hayley, 16 yrs, STPP).

Note. Cont'd = continued; PA = parents; CL = clinician; YP = young people; yrs = years.

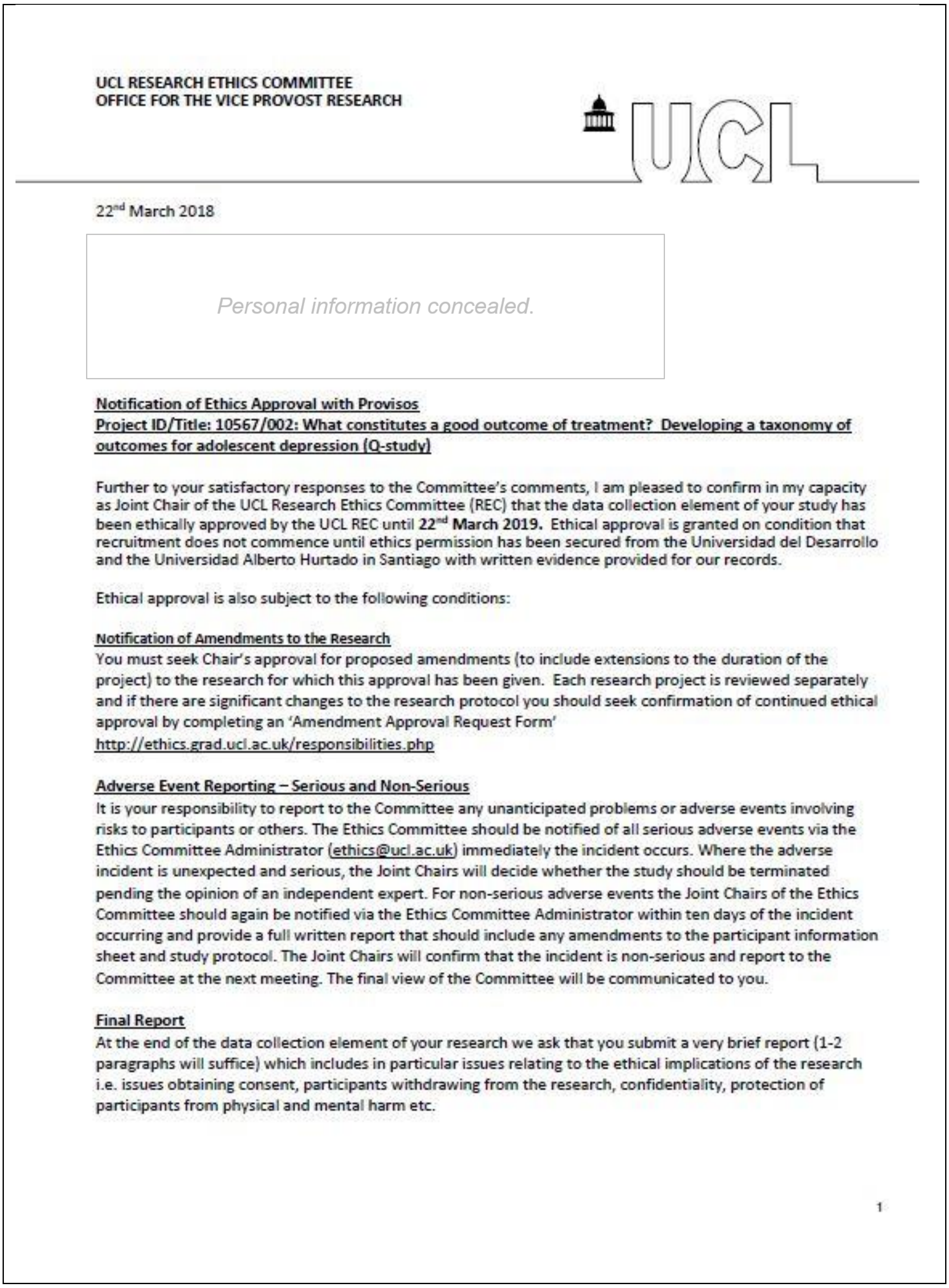
Table C.2. Salience of Outcomes Across Participant Groups and Treatment Arms

Outcome domain and subdomain	**Full sample (*n* = 102)**		**Adolescents (*n* = 34)**		**Parents (*n* = 34)**		**Clinicians (*n* = 34)**		**CBT (*n* = 27)**		**STPP (*n* = 48)**		**BPI (*n* = 27)**	
	n	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
Symptom change	81	79%	28	82%	26	77%	27	79%	24	90%	35	73%	22	82%
Mood & Affect	66	65%	22	65%	21	62%	23	68%	22	82%	27	56%	17	63%
Anger and aggression	17	17%	8	24%	8	24%	1	3%	4	15%	9	19%	4	15%
Eating and weight	13	13%	3	9%	6	18%	4	12%	3	11%	4	8%	6	22%
Sleeping and energy	20	20%	5	15%	8	24%	7	21%	6	22%	7	15%	7	26%
Self-harm	12	12%	4	12%	3	9%	5	15%	3	11%	6	13%	3	11%
Suicidality	14	14%	8	24%	2	6%	4	12%	1	4%	8	17%	5	19%
Anxiety	12	12%	4	12%	3	9%	5	15%	4	15%	1	2%	7	26%
Other comorbid issues	4	4%	2	6%	2	6%	0	0%	0	0%	2	4%	2	7%
Self-management	62	61%	24	71%	20	59%	18	53%	20	74%	25	52%	17	63%
Behavioural activation	20	20%	6	18%	6	18%	8	24%	7	26%	4	8%	9	33%
Coping and resilience	51	50%	22	65%	17	50%	12	35%	16	59%	23	48%	12	44%
Cognition and behaviour	19	19%	8	24%	6	18%	5	15%	12	44%	4	8%	3	11%
Functioning	67	66%	19	56%	26	77%	22	65%	24	89%	26	54%	17	63%
Global functioning	9	9%	2	6%	3	9%	4	12%	0	0%	4	8%	5	19%
Executive functioning	20	20%	8	24%	8	24%	4	12%	9	33%	9	19%	2	7%
Academic and vocational functioning	46	45%	10	29%	18	53%	18	53%	15	56%	19	40%	12	44%
Social functioning	36	35%	12	35%	14	41%	10	29%	13	48%	13	27%	10	37%
Personal Growth	60	59%	18	53%	21	62%	21	62%	14	52%	31	65%	15	56%
Assertiveness	13	13%	4	12%	3	9%	6	18%	1	4%	8	17%	4	15%
Autonomy and responsibility	16	16%	3	9%	10	29%	3	9%	3	11%	9	19%	4	15%
Identity	14	14%	1	3%	2	6%	11	32%	3	11%	8	17%	3	11%
Processing past and present	18	18%	6	18%	5	15%	7	21%	7	26%	6	13%	5	19%
Confidence and self-esteem	34	33%	4	12%	3	9%	6	18%	7	26%	16	33%	11	41%
Relationships	63	62%	21	62%	22	65%	20	59%	20	74%	26	54%	17	63%
Ability to talk	13	13%	4	12%	8	24%	1	3%	3	11%	6	13%	4	15%
Family functioning and relationships	50	49%	16	47%	17	50%	17	50%	18	67%	18	38%	14	52%
Friendships	28	27%	11	32%	9	27%	8	24%	10	37%	11	23%	7	26%
Other peer relationships	7	7%	5	15%	0	0%	4	9%	0	0%	1	4%	6	26%
Therapeutic process	54	53%	21	62%	17	50%	16	47%	15	56%	26	54%	13	48%
Therapeutic alliance	48	47%	21	62%	17	50%	10	29%	14	52%	24	50%	10	37%
Practical or systemic progress	9	9%	2	6%	1	3%	6	18%	2	7%	2	4%	5	19%

Outcome domain and subdomain	**Full sample (*n* = 102)**		**Adolescents (*n* = 34)**		**Parents (*n* = 34)**		**Clinicians (*n* = 34)**		**CBT (*n* = 27)**		**STPP (*n* = 48)**		**BPI (*n* = 27)**	
	n	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
Wellbeing	39	38%	10	29%	18	53%	11	32%	13	48%	16	31%	10	37%
Peace of mind	14	14%	2	6%	9	27%	3	9%	6	22%	6	13%	2	7%
Optimism and hope	12	12%	5	15%	4	12%	3	9%	6	22%	4	8%	2	7%
Future orientation	19	19%	4	12%	9	27%	6	18%	4	15%	9	19%	6	22%
Parental support and wellbeing	24	24%	3	9%	16	47%	5	15%	2	7%	1	25%	1	37%
Parental support	7	7%	0	0%	6	18%	1	3%	2	0%	12	13%	10	4%
Parental wellbeing	22	22%	3	9%	14	41%	5	15	0	7%	6	23%	1	33%

Appendix D. Supplementary Materials – Chapter 6

Figure D.1. UCL Ethics Approval Letter



UCL RESEARCH ETHICS COMMITTEE
OFFICE FOR THE VICE PROVOST RESEARCH

22nd March 2018

Personal information concealed.

Notification of Ethics Approval with Provisos
Project ID/Title: 10567/002: What constitutes a good outcome of treatment? Developing a taxonomy of outcomes for adolescent depression (Q-study)

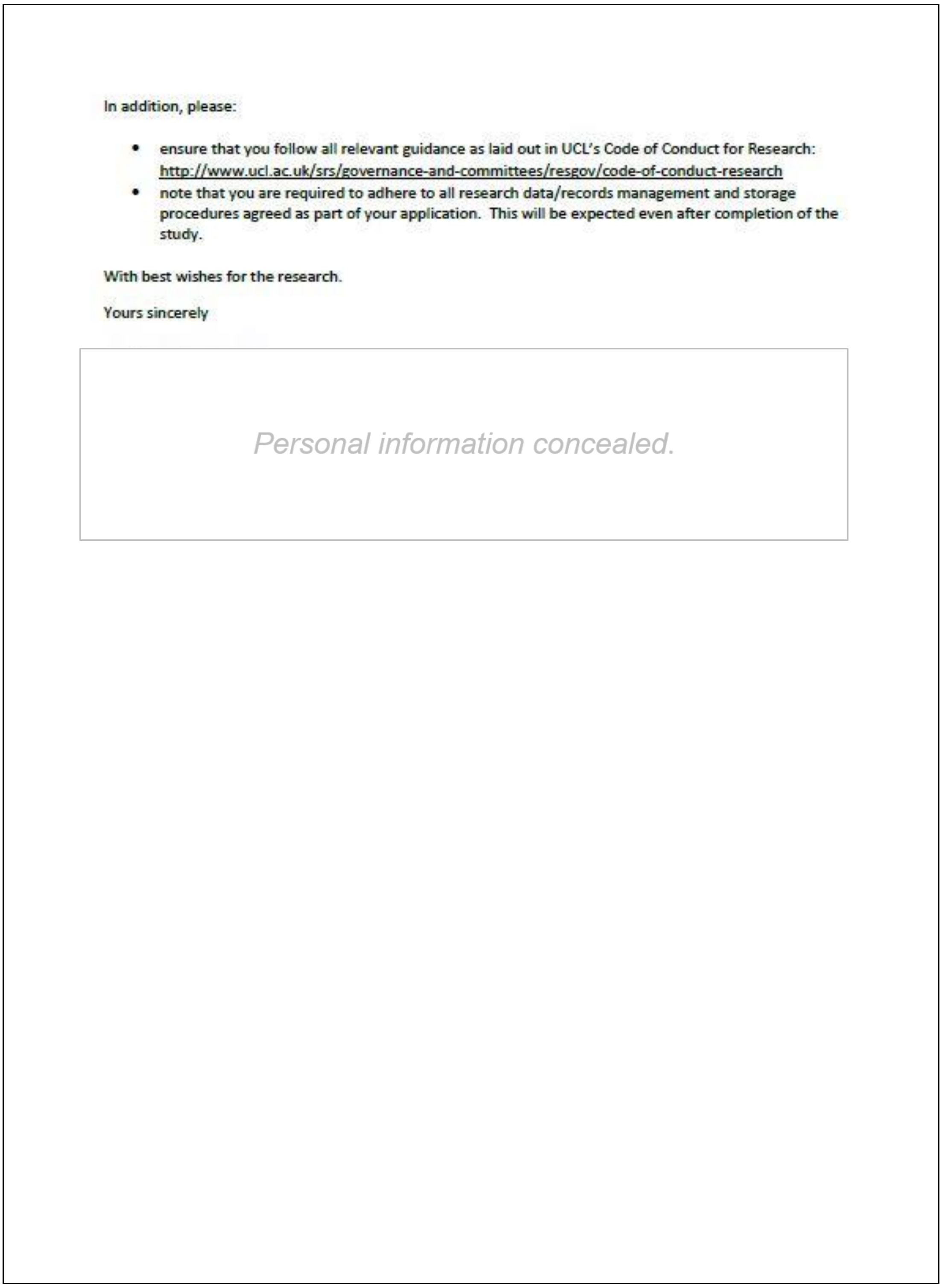
Further to your satisfactory responses to the Committee's comments, I am pleased to confirm in my capacity as Joint Chair of the UCL Research Ethics Committee (REC) that the data collection element of your study has been ethically approved by the UCL REC until **22nd March 2019.** Ethical approval is granted on condition that recruitment does not commence until ethics permission has been secured from the Universidad del Desarrollo and the Universidad Alberto Hurtado in Santiago with written evidence provided for our records.

Ethical approval is also subject to the following conditions:

Notification of Amendments to the Research
You must seek Chair's approval for proposed amendments (to include extensions to the duration of the project) to the research for which this approval has been given. Each research project is reviewed separately and if there are significant changes to the research protocol you should seek confirmation of continued ethical approval by completing an 'Amendment Approval Request Form'
http://ethics.grad.ucl.ac.uk/responsibilities.php

Adverse Event Reporting – Serious and Non-Serious
It is your responsibility to report to the Committee any unanticipated problems or adverse events involving risks to participants or others. The Ethics Committee should be notified of all serious adverse events via the Ethics Committee Administrator (ethics@ucl.ac.uk) immediately the incident occurs. Where the adverse incident is unexpected and serious, the Joint Chairs will decide whether the study should be terminated pending the opinion of an independent expert. For non-serious adverse events the Joint Chairs of the Ethics Committee should again be notified via the Ethics Committee Administrator within ten days of the incident occurring and provide a full written report that should include any amendments to the participant information sheet and study protocol. The Joint Chairs will confirm that the incident is non-serious and report to the Committee at the next meeting. The final view of the Committee will be communicated to you.

Final Report
At the end of the data collection element of your research we ask that you submit a very brief report (1-2 paragraphs will suffice) which includes in particular issues relating to the ethical implications of the research i.e. issues obtaining consent, participants withdrawing from the research, confidentiality, protection of participants from physical and mental harm etc.



In addition, please:

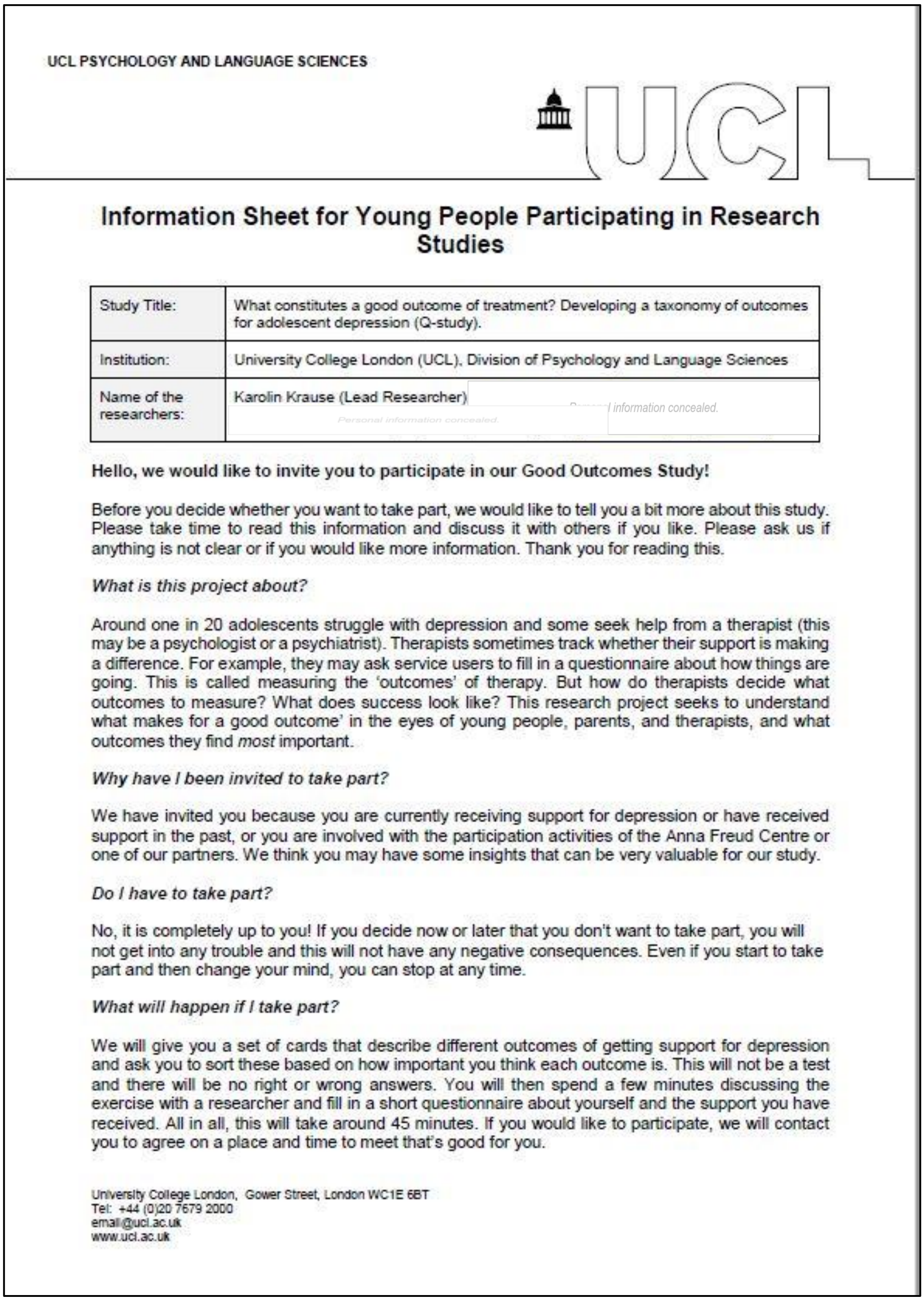
- ensure that you follow all relevant guidance as laid out in UCL's Code of Conduct for Research: http://www.ucl.ac.uk/srs/governance-and-committees/resgov/code-of-conduct-research
- note that you are required to adhere to all research data/records management and storage procedures agreed as part of your application. This will be expected even after completion of the study.

With best wishes for the research.

Yours sincerely

Personal information concealed.

Figure D.2. Participant Information Sheet for Young People



UCL PSYCHOLOGY AND LANGUAGE SCIENCES

Information Sheet for Young People Participating in Research Studies

Study Title:	What constitutes a good outcome of treatment? Developing a taxonomy of outcomes for adolescent depression (Q-study).
Institution:	University College London (UCL), Division of Psychology and Language Sciences
Name of the researchers:	Karolin Krause (Lead Researcher) [redacted]

Hello, we would like to invite you to participate in our Good Outcomes Study!

Before you decide whether you want to take part, we would like to tell you a bit more about this study. Please take time to read this information and discuss it with others if you like. Please ask us if anything is not clear or if you would like more information. Thank you for reading this.

What is this project about?

Around one in 20 adolescents struggle with depression and some seek help from a therapist (this may be a psychologist or a psychiatrist). Therapists sometimes track whether their support is making a difference. For example, they may ask service users to fill in a questionnaire about how things are going. This is called measuring the 'outcomes' of therapy. But how do therapists decide what outcomes to measure? What does success look like? This research project seeks to understand what makes for a good outcome' in the eyes of young people, parents, and therapists, and what outcomes they find *most* important.

Why have I been invited to take part?

We have invited you because you are currently receiving support for depression or have received support in the past, or you are involved with the participation activities of the Anna Freud Centre or one of our partners. We think you may have some insights that can be very valuable for our study.

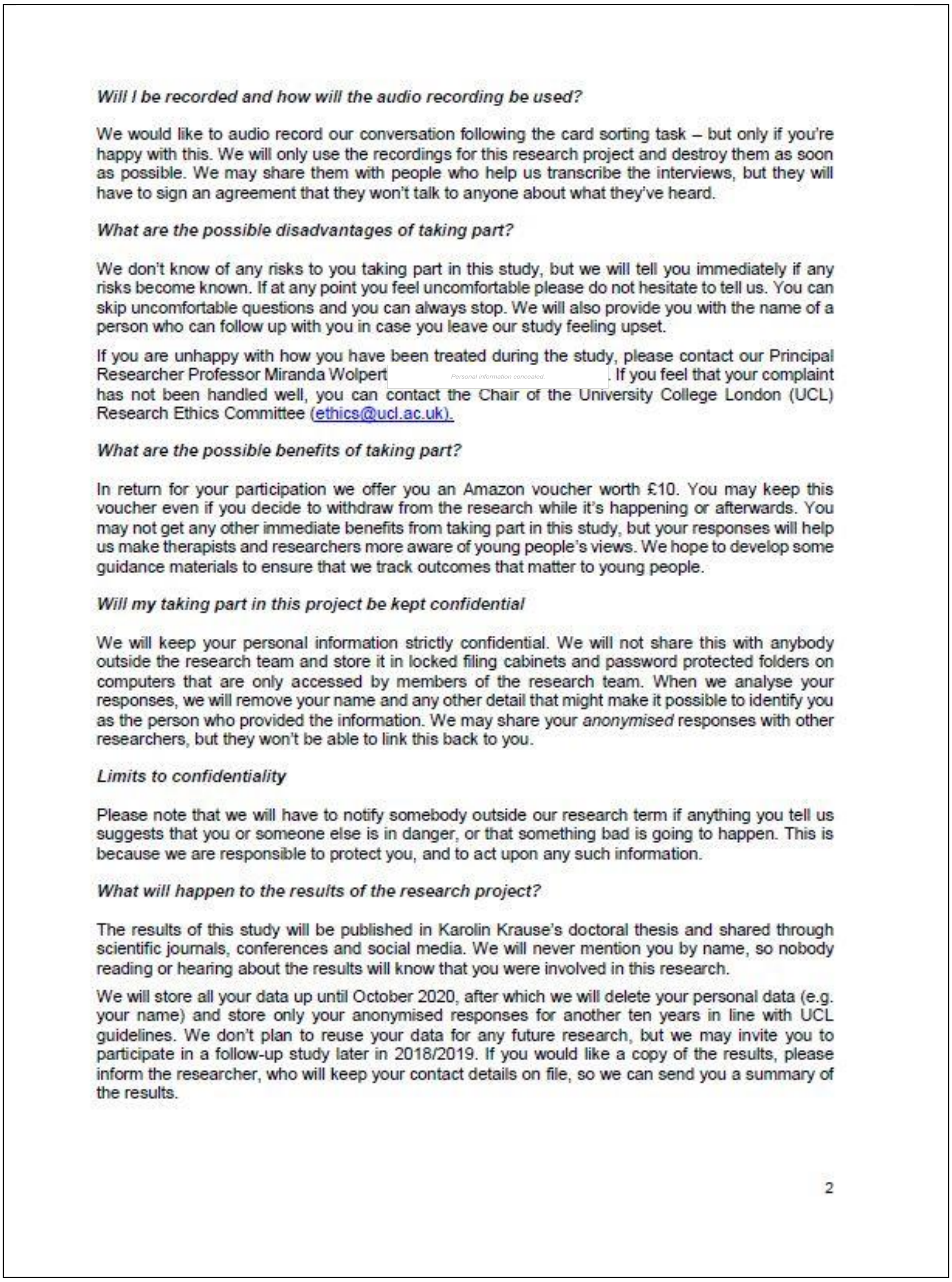
Do I have to take part?

No, it is completely up to you! If you decide now or later that you don't want to take part, you will not get into any trouble and this will not have any negative consequences. Even if you start to take part and then change your mind, you can stop at any time.

What will happen if I take part?

We will give you a set of cards that describe different outcomes of getting support for depression and ask you to sort these based on how important you think each outcome is. This will not be a test and there will be no right or wrong answers. You will then spend a few minutes discussing the exercise with a researcher and fill in a short questionnaire about yourself and the support you have received. All in all, this will take around 45 minutes. If you would like to participate, we will contact you to agree on a place and time to meet that's good for you.

University College London, Gower Street, London WC1E 6BT
Tel: +44 (0)20 7679 2000
email@ucl.ac.uk
www.ucl.ac.uk



Will I be recorded and how will the audio recording be used?

We would like to audio record our conversation following the card sorting task – but only if you're happy with this. We will only use the recordings for this research project and destroy them as soon as possible. We may share them with people who help us transcribe the interviews, but they will have to sign an agreement that they won't talk to anyone about what they've heard.

What are the possible disadvantages of taking part?

We don't know of any risks to you taking part in this study, but we will tell you immediately if any risks become known. If at any point you feel uncomfortable please do not hesitate to tell us. You can skip uncomfortable questions and you can always stop. We will also provide you with the name of a person who can follow up with you in case you leave our study feeling upset.

If you are unhappy with how you have been treated during the study, please contact our Principal Researcher Professor Miranda Wolpert Personal information concealed. . If you feel that your complaint has not been handled well, you can contact the Chair of the University College London (UCL) Research Ethics Committee (ethics@ucl.ac.uk).

What are the possible benefits of taking part?

In return for your participation we offer you an Amazon voucher worth £10. You may keep this voucher even if you decide to withdraw from the research while it's happening or afterwards. You may not get any other immediate benefits from taking part in this study, but your responses will help us make therapists and researchers more aware of young people's views. We hope to develop some guidance materials to ensure that we track outcomes that matter to young people.

Will my taking part in this project be kept confidential

We will keep your personal information strictly confidential. We will not share this with anybody outside the research team and store it in locked filing cabinets and password protected folders on computers that are only accessed by members of the research team. When we analyse your responses, we will remove your name and any other detail that might make it possible to identify you as the person who provided the information. We may share your *anonymised* responses with other researchers, but they won't be able to link this back to you.

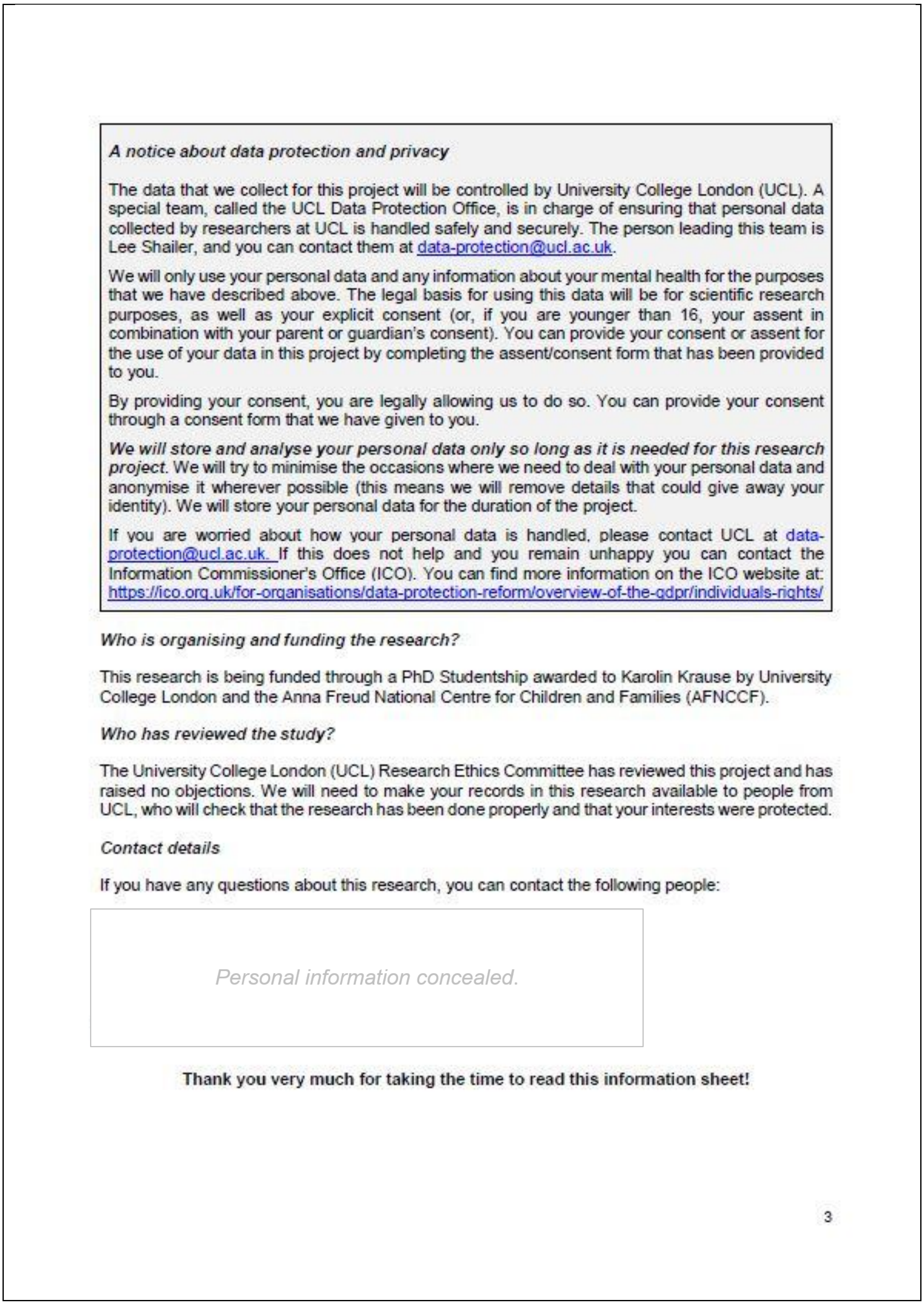
Limits to confidentiality

Please note that we will have to notify somebody outside our research term if anything you tell us suggests that you or someone else is in danger, or that something bad is going to happen. This is because we are responsible to protect you, and to act upon any such information.

What will happen to the results of the research project?

The results of this study will be published in Karolin Krause's doctoral thesis and shared through scientific journals, conferences and social media. We will never mention you by name, so nobody reading or hearing about the results will know that you were involved in this research.

We will store all your data up until October 2020, after which we will delete your personal data (e.g. your name) and store only your anonymised responses for another ten years in line with UCL guidelines. We don't plan to reuse your data for any future research, but we may invite you to participate in a follow-up study later in 2018/2019. If you would like a copy of the results, please inform the researcher, who will keep your contact details on file, so we can send you a summary of the results.



A notice about data protection and privacy

The data that we collect for this project will be controlled by University College London (UCL). A special team, called the UCL Data Protection Office, is in charge of ensuring that personal data collected by researchers at UCL is handled safely and securely. The person leading this team is Lee Shailer, and you can contact them at data-protection@ucl.ac.uk.

We will only use your personal data and any information about your mental health for the purposes that we have described above. The legal basis for using this data will be for scientific research purposes, as well as your explicit consent (or, if you are younger than 16, your assent in combination with your parent or guardian's consent). You can provide your consent or assent for the use of your data in this project by completing the assent/consent form that has been provided to you.

By providing your consent, you are legally allowing us to do so. You can provide your consent through a consent form that we have given to you.

We will store and analyse your personal data only so long as it is needed for this research project. We will try to minimise the occasions where we need to deal with your personal data and anonymise it wherever possible (this means we will remove details that could give away your identity). We will store your personal data for the duration of the project.

If you are worried about how your personal data is handled, please contact UCL at data-protection@ucl.ac.uk. If this does not help and you remain unhappy you can contact the Information Commissioner's Office (ICO). You can find more information on the ICO website at: https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/individuals-rights/

Who is organising and funding the research?

This research is being funded through a PhD Studentship awarded to Karolin Krause by University College London and the Anna Freud National Centre for Children and Families (AFNCCF).

Who has reviewed the study?

The University College London (UCL) Research Ethics Committee has reviewed this project and has raised no objections. We will need to make your records in this research available to people from UCL, who will check that the research has been done properly and that your interests were protected.

Contact details

If you have any questions about this research, you can contact the following people:

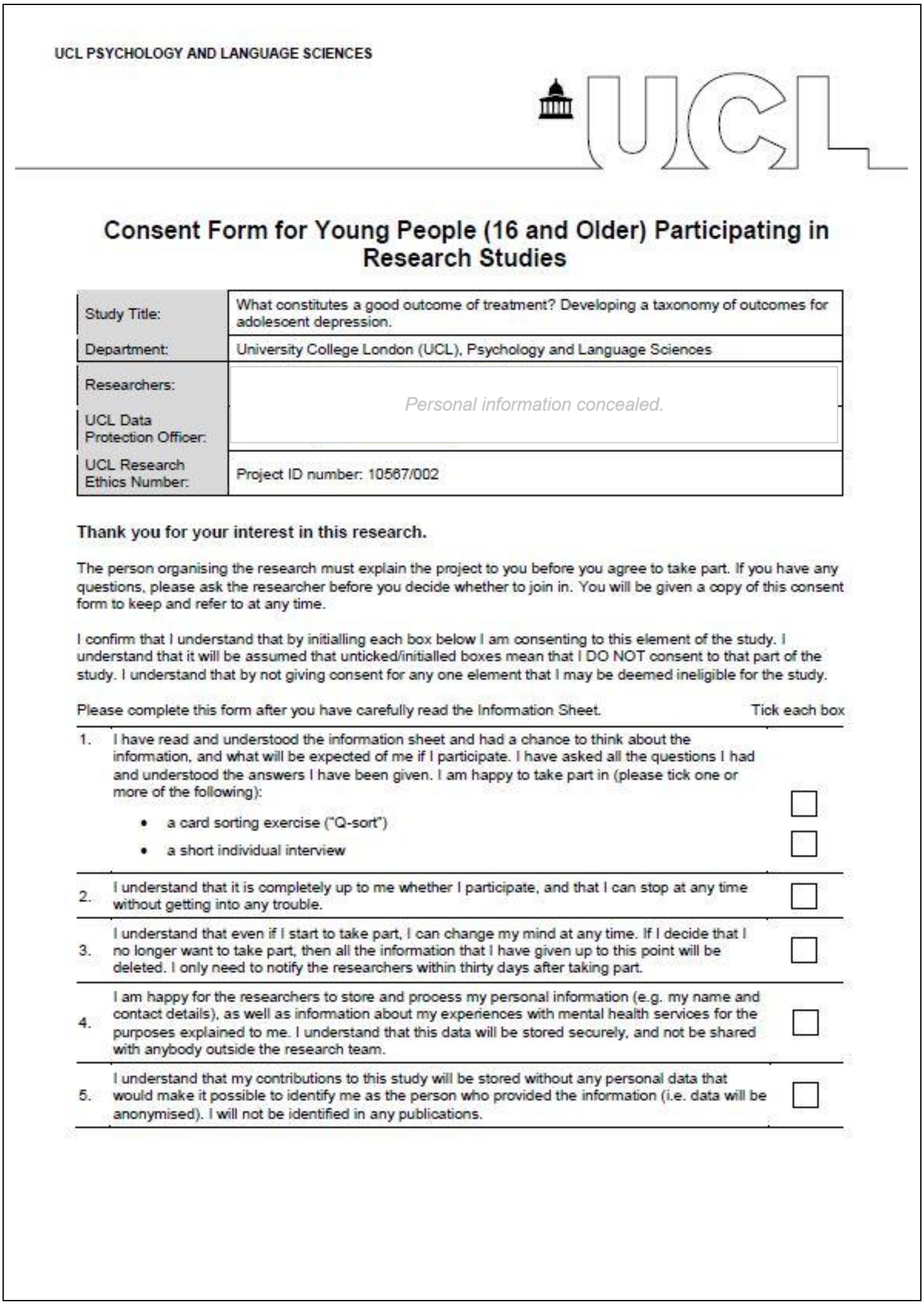
Personal information concealed.

Thank you very much for taking the time to read this information sheet!

3

Note. An adapted version of this participant information sheet was handed out to clinicians. An additional version was available for parents of youth under the age of 16.

Figure D.3. Consent Form Used with Young People



UCL PSYCHOLOGY AND LANGUAGE SCIENCES

UCL

Consent Form for Young People (16 and Older) Participating in Research Studies

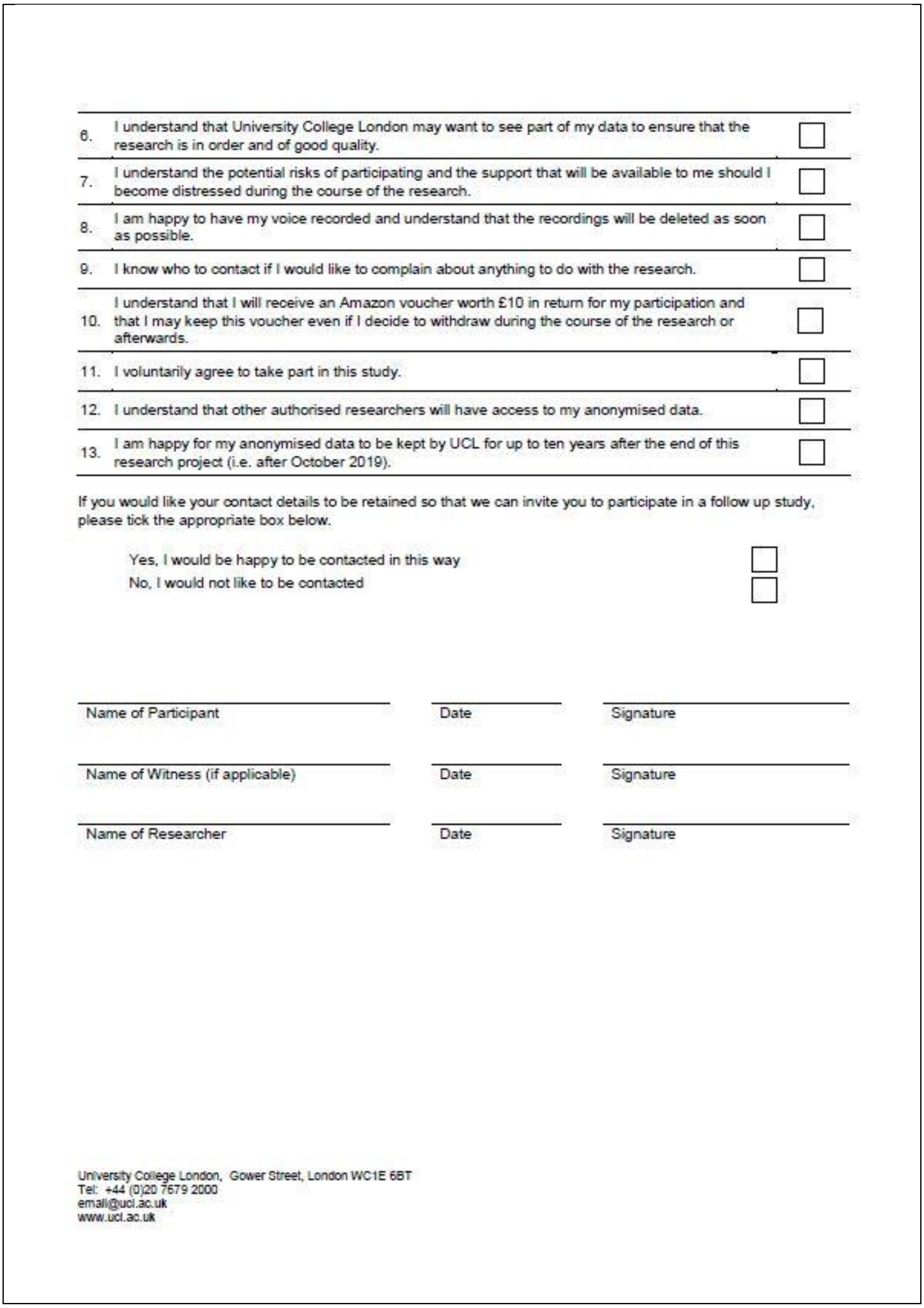
Study Title:	What constitutes a good outcome of treatment? Developing a taxonomy of outcomes for adolescent depression.
Department:	University College London (UCL), Psychology and Language Sciences
Researchers:	Personal information concealed.
UCL Data Protection Officer:	
UCL Research Ethics Number:	Project ID number: 10567/002

Thank you for your interest in this research.

The person organising the research must explain the project to you before you agree to take part. If you have any questions, please ask the researcher before you decide whether to join in. You will be given a copy of this consent form to keep and refer to at any time.

I confirm that I understand that by initialling each box below I am consenting to this element of the study. I understand that it will be assumed that unticked/initialled boxes mean that I DO NOT consent to that part of the study. I understand that by not giving consent for any one element that I may be deemed ineligible for the study.

Please complete this form after you have carefully read the Information Sheet.		Tick each box
1.	I have read and understood the information sheet and had a chance to think about the information, and what will be expected of me if I participate. I have asked all the questions I had and understood the answers I have been given. I am happy to take part in (please tick one or more of the following): • a card sorting exercise ("Q-sort") • a short individual interview	☐ ☐
2.	I understand that it is completely up to me whether I participate, and that I can stop at any time without getting into any trouble.	☐
3.	I understand that even if I start to take part, I can change my mind at any time. If I decide that I no longer want to take part, then all the information that I have given up to this point will be deleted. I only need to notify the researchers within thirty days after taking part.	☐
4.	I am happy for the researchers to store and process my personal information (e.g. my name and contact details), as well as information about my experiences with mental health services for the purposes explained to me. I understand that this data will be stored securely, and not be shared with anybody outside the research team.	☐
5.	I understand that my contributions to this study will be stored without any personal data that would make it possible to identify me as the person who provided the information (i.e. data will be anonymised). I will not be identified in any publications.	☐



6. I understand that University College London may want to see part of my data to ensure that the research is in order and of good quality. ☐

7. I understand the potential risks of participating and the support that will be available to me should I become distressed during the course of the research. ☐

8. I am happy to have my voice recorded and understand that the recordings will be deleted as soon as possible. ☐

9. I know who to contact if I would like to complain about anything to do with the research. ☐

10. I understand that I will receive an Amazon voucher worth £10 in return for my participation and that I may keep this voucher even if I decide to withdraw during the course of the research or afterwards. ☐

11. I voluntarily agree to take part in this study. ☐

12. I understand that other authorised researchers will have access to my anonymised data. ☐

13. I am happy for my anonymised data to be kept by UCL for up to ten years after the end of this research project (i.e. after October 2019). ☐

If you would like your contact details to be retained so that we can invite you to participate in a follow up study, please tick the appropriate box below.

Yes, I would be happy to be contacted in this way ☐
No, I would not like to be contacted ☐

Name of Participant　Date　Signature

Name of Witness (if applicable)　Date　Signature

Name of Researcher　Date　Signature

University College London, Gower Street, London WC1E 6BT
Tel: +44 (0)20 7679 2000
email@ucl.ac.uk
www.ucl.ac.uk

Note. An adapted version of this consent form was handed out to clinicians. Additional versions were available for youth under the age of 16 (i.e., assent) and their parents.

Figure D.4. Recruitment Flyers for the Viewpoints on Outcome Priorities Study

EBPU Evidence Based Practice Unit

A partnership of UCL Anna Freud National Centre for Children and Families

Research Participants needed!

Are you a young person aged 12-21 who has used mental health support to tackle depression? Join our Good Outcomes Study!

What is this study about?

This study is part of a PhD research project at University College London (UCL). We seek to understand what success of therapy looks like for young people with depression and their families.

What will my participation involve?

We will invite you to do a card sorting task and have a chat about what outcomes you think are important and why. Overall, this will take around 45 minutes.

Image: K. Arthur Endsley, 2017

Who can participate?

Young people aged 12-19 years who have used mental health services for depression, and their parents or carers.

Where will the study take place?

We will agree a time and place. For example, we could meet at your home, or at UCL in central London.

Will I get paid?

We will offer you a £10 Amazon voucher to thank you for your participation.

Are you interested? Please contact us!

Karolin Krause, PhD Student Personal information concealed.

EBPU Evidence Based Practice Unit

A partnership of UCL Anna Freud National Centre for Children and Families

Research Participants needed!

Are you a mental health professional supporting adolescents with depression? Join our Good Outcomes Study!

What is this study about?

This study is part of a PhD research project at University College London (UCL). We seek to understand what success of therapy looks like for young people with depression, their families, and clinicians

What will my participation involve?

We will invite you to do a card sorting task and have a chat about what outcomes you think are important and why. Overall, this will take around 30 minutes.

Image: K. Arthur Endsley, 2017

Who can participate?

Mental health professionals who routinely provide treatment to adolescents with depression.

Where will the study take place?

We will agree a time and place. For example, we could meet at your workplace, or at UCL in central London

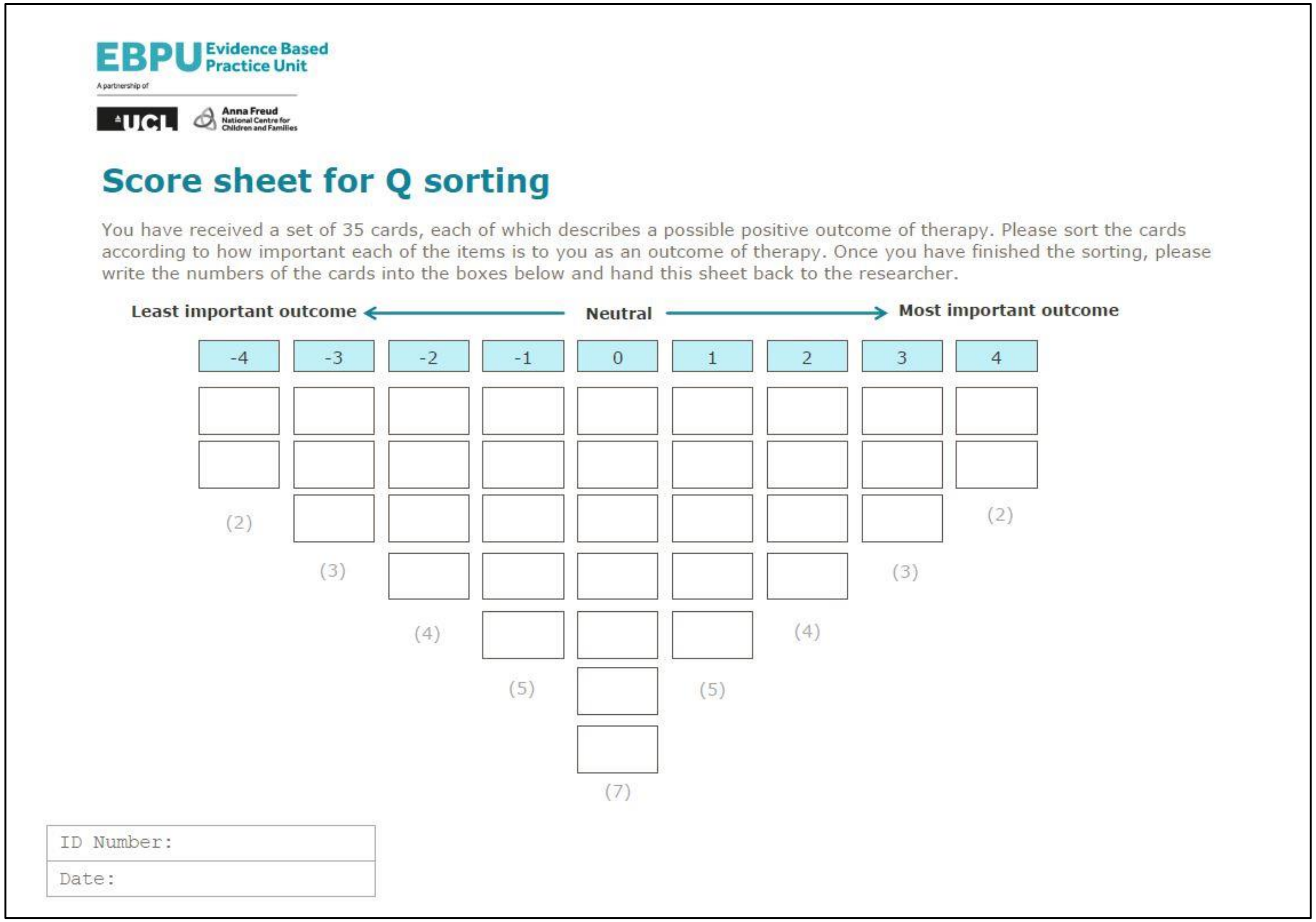
Will I get paid?

We cannot pay you any money in exchange for your participation, but we may be able to reimburse part of your travel expenses.

Are you interested? Please contact us!

Karolin Krause, PhD Student Personal information concealed.

Figure D.5. Q-Sorting Score Sheet for the Viewpoints on Outcome Priorities Study



EBPU Evidence Based Practice Unit
A partnership of
UCL
Anna Freud National Centre for Children and Families

Score sheet for Q sorting

You have received a set of 35 cards, each of which describes a possible positive outcome of therapy. Please sort the cards according to how important each of the items is to you as an outcome of therapy. Once you have finished the sorting, please write the numbers of the cards into the boxes below and hand this sheet back to the researcher.

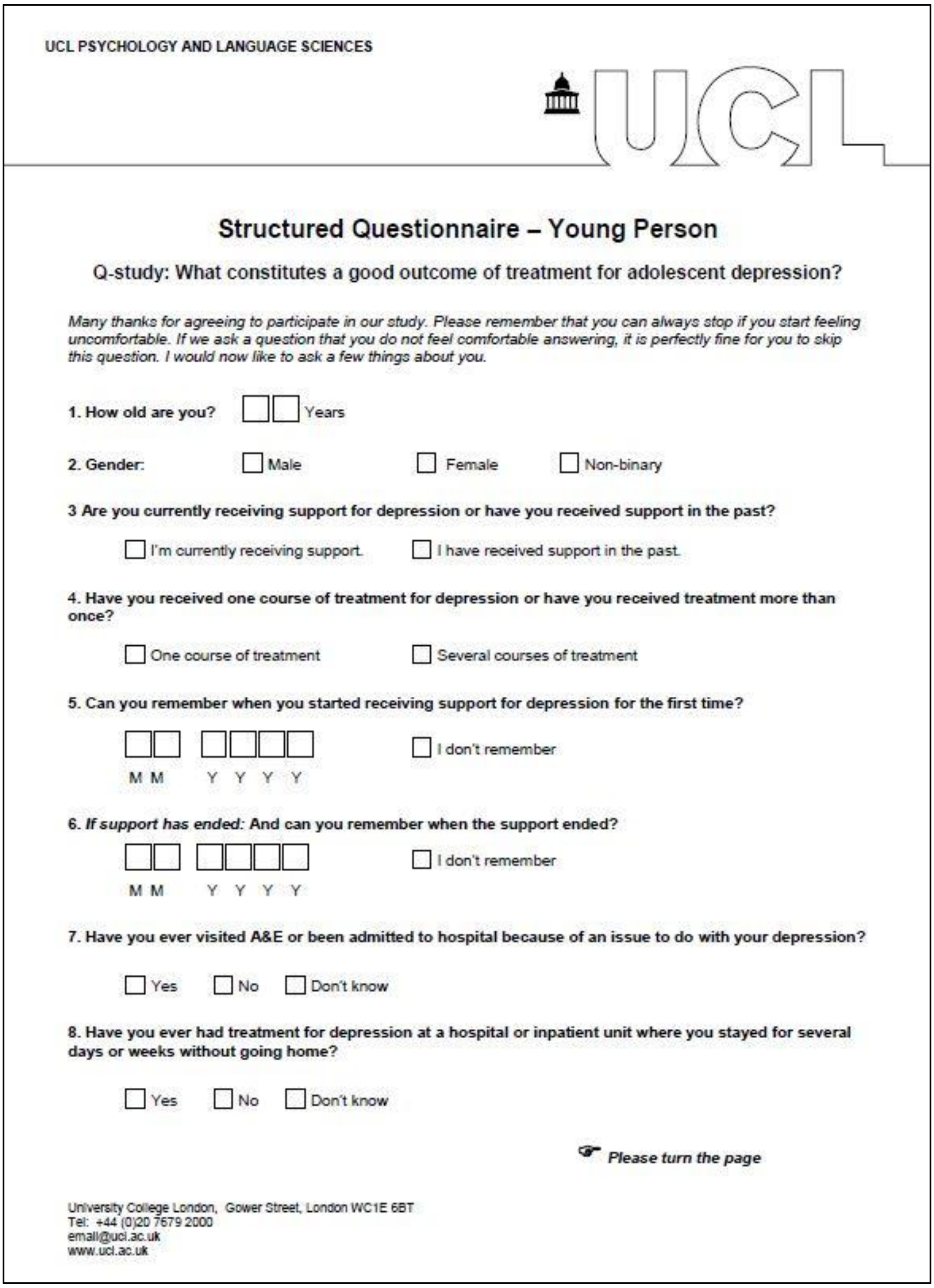
Least important outcome ← **Neutral** → **Most important outcome**

-4	-3	-2	-1	0	1	2	3	4
(2)	(3)	(4)	(5)	(7)	(5)	(4)	(3)	(2)

ID Number:
Date:

Note. This A4-sized score sheet mirrored the A2-sized Q-sorting template and was used by participants to record the position of each item card by noting the item number in the relevant box.

Figure D.6. Structured Questionnaire for Young People



UCL PSYCHOLOGY AND LANGUAGE SCIENCES

Structured Questionnaire – Young Person

Q-study: What constitutes a good outcome of treatment for adolescent depression?

Many thanks for agreeing to participate in our study. Please remember that you can always stop if you start feeling uncomfortable. If we ask a question that you do not feel comfortable answering, it is perfectly fine for you to skip this question. I would now like to ask a few things about you.

1. How old are you? ☐☐ Years

2. Gender: ☐ Male ☐ Female ☐ Non-binary

3 Are you currently receiving support for depression or have you received support in the past?

☐ I'm currently receiving support. ☐ I have received support in the past.

4. Have you received one course of treatment for depression or have you received treatment more than once?

☐ One course of treatment ☐ Several courses of treatment

5. Can you remember when you started receiving support for depression for the first time?

☐☐ ☐☐☐☐ ☐ I don't remember
M M Y Y Y Y

6. *If support has ended:* And can you remember when the support ended?

☐☐ ☐☐☐☐ ☐ I don't remember
M M Y Y Y Y

7. Have you ever visited A&E or been admitted to hospital because of an issue to do with your depression?

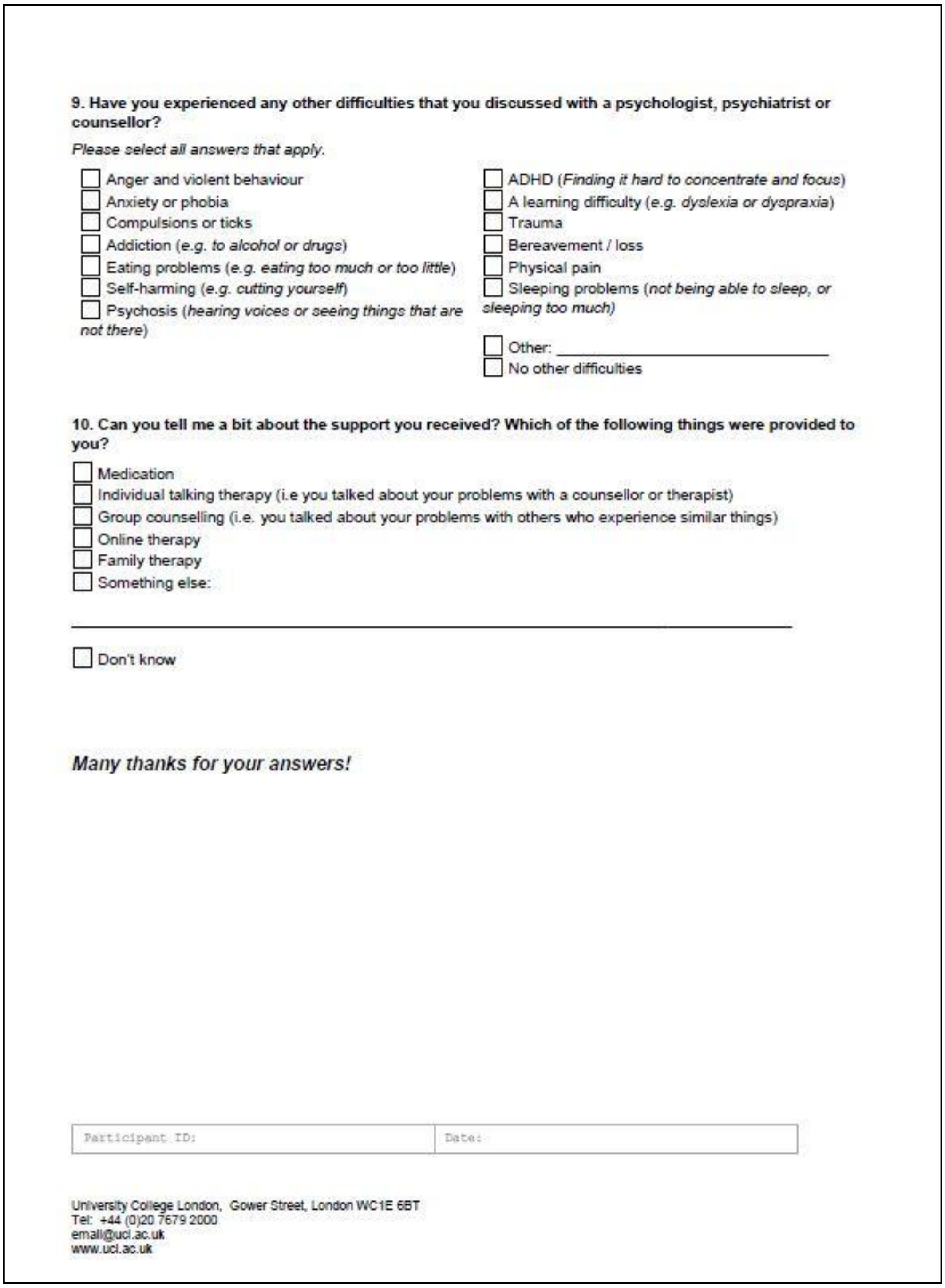
☐ Yes ☐ No ☐ Don't know

8. Have you ever had treatment for depression at a hospital or inpatient unit where you stayed for several days or weeks without going home?

☐ Yes ☐ No ☐ Don't know

☞ ***Please turn the page***

University College London, Gower Street, London WC1E 6BT
Tel: +44 (0)20 7679 2000
email@ucl.ac.uk
www.ucl.ac.uk



9. Have you experienced any other difficulties that you discussed with a psychologist, psychiatrist or counsellor?

Please select all answers that apply.

☐ Anger and violent behaviour
☐ Anxiety or phobia
☐ Compulsions or ticks
☐ Addiction (*e.g. to alcohol or drugs*)
☐ Eating problems (*e.g. eating too much or too little*)
☐ Self-harming (*e.g. cutting yourself*)
☐ Psychosis (*hearing voices or seeing things that are not there*)

☐ ADHD (*Finding it hard to concentrate and focus*)
☐ A learning difficulty (*e.g. dyslexia or dyspraxia*)
☐ Trauma
☐ Bereavement / loss
☐ Physical pain
☐ Sleeping problems (*not being able to sleep, or sleeping too much*)

☐ Other: ______________________________
☐ No other difficulties

10. Can you tell me a bit about the support you received? Which of the following things were provided to you?

☐ Medication
☐ Individual talking therapy (i.e you talked about your problems with a counsellor or therapist)
☐ Group counselling (i.e. you talked about your problems with others who experience similar things)
☐ Online therapy
☐ Family therapy
☐ Something else:

____

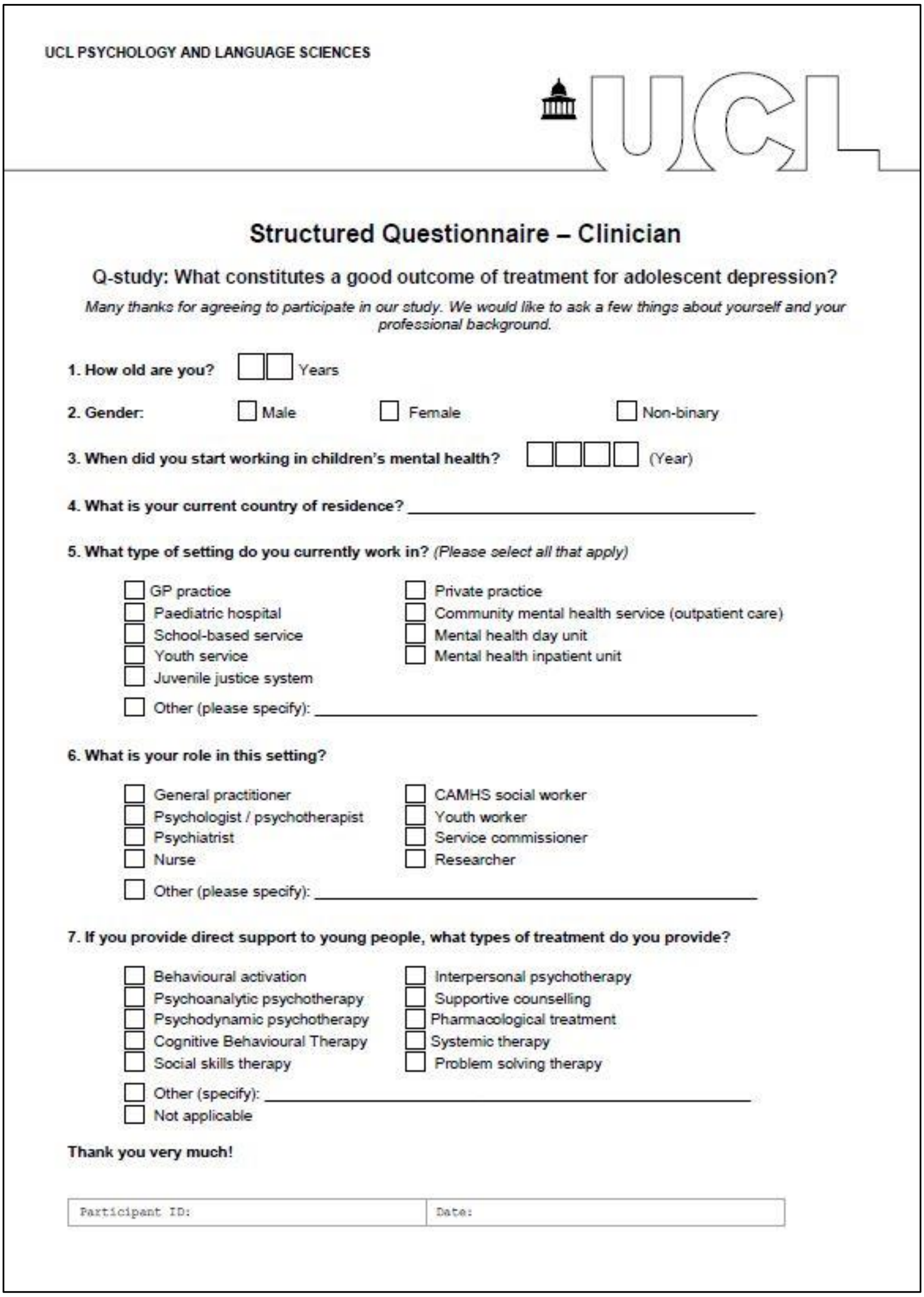
☐ Don't know

Many thanks for your answers!

Participant ID:	Date:

University College London, Gower Street, London WC1E 6BT
Tel: +44 (0)20 7679 2000
email@ucl.ac.uk
www.ucl.ac.uk

Figure D.7. Structured Questionnaire for Clinicians



UCL PSYCHOLOGY AND LANGUAGE SCIENCES

UCL

Structured Questionnaire – Clinician

Q-study: What constitutes a good outcome of treatment for adolescent depression?

Many thanks for agreeing to participate in our study. We would like to ask a few things about yourself and your professional background.

1. How old are you? ☐☐ Years

2. Gender: ☐ Male ☐ Female ☐ Non-binary

3. When did you start working in children's mental health? ☐☐☐☐ (Year)

4. What is your current country of residence? ______________________________

5. What type of setting do you currently work in? *(Please select all that apply)*

☐ GP practice
☐ Paediatric hospital
☐ School-based service
☐ Youth service
☐ Juvenile justice system
☐ Private practice
☐ Community mental health service (outpatient care)
☐ Mental health day unit
☐ Mental health inpatient unit
☐ Other (please specify): ______________________________

6. What is your role in this setting?

☐ General practitioner
☐ Psychologist / psychotherapist
☐ Psychiatrist
☐ Nurse
☐ CAMHS social worker
☐ Youth worker
☐ Service commissioner
☐ Researcher
☐ Other (please specify): ______________________________

7. If you provide direct support to young people, what types of treatment do you provide?

☐ Behavioural activation
☐ Psychoanalytic psychotherapy
☐ Psychodynamic psychotherapy
☐ Cognitive Behavioural Therapy
☐ Social skills therapy
☐ Interpersonal psychotherapy
☐ Supportive counselling
☐ Pharmacological treatment
☐ Systemic therapy
☐ Problem solving therapy
☐ Other (specify): ______________________________
☐ Not applicable

Thank you very much!

Participant ID:	Date:

Table D.1. Characteristics of Young People Associated With Each Factor

Participant Characteristics	Percentage of participants			
	F1	F2	F3	F4
Mean age (in years)	19.25	18.17	18.75	18.33
	Percentage of participants in each factor displaying the characteristic			
% Female	37.5	50.0	87.5	66.7
Currently receiving treatment	50.0	16.7	50.0	66.7
Received treatment in the past	50.0	83.3	50.0	33.3
Single course of treatment	37.5	50.0	50.0	33.3
Repeated courses of treatment	62.5	50.0	50.0	66.7
Visited A&E in relation with depression	—	16.7	37.5	66.7
Admitted to inpatient-care in relation with depression	—	—	12.5	66.7
Average number of comorbid problems	2.86	4	5.38	7.33
Anger	—	33.3	25.0	66.7
Anxiety	62.5	100.0	87.5	100.0
Compulsions	12.5	16.7	37.5	—
Substance use	12.5	16.7	25.0	33.3
Eating	50.0	66.7	50.0	100.0
Self-harm	37.5	66.7	75.0	100.0
Psychosis	12.5	16.7	12.5	66.7
ADHD	12.5	—	25.0	—
Learning difficulties	12.5	—	50.0	33.3
Trauma	—	—	25.0	66.7
Bereavement	25.0	—	12.5	33.3
Sleep	50.0	66.7	62.5	100.0
Somatic	—	—	12.5	33.3
Autism Spectrum Disorder	—	16.7	37.5	—

Table D.2. Hand Rotation Angles – Youth Sample

Factor #1	Factor #2	Angle
1	3.	.8
1	4	-.6
2	1	-.9
2	4	-.1
3	2	.6
3	4	.2

Note. Rotated in PQROT [28/01/2019]

Table D.3. Correlations Between Youth Viewpoints

Viewpoint name	Factor			
	Y.A.	Y.B.	Y.C.	Y.D.
Y.A. Becoming a 'healthier' person	1.00			
Y.B. Learning to help myself	0.33	1.00		
Y.C. Making sense of the past to embrace the future	0.26	0.30	1.00	
Y.D. Being able to do what other adolescents do	0.07	0.14	0.05	1.00

Note. Y.A. = youth viewpoint A; Y.B. = youth viewpoint B, etc.

Table D.4. Hand Rotation Angles – Professional Sample

Factor #1	Factor #2	Angle
1	2	-.5
1	3	+.1

Note. Rotated in PQROT [29/04/2019]

Table D.5. Correlations Between Professional Viewpoints

Professional viewpoints	Professional viewpoints		
	P.A/B	P.C	P.D
P.A/B. Managing risk and reducing symptoms / Empowering youth and parents	1.00		
P.C. Building skills in young people to aid coping	0.33	1.00	
P.D. Building family support around young people.	0.26	0.30	1.00

Note. P.A. = professional viewpoint A; P.B. = professional viewpoint B, etc.

Table D.6. Hand Rotation Angles – Professional Sample

Factor #1	Factor #2	Angle
1	2	-.5
2	4	+.14
3	4	-.8

Note. Rotated in PQROT [29/04/2019]

Table D.7. Correlations Between Super Viewpoints

Super Viewpoints	Super Viewpoints			
	S.A.	S.B.	S.C.	S.D.
S.A. Symptoms – Feeling better	1.00			
S.B. Self-management – Resilience through coping skills	0.37	1.00		
S.C. Parental support – Resilience through family support	0.20	0.24	1.00	
S.D. Functioning – Less interference with daily life	-0.01	0.02	-0.27	1.00

Note. S.A. = Super viewpoint A; S.B. = super viewpoint B, etc.

Appendix E. Supplementary Materials – Chapter 7

Table E.1. Composition of Reliable Change Metric for Symptoms and Functioning

Domain	Across domain comparison
Symptoms	369 (100%)
Data from SDQ Emotion only	110 (29.8%)
Data from RCADS only	17 (4.6%)
Data from both measures	242 (65.6%)
Functioning	369 (100%)
Data from SDQ Impact only	309 (83.7%)
Data from C/ORS only	18 (4.9%)
Data from both measures	42 (11.4%)

Table E.2. Levels of Reliable Change

Measure/domain	*N*	Reliable change		
		Deteriorated	Unchanged	Improved
Within symptom domain				
SDQ Emotion	943	1.7%	76.4%	22.0%
RCADS	943	7.7%	43.6%	48.7
Within functioning domain				
SDQ Impact	120	4.2%	76.7%	19.2%
C/ORS	120	5.0%	54.2%	40.8%
Across symptoms, functioning, goal progress				
Symptoms	369	6.8%	51.2%	42.0%
Functioning	369	3.3%	73.4%	23.3%
Goal progress	369	5.2%	28.5%	66.4%

N = 1,099.

Table E.3. Patterns of Data Availability

	Paired data available on each measure					Considered for comparative analysis		
n	SDQ Emotion	RCADS	SDQ Impact	C/ORS	GBO	Within symptom domain	Within functioning domain	Across symptoms, functioning, and goal progress
31	✓	✓	✓	✓	✓	✓	✓	✓
614	✓	✓	—	—	—	✓	—	—
208	✓	✓	✓	—	✓	✓	—	✓
98	✓	—	✓	—	✓	—	—	✓
49	✓	✓	✓	✓	—	✓	✓	—
38	✓	✓	—	—	✓	✓	—	—
28	—	—	✓	✓	—	—	✓	—
14	—	✓	—	✓	✓	—	—	✓
11	✓	—	✓	✓	✓	—	—	✓
3	—	✓	✓	—	✓	—	—	✓
3	✓	✓	—	✓	✓	✓	—	✓
1	—	—	✓	✓	✓	—	✓	—
1	✓	—	—	✓	✓	—	—	✓

Note. *N* = 1,099. ✓ means paired data is available; — means paired data is not available.

Table E.4. Sample with Complete Data on the C/ORS at T1 and T2 by Age

Measure	Aged 12 years		Aged 13 and older	
	N	*Mean* (*SD*)	*N*	*Mean* (SD)
CORS T1	77	19.12 (8.87)	745	18.10 (8.32)
CORS T2	15	21.72 (12.24)	129	19.59 (9.78)
ORS T1	20	19.79 (8.80)	467	19.20 (8.78)
ORS T2	13	29.82 (7.97)	315	24.33 (10.3)

Note. *N* = 1,781 within the sample of 9,904 who were moderately to severely depressed on the Current View (i.e., the eligible sample).

Table E.5. Missing Value Patterns for the C/ORS (Youth Aged 12 Years)

n	CORS T1	ORS T1	CORS T2	ORS T2
0	1	1	1	1
551	0	0	0	0
56	1	0	0	0
15	1	0	1	0
8	0	1	0	1
6	0	1	0	0
4	1	1	0	1
2	1	1	0	0
1	0	0	0	1

Note. 1 means complete; 0 means missing. *N* = 643 within the sample of 9,904 who were moderately to severely depressed on the Current View (i.e., the eligible sample).

Table E.6. Missing Value Patterns for the C/ORS (Youth Aged > 12 Years)

n	CORS T1	ORS T1	CORS T2	ORS T2
1	1	1	1	1
8,104	0	0	0	0
555	1	0	0	0
262	0	1	1	0
138	0	1	0	0
123	1	0	1	0
41	1	1	1	0
21	1	1	0	0
11	0	0	1	0
4	1	1	0	1
0	0	0	0	1

Note. 1 means complete; 0 means missing. 1 means complete; 0 means missing. *N* = 9260 within the sample of 9,904 who were moderately to severely depressed on the Current View (i.e., the eligible sample).

Table E.7. Missing Value Patterns: SDQ Emotion Assessment and Follow-up Forms

n	SDQ Ass T1	SDQ Ass T2	SDQ Fup T1	SDQ Fup T2
20	1	1	1	1
5218	1	0	0	0
2857	0	0	0	0
891	1	0	1	0
416	1	1	0	0
224	0	0	1	0
120	1	0	1	1
102	1	1	1	0
44	0	0	1	1
5	0	1	0	0
3	0	0	0	1
3	0	1	1	0
1	0	1	1	1

Note. 1 means complete; 0 means missing. *N* = 9,904. SDQ ASS stands for SDQ initial assessment template. SDQ FUP stands for SDQ follow-up template.

Table E.8. Missing Value Patterns: SDQ Impact Assessment and Follow-up Forms

n	SDQ Ass T1	SDQ Ass T2	SDQ Fup T1	SDQ Fup T2
14	1	1	1	1
4996	1	0	0	0
3260	0	0	0	0
746	1	0	1	0
371	1	1	0	0
251	0	0	1	0
107	1	0	1	1
84	1	1	1	0
40	0	0	1	1
23	0	1	0	0
4	0	0	0	1
3	0	1	1	0
3	1	0	0	1
1	0	1	1	1
1	1	1	0	1

Note. 1 means complete; 0 means missing. *N* = 9,904. SDQ ASS stands for SDQ initial assessment template. SDQ FUP stands for SDQ follow-up template.

Table E.9. Data Used to Compute the SDQ Emotion Variable

n	SDQ Ass T1	SDQ Ass T2	SDQ Fup T1	SDQ Fup T2
T1	1479	0	57	
T2	9	442	927	176

Note. *N* = 3,090. *Note.* *N* = 2,626. SDQ ASS stands for SDQ initial assessment template. SDQ FUP stands for SDQ follow-up template.

Table E.10. Data Used to Compute SDQ Impact Variable

n	SDQ Ass T1	SDQ Ass T2	SDQ Fup T1	SDQ Fup T2
T1	1268	0	45	
T2	6	385	770	152

Note. *N* = 2,626. SDQ ASS stands for SDQ initial assessment template. SDQ FUP stands for SDQ follow-up template.

Table E.11. Missing Value Patterns for the GBO at T1

n	Goal 1	Goal 2	Goal 3
530	1	1	1
7931	0	0	0
854	1	0	0
508	1	1	0
38	0	1	0
22	0	0	1
12	0	1	1
9	1	0	1

Note. 1 means complete; 0 means missing. *N* = 9,904.

Table E.12. Missing Value Patterns for the GBO at T2

n	Goal 1	Goal 2	Goal 3
336	1	1	1
8,820	0	0	0
357	1	0	0
279	1	1	0
43	0	1	0
34	0	0	1
20	0	1	1
15	1	0	1

Note. 1 means complete; 0 means missing. *N* = 9,904.