School of Population Health

An Empirical and Computational Investigation of Neurocognitive Performance Underlying Dimensional Psychopathology

Darren Haywood

0000-0002-9317-4135

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

September 2022

Declaration

To the best of my knowledge, this thesis contains no material previously published by any other person except where acknowledgement has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Human Ethics: The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. The proposed research study received human research ethics approval from the Curtin University Human Research Ethics Committee: Approval numbers HRE2021-0105.

Signature: 

Date: 27/09/22

Acknowledgements

First, I would like to my wonderful supervisors, Professor Barbara Mullan, Dr. Frank Baughman, and Associate Professor Karen Heslop. Over the years you have provided not only insight and support academically, but also on a personal level. I could have not hoped for a better team, especially during a worldwide pandemic, and I am so very proud of the work in this thesis and my growth, both academically and personally that has come throughout this journey. The work in this thesis had led me across to country to the wonderful city of Melbourne to a job I could have only dreamt about when I started this PhD, and I am so incredibly grateful for your roles in making this a reality. A special thank you goes to Dr. Frank Baughman, who over the past 7 years has been the most wonderful mentor and role model. I really cherish all our insightful conversations, disagreements, and most of all, all of your hilarious stories and anecdotes. I would also like to thank my new colleagues at St Vincent's Hospital Melbourne, in particular Professor Susan Rossell, who has provided excellent advice and insightful discussion relating to this work towards its completion.

I'd like to thank those who participated in this research, my wonderful colleagues at Curtin University who provided support and advice including Professor Peter McEvoy, Professor Moira O'Connor, Professor Clare Rees, Dr. Welber Marinovic, and Professor Lauren Breen. A thanks also go to my fellow PhD candidates, in particular Ashleigh Pantaleo, for all the laughs and shenanigans over the years. I would also like to the Australian Government, for supporting this work, through a scholarship, for the final two years.

Now to my family: First to my mom, dad, and brothers. Thank you for supporting me over these years, including through all of my stress meltdowns, long-hours, and tears. I know it is not the easiest of tasks, but you have all done everything you could to support me through these years, and I thank you for that. I would also like to thank you for supporting Zoe and I in our move across the country, it is tough not being with you, but thank you for supporting my passion. Lastly, to my wonderful partner, Zoe Cahill, who has been by my side through each of these years of this PhD, I can't express how grateful I am. You have been unwavering in your support of my career and love of this work, regardless of financial difficulties, the long-hours, and uncertainty. You have moved across the country, with me, from Perth to Melbourne, leaving your friends and family, all so I can take a job that I have dreamt of. I am so proud of the lives we are building together. Chasing our passions, uncompromisingly, side-by-side.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ALC – Alcohol

ANN – Artificial Neural Network

ASSIST – Alcohol, Smoking and Substance Involvement Screening Test

BSI – Brief Symptom Inventory

CAN – Cannabis

CD – Conduct disorder

CFA – Confirmatory factor analysis

CFI – Comparative fit index

DRG – Hard drugs

EF – Executive functioning

EFA – Exploratory factor analysis

EXT – Externalising

GAD – General anxiety disorder

Infer. Rel. – Inferring relevance

INT – Internalising

IQ – Intelligence quotient

IT – Inspection Time.

MDE – Major depressive episode

MLR – Multiple Linear Regression

OCD – Obsessive compulsive disorder

PR – Perceptual reasoning

PS – processing speed

RCS – Rate-correct score

RMSEA – Root-mean-square error of approximation

RT – Reaction time

RVIP – Rapid visual information processing

SCHIZ – Schizophrenia

Shape-Num – Shape-Number

THT – Thought disorder.

TLI – Tucker–Lewis index

TMT – Trail making test

TOB – Tobacco

VC – Verbal comprehension

Vis WM – Visual working memory

WAIS – Weschler Adult Intelligence Scale

WISC – Weschler Child Intelligence Scale

WM – Working memory

List of Appendices

Summary/Abstract

Deficits in neurocognitive abilities have been claimed to be an aetiological feature of psychopathology. However, to date there no consistent neurocognitive deficit has been found within any psychopathological disorder, indeed the associations between neurocognitive performance and psychopathology have been found to be extensively heterogeneous. One reason for this demonstrated heterogeneity may be due to the dominant use of the traditional nosological approach to diagnosis, using tools such as the Diagnostic and Statistical Manual of Mental Disorders (DSM). The traditional nosological approach has been met with multiple challenges, including high levels of comorbidity and poor diagnostic stability, meaning that the study of any single disorder is difficult. Recently, dimensional structural models of psychopathology have been developed through factor analysis that, instead of categorising psychopathology, view psychopathological experience on a dimension across multiple higher and lower order factors and indicators. Arising from these models are factors such as internalising, externalising and the *p*-factor. The *p*-factor in particular has drawn much interest and was said to be a substantive construct representing individuals' overall propensity toward psychopathology. The present thesis explores the associations between neurocognition and dimensional conceptualisations of psychopathology, including the *p*-factor, and explores our claim that non-linear multidimensional interactions between neurocognitive components underly the functional association between neurocognition and psychopathology.

To provide an initial understanding of the potential for the higher-order factors of psychopathology to have a universal substantive meaning, we explored the utility and consistency of four popular models of psychopathology in a range of population subgroups . We then examined the consistency of the neurocognitive correlates of the models' factors to provide an indication of substantive consistency. Only eight out of the sixty-three population subgroups fit any of the four popular structural models of psychopathology tested; the correlated factors model, the bifactor model, the revised bifactors model or the single factors model.. The strength of the neurocognitive correlates of the factors derived from the subgroups differed substantially from the correlations to the factors derived from the total sample. Overall, this suggests that the utility of structural models of psychopathology is best on the population level and that the substantive meaning of the factors differs between different samples. Developing a universal

substantive meaning of the higher-order factors of psychopathology may therefore be challenging and, if it was to be developed, it would only be applicable at the population level.

Informed by study one, the second study examined the potential for an alternative structural model of psychopathology, the S-1 bifactor model, to elucidate the associations between neurocognition and dimensional psychopathology. We developed S-1 bifactor models where the general factor (i.e., the *p*-factor) is defined by overall neurocognitive performance. This , therefore, mitigates the issue of a general factor without substantive meaning. We show in this study that the S-1 bifactor approach provides distinct advantages over the typical bifactor approach by allowing for the assessment of individual factor loadings on the general (cognitive) factor, and the associations between the psychopathology factors after accounting for the cognitive factor, can help form hypotheses. However, the S-1 bifactor approach is limited to providing only an understanding of the associations between psychopathology and overall neurocognition (or a single neurocognitive component) across the population. Therefore, to provide an understanding of the variability of neurocognition and its associations to psychopathology, supplementary approaches are required.

The third study looked to develop statistical models of psychopathology and explore if, the factors of psychopathology had different neurocognitive correlates. These results would elucidate potential aetiological explanations of the separate factors. We collected data participants on dimensional measures of psychopathology and substance use, as well as eight neurocognitive tasks measuring working memory, shifting, inhibition and speed of processing. Both the correlated factors model and the single factor model provided a good fit for the data. Only the tasks that measured speed of processing were significantly associated with internalising, externalising, and the *p*-factor. Therefore, on the population level, these results provided evidence that neurocognitive performance does not differentiate between the factors of psychopathology and instead speed of processing is a common correlate across the domains.

The final study directly compared two conceptualisations of the association and functionality between neurocognition and psychopathology: our non-linear multidimensional interactive conceptualisation versus the traditional linear conceptualisation. Using the neurocognitive, psychopathology, and substance use data previously collected, we compared the predictive accuracy of artificial neural network models to traditional linear models with regards to lower-level and higher-level psychopathology. Both the artificial neural network models and

the linear models had the same predictors; age, gender, and the eight neurocognitive tasks. The artificial neural network models were significantly more accurate than the linear models at predicting both (a) lower-level (i.e., depression, hostility, cannabis use) and (b) higher-level psychopathology (internalising, externalising, and the *p*-factor). These results provide support for the non-linear multidimensional interactive conceptualisation being superior to traditional linear conceptualisations of neurocognition within psychopathology. We do note however that, as shown in study one, a universal substantive meaning of the higher-level factors is not apparent, and therefore, even although a dynamic multidimensional approach is useful in predicting these factors, nuanced examinations of the multidimensional functionality of neurocognition in psychopathology may be best at the lower-level (i.e., non-factorised) domains.

Overall, this thesis provides important knowledge to forward our understanding of the structure, function, and conceptualisation of both psychopathology and neurocognition. This thesis provides an understanding of the considerations of the substantive interpretation of the factors of psychopathology, an illustration of an alternative structural approach to examine neurocognition within psychopathology, a greater understanding of the neurocognitive correlates of higher-order dimensional psychopathology across the population, and support for a multidimensional interactive conceptualisation of neurocognition within psychopathology. Future research should further examine the substantive interpretability of the factors of psychopathology, utilise the S-1 bifactor approach to further our understanding of trends of neurocognitive dysfunction within psychopathology across the population, and further examine the non-linear multidimensional integrative conceptualisation through computational modelling, descriptive, and artificial neural network approaches. Our findings, along with future research extending upon this work, may be used to inform dynamic multidimensional tools and approaches to prediction, assessment, treatment, and rehabilitation pertaining to psychopathology.

Author's Note

This thesis is presented as a hybrid thesis. The hybrid thesis includes, in the form of thesis chapters, manuscripts that have been published in peer-reviewed journals. While the thesis' chapters build upon each other, each chapter, apart from the general discussion, was produced to be a stand-alone publication. Therefore, some repetition, particularly in the background sections of each chapter is unavoidable. Each chapter is preceded by a preface that describes the contribution of the chapter to the overall aims of the thesis. The preface also acts to link each chapter, creating a unified body of research and discussion. The references for each manuscript are presented together at the end of the thesis to facilitate readability.

Publications Included in this Thesis

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2021). Psychopathology and Neurocognition in the Era of the p-factor: The Current Landscape and the Road Forward. *Psychiatry International*, *2*(3), 233-249. https://doi.org/10.3390/psychiatryint2030018

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2021). One p-factor for all? Exploring the applicability of structural models of psychopathology within subgroups of a population. *International Journal of Environmental Research and Public Health*, *18*(13), 7108. https://doi.org/10.3390/ijerph18137108

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2021). Going "up" to move forward: S-1 bifactor models and the study of neurocognitive abilities in psychopathology. *International Journal of Environmental Research and Public Health*, *18*(14), 7413. https://doi.org/10.3390/ijerph18147413

Haywood, D., Baughman, F.D., Mullan, B.A., Heslop, K.R. (2022). What Accounts for the Factors of Psychopathology? An Investigation of the Neurocognitive Correlates of Internalising, Externalising, and the *p*-Factor. *Brain Sciences. 12*(4), 421. https://doi.org/10.3390/brainsci12040421

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2022). Neurocognitive Artificial Neural Network Models Are Superior to Linear Models at Accounting for Dimensional Psychopathology. *Brain Sciences*, *12*(8), 1060. https://doi.org/10.3390/brainsci12081060

Preface to Chapter 1

The following chapter sets the scene of the current state of the literature regarding the association and dynamics of psychopathology and neurocognitive abilities, and provides a road map, in the form of four specific directions for future research. Chapter one blends evidence and theory from clinical psychology, neuropsychology, biology, and philosophy into an integrated narrative outlining the development of the literature in this area to date, including both historical perspectives and recent developments. After discussing the development of the literature, we provide several pertinent considerations and unknowns for the exploration of psychopathology and neurocognitive abilities and propose the application of a non-linear multidimensional interactive neurocognitive conceptualisation to the study of structural models of psychopathology. The proposed road forward that concludes this chapter will be followed in the remaining chapters of this thesis.

The following Introduction of chapter one (all sections until Thesis Aims) has been published in Psychiatry International.

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2021). Psychopathology and neurocognition in the era of the *p*-factor: The current landscape and the road forward. *Psychiatry International*, *2*(3), 233-249. https://doi.org/10.3390/psychiatryint2030018

Minor edits have been made to the following chapter, such as phrasing and Australian Spelling and the re-formatting of figures, to ensure consistency within the thesis.

This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

Chapter 1: Psychopathology and Neurocognition in the Era of the *p*-Factor: The Current Landscape and the Road Forward

1.1 Introduction

A consensus exists within the study of typical human development that variability in neurocognitive abilities accounts for a large proportion of individual differences in domains such as problem solving, reasoning, thinking and planning. Furthermore, deficits in neurocognitive processes have been repeatedly implicated in studies of psychopathology (Snyder et al., 2015). For example, deficits in the executive function (EF) processes of shifting, updating and inhibition have each been separately argued to explain symptoms of schizophrenia (e.g., Galletly et al., 2007; Gilleen et al., 2016; Kiehl et al., 2000), depression (e.g., De Lissnyder et al., 2010; Joormann & Gotlib, 2008; Joormann et al., 2007) and substance use disorder (e.g., Brooks et al., 2017; Mahmood et al., 2013; Noël et al., 2013). However, the literature shows little agreement as to which neurocognitive processes are of primary importance in any given disorder. One reason for this is that, within the context of clinical diagnoses, individuals diagnosed with the same psychopathological disorder can exhibit markedly different symptoms. Another reason is that many individuals diagnosed with a specific psychopathological disorder are also found to meet the criteria for other disorders (Newman et al., 1998); thus, making the *pure* study of any given disorder more challenging.

Issues to do with the heterogeneity of symptoms and the comorbidity of disorders have motivated the development of several structural models of psychopathology aimed at accounting for covariation amongst psychopathology and providing a dimensional framework that can be used for the description and understanding of psychopathology (Caspi et al., 2014; Kotov et al., 2017; Lahey et al., 2012). Whilst in some instances transdiagnostic approaches have been hailed as achieving a degree of success (e.g., Aldao et al., 2016; Mansell et al., 2012; McManus et al., 2010), an explanation of the *mechanisms* of dysfunction remains limited.

In this paper, we discuss some of the main issues that have prevailed within the classification and study of psychopathology and discuss the development of dimensional structural models of psychopathology. We use Caspi et al.'s (2014) seminal work as a basis of this paper due to its popularity and how recent literature has used their findings to further develop the understanding of structural models of psychopathology, thereby facilitating discussion of the development of this literature. We describe the rise of the *p*-factor and

components claimed to be integral to the factor's existence, as well as the debate surrounding the nature of the *p*-factor as a substantive or artefactual construct. We briefly review key neurocognitive accounts relating to the basis of psychopathology and how dimensional models may facilitate the exploration of the association between psychopathology and neurocognitive abilities and describe the *multidimensional hypothesis* (Haywood & Baughman, 2021). This hypothesis is based on the idea that psychopathologies are rarely the consequence of deficits to single neurocognitive mechanisms. Rather that cognitive dysfunctions are more often the outcome of the dynamics of a system comprised of uneven profiles in abilities. We conclude by providing a road forward for a better understanding of the relation between neurocognitive mechanisms and psychopathology.

1.2 Classifying Psychopathology

Griesinger (1817–1868) argued for psychiatric symptoms (or "madness") being the result of a singular disease, and referred to this as the "unitary psychosis" (Rybakowski, 2019). Emil Kraepelin (1856–1926) later devised the Kraepelinian Dichotomy, the characterisation of mental disorder into dementia praecox (to be later reconceived as schizophrenia) and manic-depressive. psychosis (to be later reconceived as bipolar disorder). This dichotomy led to the development of modern diagnostic manuals (Rybakowski, 2019). In the current day, psychopathology is generally defined and determined through a traditional nosological approach, classifying pathology into single, discrete categories (Krueger & Eaton, 2015). The Diagnostic and Statistical Manual (DSM) and the International Classification of Diseases (ICD) have become standard tools used to guide the diagnosis of psychopathology (Clark et al., 1995). However, the reliance on these tools have raised particular issues regarding *comorbidity* and *diagnostic stability*. For example, Newman et al.'s (Newman et al., 1998) work showed that of individuals who meet the diagnostic criteria for one DSM-3 defined disorder, approximately half will meet the criteria for a second, and approximately half of those will meet the criteria of a third disorder, and so on. These issues of comorbidity have also been seen in subsequent issues of the DSM (see (Kotov et al., 2017)). The poor stability of disorder diagnosis is a further issue for the nosological approach. For example, a high proportion of anxiety disorders transition to a different anxiety disorder over a six-year period (Hovenkamp-Hermelink et al., 2016). Aetiological similarities between disorders also suggest that disorders are not so distinct. For example, schizophrenia and bipolar affective disorder share aetiological markers across genetic,

environmental, neurological and cognitive domains (Burdick et al., 2006; Lichtenstein et al., 2009; Smucny et al., 2018). Ultimately, the high level of comorbidity between disorders, in addition to a plethora of biological, cognitive and environmental evidence, suggests that disorders are not as distinct as previously assumed (Krueger & Eaton, 2015). On a practical level, this has many implications. For example, high levels of comorbidity and low levels of diagnostic stability make the study of any individual disorder difficult, as well as complicates treatment decision making (see Newman et al., 1998).

To combat issues of comorbidity and diagnostic stability, and to better facilitate the growing aetiological evidence suggesting low-level mechanistic commonalities, several structural models of psychopathology have been developed. Kotov et al. (2021) integrates the available evidence of structural models of psychopathology, providing a synthesised model. However, the size and specifications of this structure makes it difficult to test and use in its entirety. Structural models view psychopathology as dimensional and explore hierarchical relationships among psychopathological symptoms to develop subordinate and superordinate components of psychopathology. Furthermore, it has been suggested that these dimensional models may be used to inform treatment by basing and prioritising treatment decisions on the symptom dimensions at the various levels of the models' hierarchy (see Kotov et al., 2021; Ruggero et al., 2019). The most prominent models were developed through Caspi et al.'s (2014) longitudinal research. This research saw the development and assessment of hierarchical models of psychopathology that are claimed to enhance our understanding of disorders.

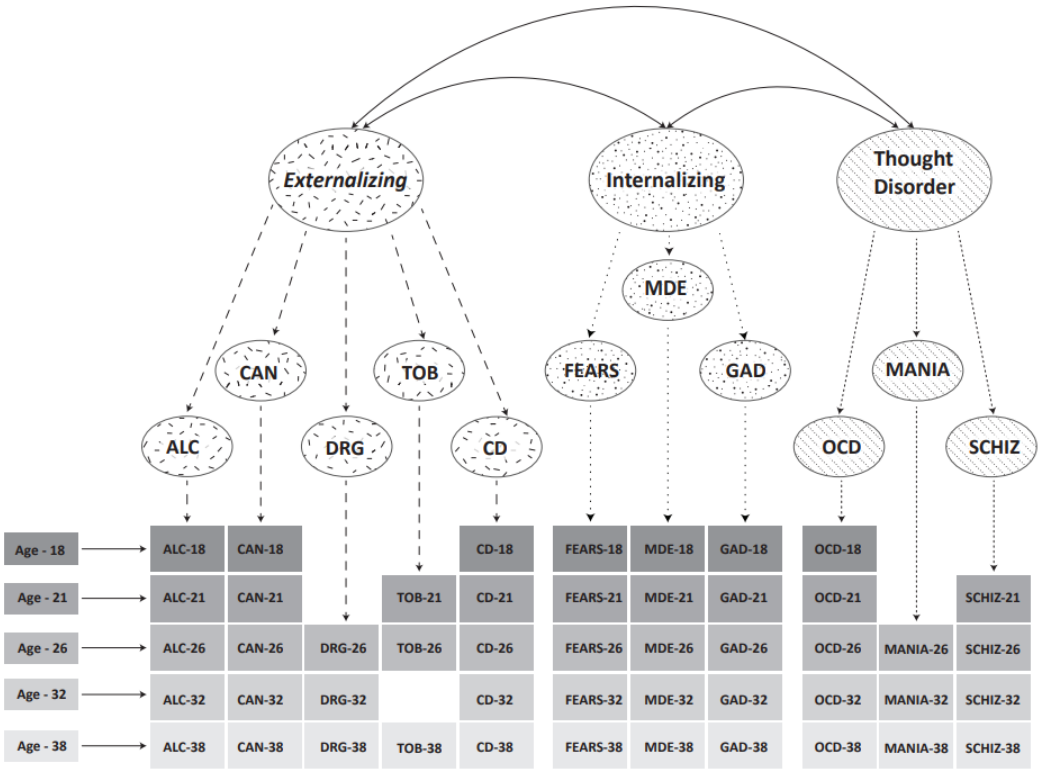
1.3 Caspi et al.'s Structural Models of Psychopathology

Caspi et al. (2014) used data from the Dunedin Multidisciplinary Health and Development Study consisting of a battery of biological, developmental, clinical, personality and neurocognitive measures administered to a representative community sample of 1000 participants across a total of 11 time points over a 35-year period (ages 3, 5, 9, 11, 13, 15, 18, 21, 26, 32 and 38). Using the Diagnostic Interview Schedule (Robins et al., 1995), clinicians counted the number of symptoms each participant reported in accordance with 11 predetermined, common, DSM defined disorders at five time points (ages 18, 21, 26, 32 and 38). Disorders and symptomology assessed included various substance use disorders (e.g., alcohol, cannabis, tobacco), conduct disorder, major depressive episode, fears and phobia symptoms, obsessive compulsive disorder, mania symptoms, and schizophrenia (Caspi et al., 2014). Caspi et al. (2014)

showed that the array of symptoms could be reliably fit to a correlated factors model, with factors pertaining to symptom counts of each disorder over time, and three higher-order factors called *internalising*, *externalising* and *thought disorder*s (see Figure 1.1). Figure 1.1 shows the 11 disorder symptom counts over time, loading onto their specific disorder factor, representing longitudinal symptomology. Figure 1.1 also shows the disorder specific factors then further loading onto one of the three higher-order factors of psychopathology.

Figure 1.1.

Correlated Factors Model Adapted from Caspi et al. (2014).

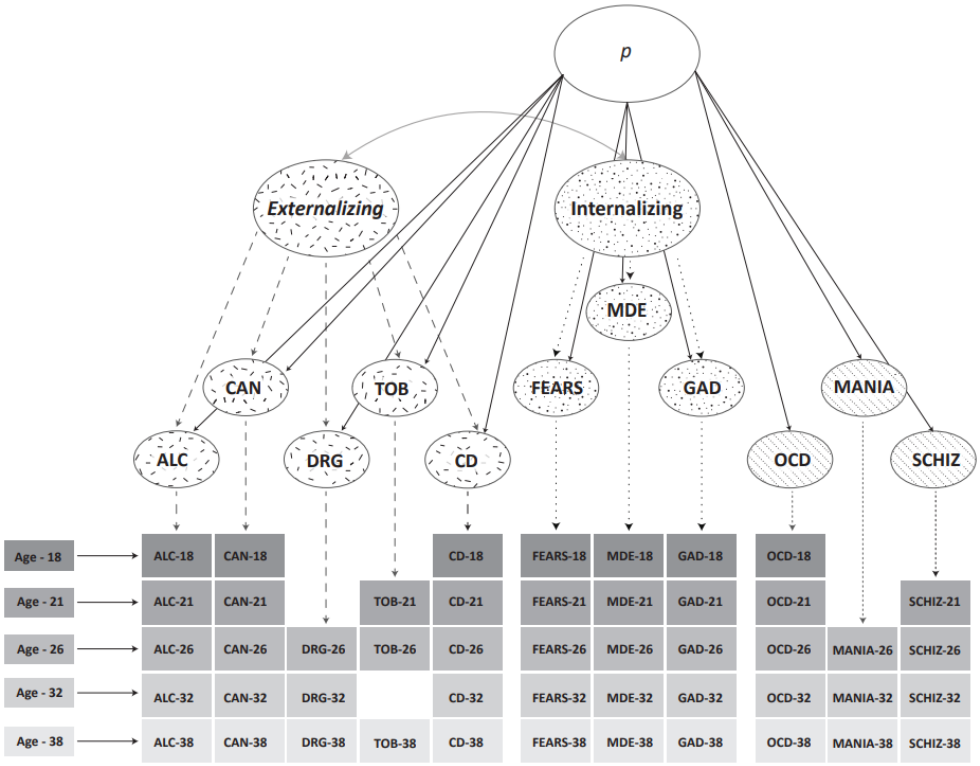


Notes. Ovals represent latent symptom factors; boxes represent the symptoms related to each disorder. The 11 disorders included in the model are: ALC = Alcohol. CAN = Cannabis. DRG = Hard drugs. TOB = Tobacco. CD = Conduct disorder. MDE = Major depressive episode. GAD = General Anxiety Disorder. OCD = Obsessive Compulsive Disorder. SCHIZ = Schizophrenia.

Caspi et al. (2014) tested another structural model of psychopathology, called the bifactor model. The bifactor model contained not only the disorder specific and higher-order factors but also a single General Psychopathology factor (see Figure 1.2). The addition of the general factor accounted for further symptom variance among all disorders included in the model, over and above that only accounted for by the internalising, externalising and thought disorder factors. The thought disorder factor was subsumed by the introduction of the general factor and so was subsequently removed from the model (Caspi et al., 2014). The bifactor model was found to be a better for the data than the correlated factors model, and has subsequently become highly popular in psychiatric and psychological research.

Figure 1.2.

Revised Bifactor Model Adapted from Caspi et al. (2014).



Note. Ovals represent latent symptom factors; boxes represent the symptoms related to each disorder. The 11 disorders included in the model are: ALC = Alcohol. CAN = Cannabis. DRG = Hard drugs. TOB = Tobacco. CD = Conduct disorder. MDE = Major depressive episode. GAD = General Anxiety Disorder. OCD = Obsessive compulsive disorder. SCHIZ = Schizophrenia.

The general psychopathology factor was named the *p*-factor, in line with its likeness to the *g*-factor, originating in the literature on intelligence. Indeed, like the *g*-factor, Caspi et al. (2014) argue that, conceptually, the *p*-factor is normally distributed within the population. Caspi et al. (2014) took the view that general psychopathology factor is a substantive construct that determines the presence and absence of all pathological symptoms. Crucially, the higher the *p*-factor, the greater the propensity towards psychopathology. Overall, Caspi et al. (2014) found that their bifactor model successfully accounted for psychopathology in a hierarchical manner. The *p*-factor accounts for the common variance of all psychopathological symptoms, while the internalising and externalising factors account for the remaining common variance of a subgroup of similar disorders, and lastly, the disorder specific factors account for symptom variance that is unique to each disorder. Recent research has also found that the *p*-factor is also supported in multi-method multi-trait modelling, providing evidence that the *p*-factor is not just the result of common method variance from the included indicators (Watts et al., 2022). Following the work of Caspi et al. (2014), a range of research attempted to discover what the substantive *p*-factor is (see Watts et al., 2020a). In other words, various theoretical explanations occurred regarding the substantive meaning of *p*.

1.4 What is the *p*-Factor?

As the *p*-factor is claimed to determine an individual's overall propensity toward psychopathology (Caspi et al., 2014), knowing the substantive meaning of *p* has potentially important implications for the understanding and treatment of psychopathology. A range of conflicting research has laid claim to the substantive meaning of *p*. For example, research has evidenced neuroticism as the primary driver of the *p*-factor (Brandes et al., 2019), and other research has offered that *p* represents functional impairment (Smith et al., 2020), impulsive responsivity to emotion (Carver et al., 2017), or disordered thought (Caspi & Moffitt, 2018). Each of these proposals make conceptual sense. However, each of the explanations are of high-level psychological domains, underpinned by a range of other mechanisms. Therefore, other lower-level mechanisms, in particular neurocognitive abilities, have been claimed to be a primary driver of the general factor (see Heinrich et al., 2020).

Indeed, each explanation for the *p*-factor, neuroticism, functional impairment, impulsive responsivity to emotion, and disordered thought are significantly accounted for by a range of neurocognitive abilities (Caspi et al., 2014; Caspi & Moffitt, 2018; Crow, 2019; Smith et al.,

2020). Furthermore, neurocognitive abilities are also significantly associated with the internalising, externalising and thought disorder factors in Caspi et al.'s (2014) correlated factors model, as well as the *p*-factor of psychopathology in the bifactor model. In fact, Caspi et al. (2014) found that, that within the tested age range, from ages 3 to 38, every direct measure of neurocognitive ability was significantly associated with the *p*-factor. Furthermore, a systematic review of risk factors predictive of the statistically derived factors of psychopathology in young people found deficits in neurocognitive abilities to be a primary risk factor for higher psychopathology factor scores (Lynch et al., 2021). Ultimately, there is evidence that neurocognitive abilities are not only related to diagnosed pathologies, but even within the general population, neurocognitive abilities are related to the proposed *propensity toward psychopathology* (Caspi et al., 2014). The importance of neurocognitive abilities in the understanding of the *p*-factor, and bifactor models of psychopathology generally, has been communicated early on from Caspi et al.'s (2014) longitudinal work. For example, Snyder et al. (2015) proposed that exploring associations between Caspi et al.'s (2014) bifactor model and executive functioning "…has the potential to greatly clarify the nature of EF impairments associated with particular forms of psychopathology, and thus accelerate progress in understanding how EF impairments may contribute to both comorbidity across disorders and heterogeneity within disorders…" (p. 17). To set the scene for our discussion of neurocognitive abilities and structural models of psychopathology, in the following section we summarise the primary neurocognitive abilities used in clinical research, as well as each ability's association with Caspi et al.'s (2014) components of psychopathology.

1.4.1 Neurocognitive Abilities as Important to the Factors of Psychopathology

Cognition partly consists of higher-level processes and components. These components include problem solving abilities and the control of attention, among other higher-level human abilities (Lezak et al., 2004; Norman & Shallice, 1986). Baddeley (1992) famously proposed a single component, termed the executive, which governs, organises and controls high-level abilities. Various accounts of the executive exist, raising contention as to whether the executive is a unitary component or a collection of components (e.g., Miyake, Emerson, et al., 2000; Norman & Shallice, 1986; Zelazo et al., 1997). However, there is broad agreement regarding the existence and importance of the executive(s) as fundamental to the control of cognition. Recently, the term executive functioning has become the norm to describe these control

processes, and most work has been focused on the three executive functioning components described by Miyake, Emerson, et al. (2000) namely, updating, shifting and inhibition. The role of each executive function component differs. Updating is considered to be involved in the removal, addition and monitoring of the contents of working memory; shifting is involved in disengaging with the present mental set and engagement with a more relevant mental set, while inhibition is described as the process that suppresses a dominant response that is not currently useful (Miyake, Emerson, et al., 2000). Additionally, other singular neurocognitive components have been considered in both the theoretical and empirical domains. The most prominent of these neurocognitive components include speed of processing (Salthouse, 1996) and working memory capacity (Baddeley, 1992). Speed of processing relates to the speed at which individuals can process information (Salthouse, 1996), while working memory capacity is considered as the amount of information that can be held in working memory and is often conceptualised and measured in accordance with working memory updating (Baddeley, 1992; Miyake, Emerson, et al., 2000).

Lezak et al. (2004) argue that the proper functioning of these neurocognitive abilities, including executive function, are crucial to everyday behaviours, including the control of appropriate, goal-oriented and responsible behaviour. It is perhaps therefore unsurprising that abnormalities in these processes have been repeatedly indicated in a variety of psychopathologies.

1.4.2 Deficits in Neurocognitive Processes and Their Relation to Psychopathology

It is often suggested that deficits in neurocognitive abilities underlie pathological symptoms across Caspi et al.'s (2014) factors of psychopathology. The following subsections present examples of symptoms of internalising, externalising and thought disorders that have been suggested to be underpinned by neurocognitive abnormalities.

1.4.2.1. Internalising

Neurocognitive deficits have been proposed to underlie a range of symptoms associated with internalising disorders. Deficits in the updating and capacity of working memory have been suggested to be central to elevated rumination in depression, due to issues in removing negative material from working memory (Joormann & Gotlib, 2008). Similarly, deficits in shifting mental set are claimed to underlie issues in shifting attention away from negative thoughts and stimuli in anxiety (Johnson, 2009). Inhibition has been seen to be an important aetiological mechanism in a

range of internalising symptoms. For example, depression is often accompanied by a range of negative attentional and memory biases and deficits in inhibition that are proposed to underlie this issue (Gilleen et al., 2016). Another salient symptom of depression is a general cognitive slowing (Tsourtos et al., 2002), and this often has a great impact on the life of the person and is said to be underlain by speed of processing deficits (Tsourtos et al., 2002).

1.4.2.2. Externalising

Externalising disorders, including behavioural and substance use disorders, are strongly associated with a range of neurocognitive abilities. For example, working memory deficits are said to mediate disinhibited decision making in externalising disorders (Endres et al., 2014; Endres et al., 2011). Another example is that deficits in shifting mental set are claimed to underlie the poor consideration of behavioural outcomes in substance addiction (Noël et al., 2013). Furthermore, the uncontrolled intake of substances has also been claimed to be associated with deficits in inhibition (Mahmood et al., 2013). Speed of processing is often associated with aspects of behavioural disorders such as attention-deficit/hyperactivity disorder. For example, speed of processing issues are claimed to underlie reading fluency issues often seen in ADHD (Jacobson et al., 2011; Shanahan et al., 2006).

1.4.2.3. Thought Disorder

Thought disorders, such as schizophrenia and mania in bipolar disorder, have been subject to a large amount of neurocognitive research. There has been broad suggestion that deficits in a variety of neurocognitive abilities are important mechanisms of the aetiology of thought disorder symptoms. For example, the difficulties people with schizophrenia have in engaging with the environment may be due to working memory deficits, resulting in a lack of flexibility toward environmental stimuli (Galletly et al., 2007). Deficits in shifting mental set are also proposed to underlie the level of insight into their disorder that people with schizophrenia have (Gilleen et al., 2016), and episodes of mania in bipolar disorder are accompanied by mental set shifting deficits (Kurtz & Gerraty, 2009). Schizophrenia is often accompanied by a range of behavioural issues and deficits in inhibition that are often claimed to be central to these issues. For example, deficits in inhibition are said to be deterministic of the poorly planned and impulsive behaviour in schizophrenia (Kiehl et al., 2000). Furthermore, deficits in a person with schizophrenia's speed of processing has been seen to mediate these broad neurocognitive deficits (Rodríguez-Sánchez et al., 2007).

Ultimately, neurocognitive abilities seem to be fundamental to understanding psychopathology symptoms across Caspi et al.'s (2014) internalising, externalising and thought disorder components. General deficits in neurocognitive abilities are robustly associated with a range of psychopathologies and their symptoms. It is important to remember, however, that Caspi et al.'s (2014) model is ultimately a description of psychopathological behaviours that often co-occur. Caspi et al. (2014) proposed that exploring how internalising, externalising and substantive *p*-factor comes to exist will require a range of measurements across biological, cognitive and environmental domains. Therefore, to examine what the hierarchical components of domains such as Caspi et al.'s (2014) bifactor model represent, a mechanistic approach exploring the association between domains such as neurocognitive abilities and the factors are required.

1.5 A Mechanistic Approach

A mechanistic alternative to descriptive models of psychopathology comes from the Research Domain Criteria (RDoC; Cuthbert, 2020; Cuthbert & Insel, 2013). The RDoC Framework reverses Caspi et al.'s (2014) top-down processes to describing psychopathology by starting with the consideration of how genetic, neurological and cognitive variation can give rise to the occurrence of psychopathological symptoms. The RDoC framework has led to programmes of research that have advanced our knowledge of the mechanisms that might underlie psychopathology (e.g., Clarkson et al., 2019; Ip et al., 2019). However, the RDoC approach is also not without limitations. Kotov et al. (2021) argue that by disregarding clinical phenotypes, and basing the exploration of psychopathology at the most basic levels, the RDoC framework has little current clinical utility. Kotov et al. (2017) and Patrick et al. (2013) suggest that the weaknesses of both the symptomatic based hierarchical structures, such as Caspi et al. (2014), and the weakness of the lower-level, mechanistically oriented RDoC framework, can be reconciled by combining the approaches. It has been suggested that joining symptomatic psychopathology structures with the RDoC constructs is likely to result in mechanisms that are measurable, consistent and explanatory of the phenotypes of psychopathology (Kotov et al., 2017; Patrick et al., 2013).

Linking descriptive (e.g., Caspi et al., 2014) and mechanistic approaches (e.g., RDoC; Cuthbert, 2020) to psychopathology requires the use of domains that are robustly associated with psychopathology at both the lower (e.g., chemical, genetic and neurological) and higher (e.g.,

psychopathological symptoms) levels. Neurocognitive abilities are included as one of the key domains in the RDoC system, as these abilities are associated with a wide range of psychopathology at each level of analysis (Genes, Molecules, Cells, Circuits, Physiology, behaviour and self-report; see Cuthbert, 2020; Cuthbert & Insel, 2013). Furthermore, neurocognitive abilities are also significantly associated with the internalising, externalising and thought disorder factors in Caspi et al.'s (2014) correlated factors model, as well as the *p*-factor of psychopathology in the bifactor model. Therefore, satisfying both criteria, neurocognitive abilities are an excellent candidate for joining the two approaches to the study of psychopathology. However, our ability to successfully link these two approaches relies on developing a thorough understanding of the meaning of *p*, and the specific factor of psychopathology. Recent literature has uncovered a range of methodological and conceptual issues that have important implications for the use of *p* as a substantive construct.

1.6 *p,* Substantive Factor, or Statistical Artifact?

In recent years several important questions and critiques have been made regarding the structural approach to psychopathology; many of these have important implications towards using these frameworks when exploring what may underpin psychopathology. Recent literature explores the question of if the *p*-factor is a substantive, meaningful construct, or rather simply a statistical artefact derived from the characteristics of the methods used. Snyder and Hankin (2017) explain that the general factor of psychopathology is dependent on the characteristics of its makeup, and therefore is an inherently inconsistent construct. Lahey et al. (2021b) describes *p* as the "weighted average" (p. 61) of the symptoms of a sample at that point in time. This conflicts with *p* being a potentially substantive construct with a consistent meaning and interpretation. Furthermore, Levin-Aspenson et al. (2020) explored the applicability of the *p*-factor among different samples. Levin-Aspenson et al. (2020) used three large datasets to conduct their exploration: (1) the National Comorbidity Survey (Kessler & Merikangas, 2004), (2) Collaborative Psychiatric Epidemiology Surveys (Heeringa et al., 2004), and (3) the Methods to Improve Diagnostic Assessment and Services (Zimmerman, 2016). The first two being large (N = 8098 and N = 19,823, respectively) epidemiological datasets, and the third being a large dataset (N = 2900) from an outpatient psychiatric hospital. The authors found bifactor models to be a good fit in each population; however, the loadings of the disorders on the *p*-factor varied extensively across the populations. Furthermore, issues have been raised regarding the indices

often used to justify the *p*-factor. Greene et al. (2019) assessed the possibility that the better fit generally found by bifactor models (those that include the *p*-factor) over correlated factor models (with no *p*-factor) may simply be due to fit indies unfairly biasing the bifactor models. This may mean that, even though bifactor models tend the fit collections of diagnoses and symptoms best, this may not be due to any substantive reason. Greene et al. (2019) found data simulated from a correlated factors model most often better fit a bifactor model rather than a correlated factors model through which the data was created. Greene et al. (2019) called for the selection of a model of psychopathology to be based on substantive interpretability and the utility of the model to facilitate the goals of the research, rather than model fit.

The applicability and substantive meaning of the *p*-factor, as well as externalising, internalising and thought disorder factors needs also to be considered *within* samples. Given that the factors of psychopathology are derived from covariation amongst of psychopathological symptoms *across* the sample, the applicability and substantive meaning of those factors will likely vary greatly for subgroups and individuals *within* the sample. While *p* and the other factors of psychopathology might do well at summarising symptomology for the population, they may be of substantially less utility for a substantial number of individuals within that sample. This consideration means it is difficult to conclude what may underpin psychopathology on the individual and subgroup level and any conclusions made might lead us astray. For example, Caspi et al. (2014) found the large majority of measures of neurocognitive ability to be significantly associated with externalising, internalising, and thought disorder factors. However, with the introduction of the *p*-factor in the bifactor model, the associations between the measures of neurocognitive ability and the internalising, externalising and thought disorder factors almost all fell to non-significant, and instead each measure of neurocognitive ability was significantly associated to the *p*-factor. This might lead us to the conclusion that neurocognitive ability has the greatest importance to psychopathology at the *p*-factor level. However, it is likely that a number of individuals in Caspi et al.'s (2014) sample had a high number of psychopathological symptoms, but a low *p* score, due to, for example, a lack of general comorbidity of symptoms and a different pattern of symptoms to the mean. We might then naively assume, due to the importance of neurocognitive abilities to psychopathology seemingly being at the *p*-factor level, that neurocognitive abilities may not be important to understanding this person's psychopathology. To summarise, the primary limitations of CFA structural models are as

follows: (a) it is unclear if the factors of psychopathology have, or can have, universal substantive meaning, (b) fit indices often used to champion one model over another are biased toward bifactor models, and (c) the applicability and consistency of structural models within subgroups of a population is not currently known.

Ultimately, for the *p*-factor to be useful in the exploration of neurocognitive abilities and psychopathology, a fuller understanding of the characteristics of the factor is needed. The methodological and conceptual issues of substantive *p* have led to a host of authors calling for a consensus on a definition on what the *p*-factor is, as well as an agreement on what should predict the general factor, and what the general factor should predict, in order to establish the factor as a substantive construct (Fried et al., 2021; Greene et al., 2019; Levin-Aspenson et al., 2020; Watts et al., 2020a). Further, other authors have argued for an alternative model to mitigate the fluidity of a general factor of psychopathology (Eid, 2020; Haywood et al., 2021a; Heinrich et al., 2020).

1.6.1 An Alternative Approach

The issues of developing a universal substantive *p* have led some authors to prioritise an alternative structural model, called the S-1 bifactor model (Burke & Johnston, 2020; Eid, 2020; Heinrich et al., 2020). The S-1 bifactor model is named as such due to it containing one less specific factor than standard bifactor models (Eid, 2020). In a traditional bifactor model, each indicator loads onto the general factor, as well as one specific factor. However, in an S-1 bifactor model, a chosen set of indicators does not load onto any specific factor and only loads onto the general factor. Eid (2020) describes these indicators as being the ‘reference domain’. The reference domain, as it only loads onto the general factor, and ‘becomes’ or defines that factor. Therefore, a researcher can pre-specify precisely what the general factor represents, circumventing the issues with an undefined general factor (e.g., the *p*-factor). The variance in an S-1 model’s specific factors reflects the common variance amongst the factor indicators after taking into account the general factor (Eid, 2020). The reference domain, and therefore the general factor, can reflect any theoretically outstanding variable of interest (Heinrich et al., 2020).

Interestingly, some traditional bifactor models have ended up transforming to S-1 bifactor models unknowingly. For example, Heinrich et al. (2020) showed that when Caspi et al. (2014) removed the thought disorder factor from their bifactor model due to a Heywood case, they turned their model into an S-1 bifactor model, as OCD, mania and schizophrenia loaded onto the

p-factor and no specific factor. Thought disorder, therefore, came to represent the general factor, and the *p*-factor was therefore not an indication of general psychopathology.

S-1 bifactor models may offer a useful way to explore how neurocognitive abilities are associated to psychopathology. It is possible to use a range of measures of neurocognitive abilities as direct indicators of the general factor, thereby defining its meaning (Haywood et al., 2021a). This could provide information that other approaches could not. For example, it would then be insightful to examine the unique variance of each symptom indicator, as well as the variance within each specific factor, after accounting for the general (neurocognitive) factor. The S-1 approach could be used with a correlated factors model to provide more information regarding the associations of specific neurocognitive components. The S-1 bifactor approach has promise for advancing our understanding of neurocognitive abilities association to psychopathology across a sample; however, it means the rejection of a general factor of psychopathology and limitations in accounting for the heterogeneity among the associations between neurocognitive abilities and psychopathology that characterise this research. The heterogeneity of neurocognitive abilities' association with psychopathology is key to developing a nuanced or mechanistic understanding of the aetiology of symptoms (Haywood & Baughman, 2021). Therefore, it is important to consider this variation and the approaches most suitable for its exploration.

1.7 Heterogeneity of Psychopathology and Neurocognition

Associations between neurocognitive abilities and psychopathologies, a direct one-to-one correspondence, or perfect association, between neurocognitive abilities and the psychological, behavioural and biological components of psychopathologies has never been found. Therefore, a neurocognitive ability cannot be seen as deterministic of psychopathology. A large body of literature has explored the specific causes of disorders across biological and cognitive mechanisms. However, finding singular mechanisms with a one-to-one, deficit-diagnosis correspondence with a disorder has been elusive. For example, at the biological level, the search for specific genes with a one-to-one correspondence with a disorder has been met with limited success (e.g., Ripke et al., 2014). The *COMT* gene, while reliably shown to be associated with a variety of disorders, does little to account for the phenotype of a disorder on an individual level (Egan et al., 2001; International Schizophrenia, 2009). Similarly, across each level of biological analysis, heterogeneity on an individual level is the rule rather than the exception (e.g., Cowen,

2016)). This means that, while certain variations may be associated with a disorder (or multiple) at a population level and may increase the risk of developing the symptoms of a disorder, that variation is not *deterministic* of psychopathology.

Cognitive endophenotype approaches have been used in attempts to uncover underlying biological mechanisms of disorders (Snyder et al., 2015). If performance on a particular neurocognitive task is associated with the genetic basis of the disorder (i.e., poor performance is seen in people with the disorder, as well as their healthy first-degree relatives), then it may be reasonable to assume that the specific neurocognitive mechanisms underlying performance on that task can be deduced to a biological basis of the disorder. However, this approach has also been met by the problem of inconsistency. For example, associating specific components of neurocognition with endophenotypic markers of psychopathology has been mixed. For example, greater than average perseveration errors on the Wisconsin Card Sorting Task are found for people with schizophrenia and their first-degree relatives (Stefanopoulou et al., 2009; Szöke et al., 2005). However, even though the WCST is a general executive function task involving the use of updating, shifting and inhibition, there is research crediting each of these components as primarily determining the amount of perseveration errors performed on the task (Barceló & Knight, 2002; Gamboz et al., 2009; Hartman et al., 2001; Manoach et al., 2002). This makes deducing the biological basis of the specific neurocognitive components contributing to perseveration errors impossible. A multitude of studies have explored the neurocognitive heterogeneity of singular disorders. At a population level, there are clear general neurocognitive deficits among psychopathologies. However, at an individual level the precise neurocognitive components that are deficit range dramatically. For example, Martino et al. (2008) found that, within bipolar disorder, 38% were not deficit in any neurocognitive domain, 40% were deficit in one to two domains and 22% were deficit in three to four domains, and the disorder was not deterministic of a deficit in any particular neurocognitive domain. Raffard and Bayard (2012) found similar heterogeneity in people with schizophrenia. Ninety four percent of people with schizophrenia had deficits in at least a single neurocognitive task, 27% showed deficits in two tasks 23% showed deficits in three tasks, while 23% showed deficits in four neurocognitive tasks (Raffard & Bayard, 2012). Furthermore, functioning in these neurocognitive domains is generally not associated with duration of the illness, current psychoticism status or medication (e.g., Raffard & Bayard, 2012). Even when comparing disorders, deficits in particular

neurocognitive domains that are able to separate the disorders are generally not uncovered (Moritz et al., 2002). The heterogeneity of the mechanisms of psychopathology has led to the call to disband the medically derived *cause model* when exploring psychopathology (e.g., Bringmann & Eronen, 2018). However, the question of "where to from here" is still unclear.

1.7.1 Multiple Realisation and Psychopathology

Perhaps embracing heterogeneity in the study of neurocognitive abilities and psychopathology, rather than seeing it as an error or something that should be minimised, would lead to a greater understanding of their associations. The notion of *multiple realisation* comes from the philosophy of mind that postulates that a mental state, event or component can be determined by *multiple different* biological states, events or components (Putnam, 1988). It has been proposed that wide, varied physicalities can each experience the same mental state, event or component form and yet share no physical similarities. For example, it is generally accepted that a wide range of creatures such as humans, birds, molluscs and amphibians experience pain, yet these creatures often share very few physical properties (Putnam, 1988). Pain can therefore be multiply realised by many different physical states, events or components. Originally based to combat reductionism, the postulate of multiple realisation has been applied to many subjects, including psychopathology.

Multiple realisation is useful in explaining the biological heterogeneity of disorders. It gives us an idea with which to explain the lack of success in finding specific biological mechanisms underlying psychopathology and provides a platform to separate mental and physical states, events or components. There has been a range of support for this concept through different methods. For example, Pavão et al.'s (2015) computational work found that 154 computational models, each representing a different grouping on brain alterations, produced activity that represented the neural activity of schizophrenia.

Application of multiple realisation at the cognitive level may also provide a platform to explain the heterogeneity of neurocognitive abilities in psychopathology. Might we also extend multiple realisation to include the same set of realisers, but at various levels of functioning? Haywood and Baughman (2021) termed this proposal as the *multidimensional hypothesis.* The multidimensional hypothesis states that various different neurocognitive components, each with different ability levels (i.e., strengths and weaknesses) can explain a psychopathological phenotype equally well (see Haywood & Baughman, 2021 for a detailed explanation).

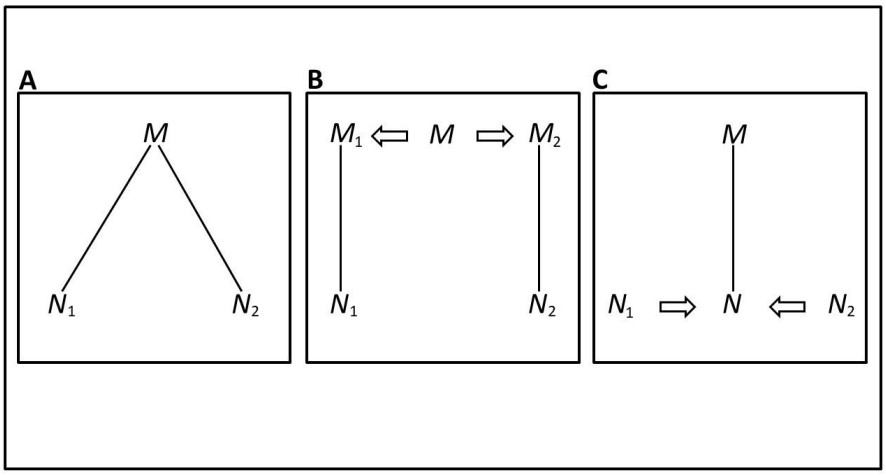
The hypothesis posits that the overall neurocognitive ability of a person, or their susceptibility to a psychopathology, cannot be explained by a single neurocognitive component, nor can it be explained by a linear, additive model. Instead, the importance lies within the *non-linear interactions between the neurocognitive components' abilities*. Testing this hypothesis, Haywood and Baughman (2021) proposed that the high amount of perseveration errors performed on the Wisconsin Card Sorting Task by people with schizophrenia and their first-degree relatives could be multiply realised by various different ability combinations among the neurocognitive components updating, shifting and inhibition. Applying computational methods, Haywood and Baughman (2021) found that the performance on the task of people with schizophrenia, their first-degree relatives and control participants' could be simulated by computational models with different levels of abilities of updating, shifting and inhibition. This suggests that general neurocognitive ability, a robust endophenotype of psychopathology, may be better explained by the interactions among neurocognitive components rather than primarily by a single deficit, thus explaining the inability to find a consistent neurocognitive ability deficit throughout individuals with a certain disorder.

It is important to note that, over time, the fundamental postulates of multiple realisation have been questioned (e.g., Bechtel & Mundale, 1999; Bregant, 2006; Polger & Shapiro, 2016)). It has been suggested that many cases of multiple realisation (Figure 1.3*A*) can be explained by either splitting the mental state, event or component into two or more states, events or components (Figure 1.3*B*), or merging the realisers (Figure 1.3C; see Pernu, 2019 for a summary). Splitting is done if it is found that the mental state, event or component is better seen as multiple. Take, for example, if (M; Figure 1.3*B*) working memory is split into (M1) working memory capacity and (M2) working memory updating, we might find each mental component to be realised by separate physical properties (i.e., N1 and N2, respectively). Merging is done if the realisers are found to be the same physically (see Figure 1.3*C*). Pernu (2019) provides the example that the intention to grasp an object (M; Figure 1.3*C*) can be found with the mean neural activity of some specific neuronal structures (N1 and N2). N1 and N2 in this case will be *merged*, resulting in a singularly realised (N) component. However, it seems that there are many contexts in which neither splitting nor merging can be easily applied and conform to existing

empirical evidence (e.g., hunger; see Putnam, 1988); in these cases, multiple realisation gives us a useful platform to understand heterogeneity.[1]

Figure 1.3.

Splitting and Merging Realisers



Note. Depiction of splitting and merging realisers accounting for multiple realisation. Panel A: Working memory (*M*) being realised by two separate physical properties (N_1 and N_2, respectively). Panel B: Working memory (*M*) being split into working memory updating (M_1) and working memory capacity (M_2), M_1 and M_2 being realised by two separate physical properties (N_1 and N_2, respectively). Panel C: The intention to grasp an object (*M*) being realised by the mean neural activity of some specific neuronal structures (N_1 and N_2), resulting in a singular unified realiser (*N*).

[1] There is also an inverse proposition of multiple realisation, namely reverse multiple realisation (Pernu, 2019). Reverse multiple realisation is the claim that the same physical states, events or components could realise different mental states, events and components. Bringing this concept to psychopathology would suggest that the same biological properties could underlie different mental disorders. Pernu (2019) points out that reverse multiple realisation has support within the neuroplasticity and neural reuse literature. For example, Anderson (2010) illustrates that neural circuits can be deployed, over time, for a different purpose if the need arises. Therefore, the same physical states, events or components can realise multiple different mental states, events or components.

Ultimately, the notion of multiple realisation provides a platform for questioning the relationship between neurocognitive abilities and psychopathology. Other than the preliminary evidence in support of the multidimensional hypothesis (Haywood & Baughman, 2021), little is known about the applicability of multiple realisation to solely cognitive and psychological states, events or components. That is, can a psychological or cognitive state, event or component be multiply realised by various other cognitive or psychological states, events or components at different levels of functioning? Further research exploring neurocognitive variability in psychopathology is therefore needed.

Another possibility is that the heterogeneity of neurocognitive abilities within disorders may be minimised when assessing statistically derived symptomatic components of psychopathology (i.e., internalising, externalising and thought disorder), rather than DSM defined disorders with a great level of overlap. For example, there may be clear patterns of neurocognitive ability profiles within the internalising, externalising and thought disorder components of psychopathology, and these patterns may help explain those symptom clusters and their aetiology. However, the neurocognitive heterogeneity within DSM disorders might also be seen in the statistically derived components of psychopathology, as per the multidimensional hypothesis (Haywood & Baughman, 2021).

1.8 The Road Forward

Structural models of psychopathology provide a promising framework to advance our understanding of the relation between neurocognitive abilities and psychopathology. Finding reliable, specific associations or patterns of association, and supporting causal explanations between neurocognitive abilities and psychopathology, is unlikely if explorations continue to be based upon DSM/ICD defined disorders. Take, for example, that within the DSM there are a total of 227 different possible symptom combinations that fulfil the criteria for a diagnosis of major depressive disorder (Park et al., 2017). Therefore, at the phenotype level, the symptom heterogeneity and lack of stability as well as the comorbidity between disorders, means that consistent associations between neurocognition and DSM/ICD defined disorders are unlikely. However, as per the call of Levin-Aspenson et al. (2020), a consensus around the substantive meaning of the *p*-factor is needed. The uncertainty of the meaning and applicability of the factors of psychopathology greatly limits our confidence to draw conclusions. Future research should first assess the applicability of the factors of psychopathology within subgroups of a community

sample. This will advance our understanding of how the sample derived factors of psychopathology reflect individuals and subgroups within the sample. Future research should also assess how neurocognitive abilities are related to these factors of psychopathology *within* the subgroups and explore association differences *between* these subgroups. The differences in the associations between the subgroups may illuminate, not only the applicability and utility of the factors within subgroups, but also provide useful knowledge on the substantive meaning of the factors and how this might differ depending on subgroup and population.

S-1 bifactor models offer a promising method of explaining neurocognitive abilities' associations with psychopathology, while mitigating the questionable substantive validity of the undefined *p*-factor (Haywood t al., 2021a). The S-1 bifactor model, supplemented by the correlated factors model, seems particularly useful at the population level for examining how cognitive functioning may be associated with psychopathology. While the traditional use of the structural approaches are limited in explaining the neurocognitive heterogeneity within psychopathology, is it possible to use these approaches to elucidate potential neurocognitive performance patterns within psychopathology. If somewhat reliable patterns of associations, and causal accounts between neurocognitive components and psychopathology, are to be supported, it might be at the level of internalising, externalising and thought disorders, rather than at the level of individual disorders. However, neurocognitive abilities' association with psychopathology may also be explained by non-linear dynamic interactions as implied by the multidimensional hypothesis (Haywood & Baughman, 2021), and the two possibilities are not mutually exclusive. The non-linear multidimensional conceptualisation is applicable at the internalising, externalising, and thought disorder, and *p*-factor levels, and it will be further supported if it is seen that, even within statistically derived components of psychopathology, non-linear multidimensional interactions are superior to linear conceptualisations at predicting psychopathology. Ultimately, in our view, to progress knowledge about the underpinnings of psychopathology, future research should:

a. Further examine if a universal substantive *p* (and specific factors) could be developed by the assessment of the utility and consistency of structural models of psychopathology in subgroups.
b. Use the S-1 bifactor model in explorations of neurocognitive ability and psychopathology.

c. Assess if the factors of psychopathology can be partly explained by discrete association patterns of neurocognitive performance.
d. Assess if each factor of psychopathology (e.g., internalising, externalising, thought disorder and the *p*-factor) is explained by multidimensional interactive components of neurocognition.

1.9 Thesis Outline and Aims

The overarching objective of this thesis is to provide an improved understanding of the associations between neurocognition and psychopathology, taking into consideration the rise of dimensional and structural psychopathology. The findings of the thesis will inform theoretical understandings of the dynamics of neurocognition and psychopathology as well as advance our understanding of the applied relevance of the association. In this thesis, I use a simulation, computational, and empirical techniques to achieve the overarching objective. Underlying this objective are four specific aims of this thesis:

(1) To further examine if a universal substantive *p* (and specific factors) could be developed by the assessment of the utility and consistency of structural models of psychopathology in subgroups.
(2) To use the S-1 bifactor model in explorations of neurocognitive ability and psychopathology.
(3) To assess if the factors of psychopathology can be partly explained by discrete association patterns of neurocognitive performance.
(4) To assess if each factor of psychopathology (e.g., internalising, externalising, thought disorder and the *p*-factor) is usefully explained by multidimensional interactive components of neurocognition.

Each of the thesis aims will be explored within a separate chapter. The thesis is comprised of four studies, one examining each aim. Preceding each chapter is a preface, which provides an outline, rationale, and goal of the chapter.

Preface to Chapter 2

In chapter two the first direction of future research in chapter one of our road forward for the exploring of psychopathology and neurocognitive abilities is examined. This direction of future research coincides with the first aim of the thesis. Chapter two examines the important question of whether a universal substantive *p*-factor (and specific factors) could be developed by assessing the applicability, loading pattern consistency, and neurocognitive correlates of four popular structural models of psychopathology within population subgroups. We use data simulation methods to create a large dataset that is capable of being split into groups of individuals with various levels and types of symptom heterogeneity. We then assess the applicability of the structural models within these subgroups and examine the consistency of factor loadings and neurocognitive correlates between the groups. This chapter,, therefore, provides important information for the potential development of a consistent substantive *p*-factor with which associations between the *p*-factor and neurocognitive components could be interpreted consistently.

The following chapter has been published in International Journal of Environmental Research and Public Health in the special issue titled Diagnosis and Advances in Research on Human Behavior.

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2021). One p-Factor for all? Exploring the applicability of structural models of psychopathology within subgroups of a population. *International Journal of Environmental Research and Public Health, 18*(13), 7108. https://doi.org/10.3390/ijerph18137108

Minor edits have been made to the following chapter, such as phrasing and Australian Spelling and the re-formatting of figures, to ensure consistency within the thesis. This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

Chapter 2 (Study 1): One *p*-Factor for All? Exploring the Applicability of Structural Models of Psychopathology within Subgroups of a Population

2.1 Introduction

Methodologies using confirmatory factor analytic (CFA) techniques have recently arisen as alternatives to traditional diagnostic approaches in the study of psychopathology. These approaches have risen to mitigate the issues of extensive comorbidity and diagnostic instability found when using traditional categorical systems, such as the *Diagnostic and Statistical Manual* (see Caspi et al., 2014; Lahey et al., 2021b). Factor analytical approaches typically take a dimensional approach to the presentations of psychopathology and organise hierarchical structures of specific and more general factors (Lahey et al., 2021b). In essence, these approaches examine the number (and type) of psychological symptoms that are present in a population and then fit these symptoms onto statistical models that posit different structural relations between higher order factors. The bifactor model of psychopathology, comprising of a hierarchical structure of a range of psychopathological symptoms and a smaller collection of higher-order factors, has emerged as the preferred CFA model to summarise psychopathology (Fried et al., 2021; Greene et al., 2019; Lahey et al., 2021b). Particularly noteworthy from the bifactor literature are (1) the findings that a single higher-order factor emerges from the statistical analysis of symptoms, and (2) the claim that this factor (subsequently named the *p*-factor; as in 'psychopathology factor') may refer to an individual's propensity towards psychological disease and illness (see Caspi et al., 2014; Lahey et al., 2021b). Whilst both these points relate to issues that occur when cognitive processes go awry, they are highly similar in nature to claims regarding the structure of the typical development of intelligence. For instance, both the *p*-factor and the *g*-factor (the general factor of intelligence) emerge from the analysis of population data (Caspi et al., 2014). However, whereas the *g*-factor is assumed to relate to the typical development of intellectual abilities, the *p*-factor is held to represent atypical psychological functioning. Additionally, similar to work on the structure of intelligence is the claim that levels of *p* at the individual level are normally distributed within the population (Caspi et al., 2014).

The key reason why *p* is of interest is the claim that one's *p* represents a substantive property of the system that determines one's propensity towards psychopathology (Caspi et al., 2014; Caspi & Moffitt, 2018). In intelligence, *g* has been variously argued to relate to properties

such as the capacity of the system, the speed of information processing, or the complexity of information that is represented, (e.g., Gottfredson, 1998). In the study of the *p*-factor, claims for *p* representing a substantive construct are popular. For example, it has been suggested that *p* may primarily represent functional impairment (Smith et al., 2020), neuroticism (Brandes et al., 2019), cognitive abilities (see Lahey et al., 2021b), impulsive responsivity to emotion (Carver et al., 2017), and disordered thought (Caspi & Moffitt, 2018). However, there is a range of evidence suggesting that, in a purely statistical state, the *p*-factor may not realistically represent a substantive construct or property (e.g., Watts et al., 2020b).

The *p*-factor is a function of the dataset from which it is extracted (Murray et al., 2016; Snyder & Hankin, 2017), and therefore it is inherently fluid and is a different factor depending on the characteristics of the sample and the methods of extraction (Levin-Aspenson et al., 2020). Lahey et al. (2021b) describes *p* as a "weighted-average" (p. 61) of aspects of all the symptoms assessed. Indeed, Fried et al. (2021) showed that *p* derived from bifactor models is not notably different to a simple sum of symptoms. Furthermore, even though a bifactor model is often chosen over a correlated factors model (a model with no *p*-factor and correlated specific factors) due to better fit (e.g., Caspi et al., 2014), Greene et al. (2019) showed that fit statistics unfairly favour the more accommodating bifactor model. These issues have led to multiple authors claiming that if the *p*-factor is to be a substantive construct, substantive *p* must be built "on-top" of statistical *p* (Fried et al., 2021; Greene et al., 2019; Watts et al., 2020a). That is, a theoretical construct of *p* not only must be informed by its statistical make up but also must incorporate predefined predictors and boundaries so that *p* can be falsified (Fried et al., 2021; Levin-Aspenson et al., 2020; Watts et al., 2020a). The development of a universal substantive *p* may have important implications for etiological and treatment domains of psychopathology (Fried et al., 2021; Lahey et al., 2021b). However, for substantive *p* to be useful in the treatment setting, it must be a construct applicable to every individual in the population. On the individual level, a person's *p* score is a function of the sample, and therefore would differ depending upon the attributes of the dataset in which they lie.

Previously, we have claimed that there is a need to assess the applicability and consistency of structural models in heterogeneous subgroups of a population (Haywood et al., 2021c). Bifactor models have been successfully fitted to normative and clinical populations (even though the makeup of *p* changes); however, it is unclear if structural models can fit

heterogeneous subgroups *within* a population (Haywood et al., 2021c). It is possible that the factors of psychopathology within a dataset are differentially applicable to a number of individuals in the sample. For example, the *p*-factor might not have substantial utility for a number of individuals or subgroups within a sample, even though the factor has utility for the sample as a whole. Specifically, this may mean factor scores derived from within a subgroup may be better (or worse) at representing those individuals' symptomologies when compared with factor scores derived from the total sample. However, an alternative model (e.g., a correlated factors model) may be a better fit for some subgroups within a sample or, indeed, no substantial model would adequately account for the subgroups' symptoms. Practically, this may mean that when parsing out the remaining individuals from a population for a collection of individuals or subgroups within a dataset, the *p*-factor and/or the specific factors of psychopathology might (a) not be applicable, (b) have different factor loading patterns, and/or (c) have different correlates or correlates with different utility. Any of these possibilities would suggest that the factors would have a different substantive meaning within each subgroup when compared with the sample as a whole. However, it is unknown the degree to which subgroups with clear symptomatic boundaries, but with adequate variation to use CFA techniques, could be fit to structural models of psychopathology. This has implications for the representativeness of both statistical *p* and any potential, universal, substantive *p.* Therefore, consideration must also be extended from not just what the *p*-factor might represent *between* samples (e.g., Levin-Aspenson et al., 2020)) but also to what it might represent *within* samples.

Levin-Aspenson et al. (2020) call for an agreed upon definition of the *p*-factor with corresponding expectations of what constructs should, and should not, correlate with the factor. However, if the *p*-factor has substantial differences in factor loading patterns within a subgroup of a sample, or if a meaningful number of similar individuals within the sample are not adequately represented by the *p*-factor (or the second-order factors), there is unlikely to be a universal (i.e., over the population and different groups) agreement on what the attributions of the *p*-factor (and the specific factors) are. In order to develop a universal substantive meaning of the factors of psychopathology, a within-sample exploration of the factors' applicability and loading characteristics, as well as the stability of their predictors is needed. However, the issue of poor statistical power arises from fitting CFA models to large subgroups within a sample. Data simulation methodologies not only allow adequate power by having a sample size fit to purpose

but also allow the user to control the characteristics of the data. These aspects make simulation methodologies particularly suited to questions related to structural models of psychopathology (e.g., see Greene et al., 2019)).

To inform considerations for the use of the factors of psychopathology as both a statistical and substantive construct, in this study we aimed to (a) explore the applicability, (b) loading patterns, and (c) factor associations of various popular structural models of psychopathology to subgroups of a sample using a simulated data approach. The findings of this research may be used to assess the likelihood that consistent, universal, substantive constructs of psychopathology may be developed.

2.2 Materials and Methods

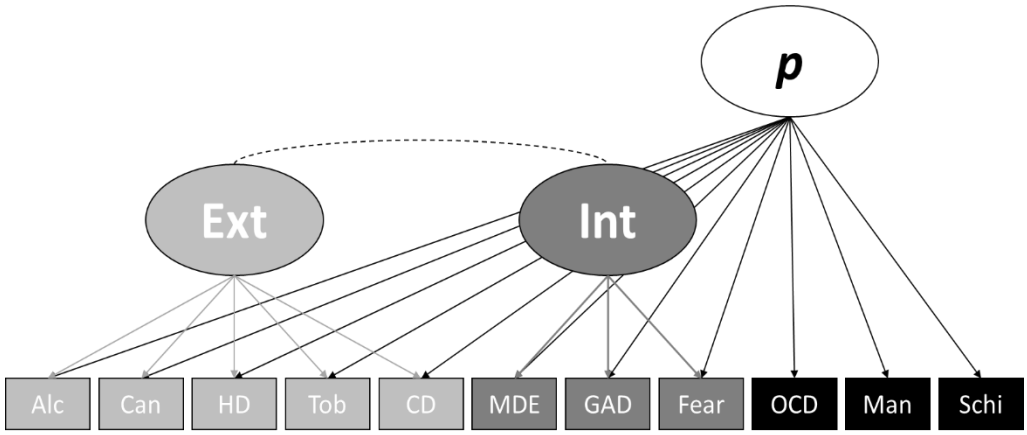
2.2.1 Data Generation

2.2.1.1. Symptoms

Caspi et al. (2014) used data from the Dunedin Multidisciplinary Health and Development Study (n = 1000) approved by the University of Otago Ethics Committee, including symptom counts using the Diagnostic Interview Schedule (Robins et al., 1995) and a range of potential correlates of psychopathology (see Caspi et al., 2014 for futher details). Similar to Greene et al. (2019), we took a top-down data simulation approach whereby the use of existing factor loadings provided by the final bifactor model of Caspi et al. (2014) (see Figure 2.1) allowed us to generate a population dataset consisting of distributions of symptom counts of the 11 disorder categories for 100,000 individual subjects. The distribution of symptom counts were positively skewed to approximately 2.0 across the population (following Greene et al., 2019). This level of positive skew is representative of distributions of symptoms typically found in the population (Curran et al., 1996). The resulting population dataset therefore, represented a summary of symptom counts from Caspi et al. (2014).

Figure 2.1.

Bifactor Model Used for the Development of the Simulated Data.



Note. Depiction of the received bifactor model from Caspi et al. (2014) used to develop the simulated data. Alc = alcohol; Can = cannabis; HD = hard drugs; Tob = tobacco; CD = conduct disorder; MDE = major depressive episode; GAD = generalized anxiety disorder; Fear = fears and phobias; OCD = obsessive-compulsive disorder; Man = mania; Schi = schizophrenia; Ext = externalizing; Int = internalizing.

2.2.1.2. Intelligence

To explore potential variation in correlates of the factors of psychopathology between samples, we produced IQ scores (analogous to Wechsler Adult Intelligence Scale-IV; WAIS-IV scores) that consisted of normally distributed composite scores (mean = 100, SD = 15) for the Verbal Comprehension (VC), Perceptual Reasoning (PR), Working Memory (WM), and Processing Speed (PS). Each IQ score was generated by matching the associations reported between each subscale score and the externalising factor, internalising factor, and the *p*-factor of Caspi et al. (2014). RStudio was used with the Lavaan package (Rosseel, 2012) to develop and analyse the simulation data.

2.2.2 Summary of Generated Population Data

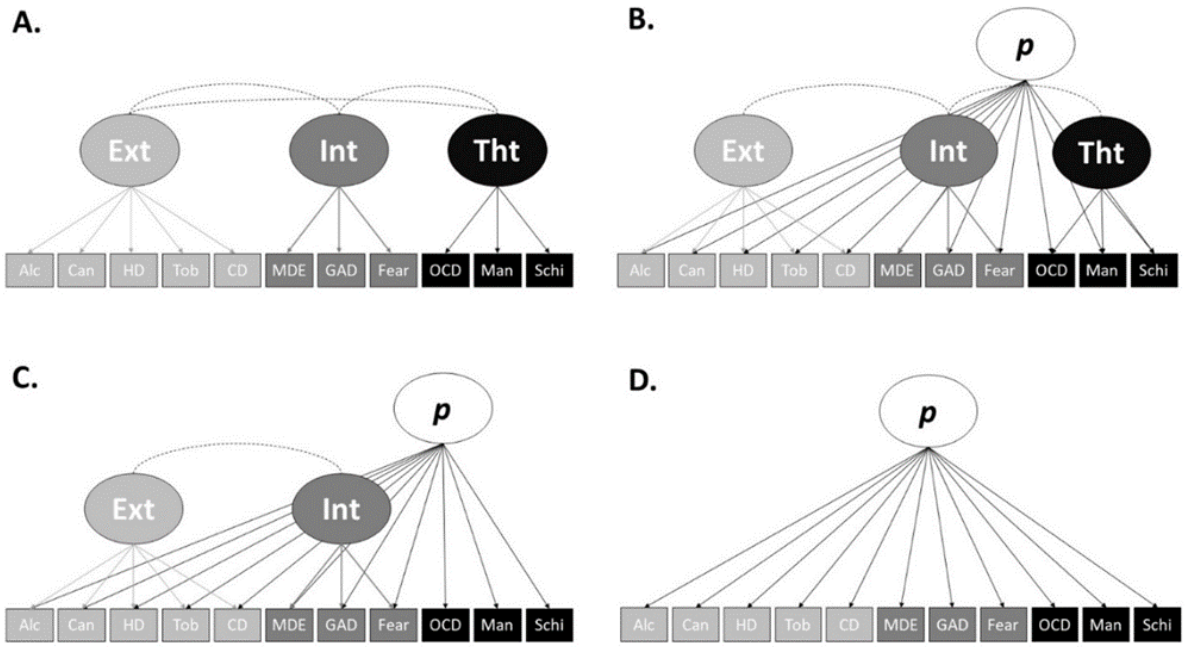
Diverging from Greene et al. (2019), we did not use Monte Carlo simulations. To better fit the purposes of this research, we produced a single dataset to represent data collected from a single sample as per empirical research and to facilitate a detailed exploration of subgroup characteristics. Our final population dataset comprised of data for 100,000 subjects that included

simulated scores on the WAIS-IV subscales, each normally distributed, and the symptoms counts of 11 disorder categories (alcohol, cannabis, hard drugs, tobacco, conduct disorder, fears and phobias, major depressive episode, generalised anxiety disorder, obsessive-compulsive disorder, mania, and schizophrenia), with an average positive skew of approximately 2.0. WAIS-IV Scores and symptom counts within the population were simulated from the revised bifactor model and resulting associations from Caspi et al. (2014).

The characteristics of the simulated data closely resembled that of Caspi et al. (2014) and had a mean skew of 1.99 (SD = .35) between the 11 observed variables. Table 2.1 below presents the revised bifactor model factor loadings of Caspi et al. (2014), and loadings of the full simulated dataset fitted to the same model are shown in Figure 2.1. Table 2.1 also presents bivariate associations between each WAIS-IV subscale and internalising factor, externalising factor, and the *p*-factor for both Caspi et al. (2014) and our simulated data (see the Analysis and Subgroup Generation section for details regarding the CFA estimation method and fit indicators used). As shown in Table 2.1, there is slight divergence between the factor loadings and associations of Caspi et al. (2014) and our simulated data. This is due to both random data generation factors, data constraints, potential differences in skew between each Caspi et al. (2014) variable and the simulated data, and the use of continuous variables and different estimators (WLSMV vs. MLR). However, the Caspi et al. (2014) revised bifactor model of psychopathology fit the simulation data well ($\chi 2(35, N = 100{,}000) = 41.82$, CFI = 1.00, TLI = 1.00, SRMR = .002, RMSEA = .001, 90% CI = [.000, .003]), and the simulated data largely retained the factor loading and relationship patterns of Caspi et al. (2014). Furthermore, mirroring Caspi et al. (2014), the correlated factors model fit our simulated data well ($\chi 2(41, N = 100{,}000) = 4484.98$, CFI = .989, TLI = .986, SRMR = .031, RMSEA = .033, 90% CI = [.032, .034]) (see Table 2.2 for loading comparisons), and like Caspi et al. (2014) the original bifactor model (Figure 2.2B) had a convergence issue due to the thought disorder factor being subsumed by the *p*-factor. Lastly, like Caspi et al. (2014), the single-factor model (Figure 2.2D) did not fit the data well ($\chi 2(44, N = 100{,}000) = 86{,}267.82$, CFI = .792, TLI = .740, SMRM = .116, RMSEA = .140, 90% CI = [.139, .141]). Therefore, we conclude that our simulated data are a good representation of the data used by Caspi et al. (2014). Following this, the factor loadings were saved back to the dataset, then, as per Caspi et al. (2014), we standardised the *p*-factor loadings to a mean of 100 and a standard deviation of 15.

Figure 2.2.

Four Structural Models from Caspi et al. (2014)



Note. Depiction of the four structural models used by Caspi et al. (2014) and tested in this research. Panel A: Correlated factors model; Panel B: Full bifactor model; Panel C: Revised bifactor model; Panel D: Single-factor model. Alc = alcohol; Can = cannabis; HD = hard drugs; Tob = tobacco; CD = conduct disorder; MDE = major depressive episode; GAD = generalized anxiety disorder; Fear = fears and phobias; OCD = obsessive-compulsive disorder; Man = mania; Schi = schizophrenia; Ext = externalizing; Int = internalizing; Tht = thought disorder.

Table 2.1

Comparison Between Caspi et al.'s (2014) Revised Bifactor Model in the Original and Simulated Data.

Dataset	Factor	Alc	Cann	HD	Tob	CD	Dep	GAD	Fears	OCD	Mania	Schiz	VC~	PR~	WM~	PS~	Ext~ Int
Caspi et al. (2014) (N = 1000)																	
																	−.471
	p	.397	.455	.452	.504	.557	.835	.812	.623	.725	.973	.819	−.129	−.129	−.183	−.176	
	Ext	.626	.811	.709	.420	.691	-	-	-	-	-	-	−.084	−.054	−.028	−.035	
	Int	-	-	-	-	-	.340	.497	.441	-	-	-	.112	.062	−.027	.019	
Simulation Data (N = 100,000)																	
																	−.387 **
	p	.294	.309	.320	.398	.394	.605	.582	.476	.709	.969	.805	−.120 **	−.123 **	−.176 **	−.167 **	
	Ext	.536	.629	.577	.368	.559	-	-	-	-	-	-	−.120 **	−.068 **	−.015 **	−.035 **	
	Int	-	-	-	-	-	.247	.388	.347	-	-	-	.110 **	.060 **	−.007 *	.032 **	

Note. Significance of the correlations was not reported by Caspi et al. for each variable. ~ = Correlation; Alc = alcohol; Cann = cannabis; HD = hard drugs; CD = conduct disorder; Dep = major depressive episode; GAD = generalized anxiety disorder; Fears = fears and phobias; OCD = obsessive-compulsive disorder; Schiz = schizophrenia; VC = Verbal Comprehension; PR = perceptual reasoning; WM = Working Memory; PS = Processing Speed; Ext = externalizing; Int = internalizing.

All factor loadings are significant at $p < .01$. * = $p < .05$. ** = $p < .01$.

Table 2.2

Comparison Between Caspi et al.'s (2014) Correlated Model in the Original and Simulated Data.

Dataset	Factor	Alc	Cann	HD	Tob	CD	Dep	GAD	Fears	OCD	Mania	Schiz	VC~	PR~	WM~	PS~	Ext~ Int	Ext~ Tht	Int~ Tht
Caspi et al. (2014) (N = 1000)																			
																	.328	.577	.849
	Ext	.733	.885	.839	.668	.909	-	-	-	-	-	-	−.139	−.166	−.126	−.126			
	Int	-	-	-	-	-	.972	.934	.704	-	-	-	−.049	−.077	−.154	−.134			
	Tht	-	-	-	-	-	-	-	-	.726	.982	.826	−.115	−.116	−.171	−.166			
Simulation Data (N = 100,000)																			
																	.296 **	.531 **	.869 **
	Ext	.604	.678	.652	.545	.694	-	-	-	-	-	-	−.166 **	−.125 **	−.114 **	−.125 **			
	Int	-	-	-	-	-	.682	.677	.563	-	-	-	−.075 **	−.094 **	−.166 **	−.148 **			
	Tht	-	-	-	-	-				.709	.968	.805	−.120 **	−.123 **	−.176 **	−.167 **			

Note. Significance of the associations was not reported by Caspi et al. for each variable. ~ = correlation; Alc = alcohol; Cann = cannabis; HD = hard drugs; CD = conduct disorder; Dep = major depressive episode; GAD = generalized anxiety disorder; Fears = fears and phobias; OCD = obsessive-compulsive disorder; Schiz = schizophrenia; VC = Verbal Comprehension; WM = Working Memory; PS = Processing Speed; Ext = externalizing; Int = internalizing; Tht = thought disorder.

All factor loadings are significant at $p < .01$. * = $p < .05$. ** = $p < .01$.

2.2.3 Analysis and Subgroup Generation

2.2.3.1. CFA

As per Greene et al. (2019), we used maximum likelihood estimation (MLR) with robust standard errors and Pearson's correlations for our CFAs to assess the fit of the models of psychopathology. MLR was chosen for its corrections to chi-square statistics and standard errors when used with skewed data. Sample and estimated variance and covariance differences were assessed via the standardized root-mean-square residual (SRMR; Hu & Bentler, 1995), degree of model fit was determined by root-mean-square error of approximation (RMSEA; Steiger, 1990), and the comparative fit index (CFI; Bentler, 1990) and Tucker–Lewis index (TLI) was used to assess fit improvement in relation to a model that is saturated. RMSEA values of <.05 (Bollen & Curran, 2006) and CFI and TLI values of >.95, with all significant loadings, no negative loadings, no non-positive-definite identification issues, and no negative variance were used as thresholds for a good fit and model utility (Hu & Bentler, 1995).

2.2.3.2. Models Tested

Caspi et al. (2014) attempted to fit their full sample's data to four models of psychopathology. A correlated factors model (Figure 2.2A), a bifactor model (with correlated higher-order factors; Figure 2.2B), a revised bifactor model (with correlated higher-order factors; Figure 2.2C) and a single-factor model (Figure 2.2D). Caspi et al. (Caspi et al., 2014) found the correlated factors model ($\chi2(1018, N = 1000) = 1737.16$, CFI = .962, TLI = .958, RMSEA = .027, 90% CI = [.024, .029]), and the revised bifactor model ($\chi2(1012, N = 1000) = 1652.59$, CFI = .966, TLI = .963, RMSEA = .025, 90% CI = [.023, .027]) fit the data well. The original bifactor model did not successfully converge (the thought disorder factor was subsumed by the *p*-factor), and the single-factor model did not fit the data well ($\chi2(1021, N = 1000) = 3404.57$, CFI = .875, TLI = .862, RMSEA = .048, 90% CI = [.047, .050]). To assess the applicability of each model and the factors of psychopathology *within* a sample, we tested the fit of each of these models on each of our devised subgroups.

2.2.3.3. Subgroup Determination

To ensure that we captured the heterogeneity and the variation of comorbidity of symptoms within a population, we created overlapping subgroups from our simulated data (i.e., a single case or subject could appear in more than one subsample) determined by thirds of the total sample's disorder symptom and factor scores. Specifically, each case was characterised as within the lower, middle, or upper third of scores for (1) at least one disorder reflecting a (a) internalising, (b) externalising, or (c) thought disorder, and further by (2) their

scores on the (a) externalising and (b) internalising factors and by (3) their *p*-factor score. These permutations resulted in 63 subgroups. Twenty-seven of the subgroups were externalising variants, created using all permutations of upper, middle, and lower scores for at least one externalising disorder, the externalising factor, and *p*. A further 27 were internalising variants, resulting from all permutations of upper, middle, and lower scores for at least one internalising disorder, the internalising factor, and *p*. Lastly, 9 further subgroups were thought disorder variants, resulting from all permutations of the upper, middle, and lower scores for at least one thought disorder and *p*. The thought disorder factor did not appear in the Caspi et al. (2014) revised bifactor model; thus, thought disorder factor permutations were not possible. In this paper, we refer to subgroups in a A(x)-B(x)-C format, where A represents either the upper (1), middle (2), or lower (3) third of the symptomology of at least one externalising factor, internalising factor, or thought disorder, represented by "x" ("ext", "int", or "tht"). B represents either the upper (1), middle (2), or lower (3) third of externalising or internalising factor scores, again with the factor represented by "x". Lastly, C represents either the upper (1), middle (2), or lower (3) third of *p*-factor scores.

Each of the 63 subgroups were fit to each of the four models of psychopathology to explore the applicability of the sample level structure to the subgroups and the precise factor loading characteristics of each subgroup. Furthermore, bivariate associations between the four WAIS-IV subscales and internalising, externalising, thought disorder, and the *p*-factor in each subgroup were examined.

2.3 Results

Of the 63 subgroups, only 8 fit at least one of the four structural models, with all significant loadings, no negative loadings, no negative variance, and no non-positive-definite identification issues. All eight subgroups fit the correlated factors model (Figure 2.2A), none fit the original bifactor model (Figure 2.2B), four subgroups fit the revised bifactor model (Figure 2.2C), and one subgroup fit the single-factor model (Figure 2.2D). Of the eight subgroups, four fit only one model, and four fit two of the models of psychopathology. Five externalising and three thought disorder subgroup variants fit the correlated factors model (A), while one externalising and three thought disorder subgroup variants fit the revised bifactor model (C). Lastly, only a single subgroup, an externalising variant, fit the single-factor model. None of the four CFA models of psychopathology fit any of the internalising subgroup variants. The fit indices for the subgroups that fit at least one model are presented in Table 2.3, and their loadings and associations are presented in Table 2.4.

Table 2.3

Fit Statistics for Subgroups That Fit At Least One Model Well

Model	Subgroup	Chi-Square	*df*	CFI	TLI	SRMR	RMSEA [90% CI]
Correlated Factors Model (A)							
	1(ext)/1(ext)/3 (*n* = 10,684)	133.73	41	.992	.990	.012	.015 [.012, .017]
	3(ext)/3(ext)/1 (*n* = 9,163)	93.22	41	.990	.987	.012	.012 [.009, .015]
	3(ext)/3(ext)/2 (*n* = 11,738)	36.81	41	1.00	1.00	.007	.000 [.000, .005]
	3(ext)/3(ext)/3 (*n* = 12,305)	201.57	41	.990	.987	.015	.018 [.015, .020]
	2(ext)/1(ext)/3 (*n* = 11,306)	122.38	41	.996	.994	.011	.013 [.011, .016]
	1(tht)/X/1(tht) (*n* = 33,218)	525.24	41	.991	.987	.017	.019 [.017, .020]
	2(tht)/X/2(tht) (*n* = 32,610)	52.22	41	1.00	.999	.005	.003 [.000, .005]
	3(tht)/X/3(tht) (*n* = 33,255)	791.28	41	.989	.985	.022	.023 [.022, .025]
Revised Bifactor Model (C)							
	3(ext)/3(ext)/3 (*n* = 12,305)	35.94	35	1.00	1.00	.005	.001 [.000, .007]
	1(tht)/X/1(tht) (*n* = 33,219)	6.07	35	1.00	.999	.004	.005 [.003, .007]
	2(tht)/X/2(tht) (*n* = 32,610)	24.32	35	1.00	1.00	.003	.000 [.000, .002]
	3(tht)/X/3(tht) (*n* = 33,255)	23.91	35	1.00	1.00	.003	.000 [.000, .001]
Single-Factor Model (D)							
	2(ext)/1(ext)/3 (*n* = 11,306)	545.76	44	.974	.967	.023	.032 [.019, .034]

Note. *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker–Lewis index; SRMR = standardized root-mean-square residual; RMSEA = root-mean-square error of approximation; 90% CI = 90% confidence interval; Ext = externalizing; Int = internalizing; Tht = thought disorder

Table 2.4

Loadings and Associations for Subgroups That Fit At Least One Model Well

Model	Subgroup	Factor	Alc	Cann	HD	Tob	CD	Dep	GAD	Fears	OCD	Mania	Schiz	Ext~ Int	Ext~ Tht	Int~ Tht
Correlated Factors Model (A)																
	1(ext)/1(ext)/3 (*n* = 10,684)													.641 **	.955 **	.674 **
		Ext	.206	.255	.234	.259	.307	-	-	-	-	-	-			
		Int	-	-	-	-	-	.506	.495	.377	-	-	-			
		Tht	-	-	-	-	-	-	-	-	.462	.916	.584			
	3(ext)/3(ext)/1 (*n* = 9163)													.257 **	.655 **	.590 **
		Ext	.156	.161	.182	.209	.222	-	-	-	-	-	-			
		Int	-	-	-	-	-	.438	.467	.426	-	-	-			
		Tht	-	-	-	-	-	-	-	-	.359	.846	.481			
	3(ext)/3(ext)/2 (*n* = 11,738)													−.116 **	.258 **	.253 **
		Ext	.215	.341	.272	.132	.221	-	-	-	-	-	-			
		Int	-	-	-	-	-	.312	.418	.412	-	-	-			
		Tht	-	-	-	-	-	-	-	-	.115	.512	.196			
	3(ext)/3(ext)/3 (*n* = 12,305)													.455 **	.733 **	.721 **
		Ext	.290	.374	.351	.307	.415	-	-	-	-	-	-			
		Int	-	-	-	-	-	.527	.535	.401	-	-	-			
		Tht	-	-	-	-	-	-	-	-	.482	.916	.602			

Model	Subgroup	Factor	Alc	Cann	HD	Tob	CD	Dep	GAD	Fears	OCD	Mania	Schiz	Ext~ Int	Ext~ Tht	Int~ Tht
Revised Bifactor Model (C)	2(ext)/1(ext)/3 (*n* = 11,306)													.800 **	.992 **	.732 **
		Ext	.316	.373	.349	.327	.406	-	-	-	-	-	-			
		Int	-	-	-	-	-	.528	.510	386.	-	-	-			
		Tht	-	-	-	-	-	-	-	-	.522	.935	.628			
	1(tht)/X/1(tht) (*n* = 33,218)													−.135 **	.265 **	.620 **
		Ext	.564	.652	.618	.429	.621	-	-	-	-	-	-			
		Int	-	-	-	-	-	.465	.508	.437	-	-	-			
		Tht	-	-	-	-	-	-	-	-	.403	.879	.525			
	2(tht)/X/2(tht) (*n* = 32,610)													−.381 **	.062 **	.181 **
		Ext	.553	.663	.605	.405	.603	-	-	-	-	-	-			
		Int	-	-	-	-	-	.336	.466	.395	-	-	-			
		Tht	-	-	-	-	-	-	-	-	.109	.568	.183			
	3(tht)/X/3(tht) (*n* = 33,255)													−.050 **	.317 **	.710 **
		Ext	.577	.666	.623	.448	.642	-	-	-	-	-	-			
		Int						.536	.556	.430	-	-	-			
		Tht	-	-	-	-	-	-	-	-	.496	.921	.616			
	3(ext)/3(ext)/3 (*n* = 12,305)													−.148 **		
		p	.199	.261	.249	.258	.309	.399	.376	.275	.483	.916	.602			
		Ext	.263	.338	.268	.079	.243	-	-	-	-	-	-			
		Int	-	-	-	-	-	.248	.476	.322	-	-	-			
	1(tht)/X/1(tht) (*n* = 33,219)															

Model	Subgroup	Factor	Alc	Cann	HD	Tob	CD	Dep	GAD	Fears	OCD	Mania	Schiz	Ext~ Int	Ext~ Tht	Int~ Tht
														−.391 **		
		p	.138	.137	.152	.189	.184	.320	.302	.249	.402	.881	.523			
		Ext	.547	.645	.600	.387	.590	-	-	-	-	-	-			
		Int	-	-	-	-	-	.289	.436	.391	-	-	-			
	2(tht)/X/2(tht) (*n* = 32,610)															
														−.399 **		
		p	.036	.034	.039	.047	.030	.092	.080	.051	.109	.573	.181			
		Ext	.552	.662	.604	.403	.602	-	-	-	-	-	-			
		Int	-	-	-	-	-	.323	.459	.395	-	-	-			
	3(tht)/X/3(tht) (*n* = 33,255)															
														−.401 **		
		p	.156	.175	.183	.230	.230	.410	.383	.280	.496	.922	.615			
		Ext	.560	.651	.597	.387	.594	-	-	-	-	-	-			
		Int	-	-	-	-	-	.277	.449	.377	-	-	-			
Single Factor Model (D)																
	2(ext)/1(ext)/3 (*n* = 11,306)															
		p	.318	.377	.352	.330	.409	.421	.389	.287	.527	.915	.635			

Note. ~ = correlation; Alc = alcohol; Cann = cannabis; HD = hard drugs; CD = conduct disorder; Dep = major depressive episode; GAD = generalized anxiety disorder; Fears = fears and phobias; OCD = obsessive-compulsive disorder; Schiz = schizophrenia; Ext = externalizing; Int = internalizing; Tht = thought disorder.

All factor loadings significant at $p < .01$. * = $p < .05$. ** = $p < .01$.

2.3.1 Correlated Factors Model (A)

Regarding the correlated factors model (A), symptom loadings on externalising varied substantially between the total sample and the subgroups. Loadings on externalising in the total sample ranged from .545 to .694, while the loadings on externalising among the subgroups ranged from .156 to .666. As expected, externalising variant subgroups had the largest externalising loading differentiation (ranging from .105 to .406), while the thought disorder variants (loadings ranging from .405 to .666) had more similar factor loadings to the total sample. Internalising factor loadings also varied between the total sample and the subgroups. Loadings on internalising for the total sample were typically higher than for the subgroups. The loadings on internalising for the total sample ranged from .563 to .682, while the loadings on internalising for the subgroups ranged from .312 to .556. Loadings of externalising (ranging from .312 to .535) and thought disorder (ranging from .336 to .556) subgroup variants on the internalising factor were similar. Lastly, factor loadings on thought disorder between the full sample and the subgroups varied considerably. Loadings on thought disorder in the total sample ranged from .709 to .968, while the loadings on thought disorder among the subgroups ranged from .109 to .935. Thought disorder variant subgroups (loadings ranging from .109 to .921) and the externalising variants (loadings ranging from .115 to .935) had similar variation in factor loadings on thought disorder. Subgroup variants of higher thought disorder symptoms did not typically load higher than the other subgroups on the thought disorder factor, instead the combination of total sample *p*-factor level and thought disorder symptom level seemed to drive the factors loadings. For example, subgroup 1(tht)-X−1 had thought disorder loadings ranging from .403 to .879, while subgroup 2(tht)-X−2 had thought disorder loadings ranging from .109 to .568. Similarly, externalising variant 1(ext)−1(ext)−3 had higher factor loadings on externalising (ranging from .206 to .307) than subgroup variant 3(ext)−3(ext)−1 (ranging from .156 to .222).

Out of the eight subgroups that fitted the correlated factors model well, four had a negative association between internalising and externalising. This contrasts with the positive association of .296 between the externalising and internalising factors of the full simulation sample. Externalising–internalising association among the four subgroups with a positive relationship ranged from .257 to .800 and included only externalising variant subgroups. The associations among the four subgroups with negative relationships ranged from −.381 to −.050

and included one externalising and three thought disorder variant subgroups and had no consistencies in thirds rankings. This suggests that even though the models with negative relationships between externalising and internalising fit the correlated factors model well, the characteristics of these models are fundamentally different from that of the total sample.

2.3.2 Revised Bifactor Model (C)

Regarding the revised bifactor model (C), one externalising subgroup variant and three thought disorder variants fit the model well, and, like the correlated factors model, symptom loadings on externalising for the revised bifactor model (C) varied substantially between the total sample and the subgroups. Loadings on externalising in the total sample ranged from .368 to .629, while the loadings on externalising among the subgroups ranged from .079 to .662. Loadings on internalising for the subgroups were typically higher than the full sample, ranging from .248 to .459, while the total sample ranged from .247 to .347; however, loadings were closer to the full sample loadings when compared to externalising. This may be expected as no internalising variant subgroups fit the revised bifactor model. The only externalising subgroup variant that fit the revised bifactor model was the subgroup with the lowest externalising loadings (ranging from .079 to .338), even though this subgroup only contained participants in the upper third for at least one externalising disorder, as well as the upper third for the externalising factor score and *p*-factor score (3(ext)/3(ext)/3). The other subgroups that fit the bifactor model well were all thought disorder variants, with matching thirds between their thought disorder symptoms and their *p*-factor scores. This supports the suggestion that in the Caspi et al. (Caspi et al., 2014) revised bifactor model, the *p*-factor largely represents thought disorder (Caspi & Moffitt, 2018).

Symptom loadings on the *p*-factor differed substantially between the subgroups and the total sample. Externalising disorder symptom loadings on *p* ranged from .030 to .309 in the subgroups and .294 to .398 in the total sample. Loadings of externalising disorder symptoms on *p* were generally lower than in the total sample, with the 2(tht)-X−2 subgroup providing the lowest externalising disorder loadings on *p* ranging from .030 to .047. Internalising disorder symptom loadings on *p* ranged from .051 to .476 in the subgroups and .476 to .605 in the total sample. Like externalising, loadings of internalising disorder symptoms on *p* were generally lower than in the total sample, and once again the 2(tht)-X−2 subgroup provided the lowest

internalising disorder loadings on *p*, ranging from .051 to .092. Thought disorder symptom loadings on *p* once again differed substantially between the subgroups, ranging from .109 to .922, and the total sample, ranging from .709 to .805. Again, the 2(tht)-X−2 subgroup provided the lowest thought disorder loadings, ranging from .109 to .573. The loading characteristics of 2(tht)-X−2 provide an example of a subgroup where, although the bifactor model fit well, the *p*-factor is not representative of the participants' propensity towards psychopathology, and rather the lower-order factors (i.e., externalising and internalising) are better representations of the subjects' collective symptoms.

The association between externalising and internalising for all four subgroups that fit the bifactor model well were negative and ranged from −.148 to −.401, compared with the full sample externalising–internalising association of −.387. Only the externalising subgroup variant had the lowest externalising–internalising relationship (−.148) and differed the greatest when compared with the total sample. The thought disorder variants had an externalising–internalising relationship ranging from −.391 to −.401, closely resembling the total sample and providing more evidence that thought disorder drives the *p*-factor in this revised bifactor model.

2.3.3 Single-Factor Model (D)

Only a single subgroup, 2(ext)/1(ext)/3, fit the single-factor model well, indicating that *p* alone is generally unable to successfully account for the symptoms of psychopathology at the total sample and the subgroup level. Within subgroup 2(ext)/1(ext)/3, externalising and internalising disorders loaded on *p* at a similar level. Externalising disorders loadings on *p* ranged from .318 to .409, and internalising disorders loadings on *p* ranged from .287 to .421; however, similar to the correlated factors and bifactor models, the thought disorders loaded on *p* to a greater extent. The three thought disorders loadings on *p* ranged from .527 to .915, suggesting that whether second order factors are included (e.g., the bifactor model) or not (e.g., the single-factor model), within the total sample and the subgroups, *p* in this dataset is largely defined by thought disorder symptomology. Given that the single-factor model was not a good fit for the total sample, any comparison of factor loadings between the total sample and the subgroup would not be appropriate.

2.3.4 Factor Associations

The associations between the factor scores derived from within each subgroup and the factor scores for the subgroup derived from the total sample are presented in Table 2.5.

Regarding the correlated factors model (A), the data in Table 2.5 show that while the subgroup-derived and total-sample-derived factor scores for internalising and thought disorder are very similar (internalising ranging from .921 to .999; thought disorder ranging from .983 to 1.00), subgroup- and total-sample-derived externalising scores vary (ranging from .663 to .999). Externalising factor score associations tended to be lower in the externalising subgroup variants (ranging from .663 to .973) than in the thought disorder subgroup variants (ranging from .997 to .999). This may suggest that externalising factor scores are particularly susceptible to individual and subgroup differences, while internalising and thought disorder scores largely follow the same pattern of scores within the total sample and the subgroups in a correlated factors model. Regarding the revised bifactor model (C), associations between the total-sample- and subgroup-derived factor scores were generally consistent and high (ranging from .941 to 1.00). This suggests the pattern of externalising, internalising, and *p*-factor scores were generally consistent in the total sample and the subgroups. This mirrors the factor loading characteristics that showed, generally, that subgroup *p*-factor loadings followed a similar pattern to the total sample, even when the magnitude of the loadings differed. Similarly, even though the single-factor model did not fit the total sample well, the subgroup-derived *p*-factor for the single subgroup that fit the model well had a very strong relationship with the total-sample-derived *p*-factor (.996).

Table 2.5

Total Sample and Within Subgroup p-factor Associations

Model	Subgroup		Total Sample EXT	Total Sample INT	Total Sample THT	Total Sample *p*
Correlated Factors Model (A)						
	1(ext)/1(ext)/3 (*n* = 10,684)					
		Subgroup EXT	.759 **	.849 **	.992 **	-
		Subgroup INT	.495 **	.996 **	.869 **	-
		Subgroup THT	.723 **	.858 **	.997 **	-
	3(ext)/3(ext)/1 (*n* = 9163)					
		Subgroup EXT	.663 **	.538 **	.924 **	-
		Subgroup INT	.149 **	.976 **	.789 **	-
		Subgroup THT	.348 **	.778 **	.999 **	-
	3(ext)/3(ext)/2 (*n* = 11,738)					
		Subgroup EXT	.973 **	−.220 **	.364 **	-
		Subgroup INT	−.136 **	.921 **	.406 **	-
		Subgroup THT	.279 **	.592 **	.996 **	-

Model	Subgroup		Total Sample EXT	Total Sample INT	Total Sample THT	Total Sample *p*
	3(ext)/3(ext)/3 (*n* = 12,305)					
		Subgroup EXT	.887 **	.687 **	.890 **	-
		Subgroup INT	.425 **	.990 **	.887 **	-
		Subgroup THT	.605 **	.888 **	.999 **	-
	2(ext)/1(ext)/3 (*n* = 11,306)					
		Subgroup EXT	.825 **	.908 **	.996 **	-
		Subgroup INT	.691 **	.990 **	.902 **	-
		Subgroup THT	.831 **	.883 **	.997 **	-
	1(tht)/X/1(tht) (*n* = 33,218)					
		Subgroup EXT	.999 **	−.152 **	.343 **	-
		Subgroup INT	−.126 **	.999 **	.771 **	-
		Subgroup THT	.312 **	.764 **	.999 **	-
	2(tht)/X/2 (*n* = 32,610)					
		Subgroup EXT	.997 **	−.454 **	.143 **	-
		Subgroup INT	−.535 **	.949 **	.297 **	-
		Subgroup THT	.136 **	.504 **	.983 **	-
	3(tht)/X/3 (*n* = 33,255)					
		Subgroup EXT	.998 **	.050 **	.368 **	-
		Subgroup INT	.032 **	.998 **	.843 **	-
		Subgroup THT	.413 **	.873 **	1.00 **	-
Revised Bifactor Model (C)						
	3(ext)/3(ext)/3 (*n* = 12,305)					
		Subgroup EXT	.986 **	−.501 **	-	.029 **
		Subgroup INT	−.239 **	.941 **	-	.072 **
		Subgroup *p*	.163 **	−.139 **	-	.999 **
	1(tht)/X/1(tht) (*n* = 33,219)					
		Subgroup EXT	.999 **	−.652 **	-	.066 **
		Subgroup INT	−.592 **	.997 **	-	.033 **
		Subgroup *p*	.067 **	.033 **	-	.998 **
	2(tht)/X/2 (*n* = 32,610)					
		Subgroup EXT	.998 **	−.622 **	-	.097 **
		Subgroup INT	−.606 **	.992 **	-	.136 **
		Subgroup *p*	−.020 **	−.032 **	-	.983 **
	3(tht)/X/3 (*n* = 33,255)					
		Subgroup EXT	.997 **	−.587 **	-	.027 **
		Subgroup INT	−.620 **	.998 **	-	.052 **
		Subgroup p	−.052 **	.003	-	1.00 **
Single-Factor Model (D)						
	2(ext)/1(ext)/3 (*n* = 11,306)					
		Subgroup *p*	-	-	-	.996 **

Note. EXT = externalizing; INT = internalizing; THT = thought disorder.

* = $p < .05$. ** = $p < .01$.

2.3.5 The Factors of Psychopathology and Intelligence

Next, we compared the total-sample-derived factor scores and subgroup-derived factor scores' associations with intelligence for each successfully fitted subgroup. The results are presented in Table 2.6. Regarding the correlated factors model, associations between the total sample factor scores and the WAIS subscales varied greatly from the subgroup *p*-factor scores derived from both the (a) total sample and the (b) subgroups. For example, the associations between the verbal comprehension (VC) subscale and the subgroups' externalising score derived from the total sample (ranging from −.076 to −.176) and subgroup (ranging from .055 to −.177) varied greatly. Furthermore, only the higher end of that relationship range was representative of the total sample's externalising relationship with the VC subscale. This may suggest that, within subgroups, factors of psychopathology might be differentially predicted by a range of constructs and thereby have a different substantial meaning for each subgroup. For the correlated factors model, the subgroups' associations between the WAIS subscales and the total-sample-derived factor scores and subgroup-derived factor scores were generally quite consistent. For example, the subgroup-derived thought disorder factor scores and the total-sample-derived thought disorder factor scores' associations with the WAIS subscales never differed by more than .004. However, in other instances associations differed between the total-scale-derived and subgroup-derived factors and the WAIS subscales. For example, the relationship between working memory (WM) and the total sample (−.085) and subgroup-derived externalising score (−.135) for subgroup 1(ext)/1(ext)/3 differed by .05.

Regarding the revised bifactor model (C), associations between the total sample factor scores and the WAIS subscales also varied from the subgroup *p*-factor scores derived from both the (a) total sample and the (b) subgroups. For example, the associations between the processing speed (PS) subscale and the subgroups' *p*-factor score derived from the total sample (ranging from −.030 to −.170) and subgroup (ranging from .029 to −.171) varied greatly compared with the relationship between the total sample's *p*-factor score and the PS subscale (−.167). However, the subgroups' associations between the WAIS subscales and the total-sample-derived and subgroup-derived *p*-factor scores were very similar, never differing by more than .003. This suggests that while the characteristics of *p* might differ for each subgroup, as indicated by

heterogeneous subgroup associations to the WAIS subscales, *p* derived from within a subgroup and the total sample may reflect the same construct.

Lastly, even though the single-factor model (D) was not a good fit for the total sample and was only a good fit for a single subgroup, associations between the WAIS subscales and the total sample and subgroup-derived *p* scores for the subgroup were very similar. The largest WAIS subscale relationship difference between the total-sample-derived *p* and the subgroup-derived *p* was .005, once again suggesting that within-group-derived *p* is the same construct as total sample *p*.

Table 2.6

Psychopathology Factors and WAIS Subscale Associations

Model	Subgroup		VC	PR	WM	PS
Correlated Factors Model (A)						
	1(ext)/1(ext)/3 (n = 10,684)					
		Total sample EXT	−.083 **	−.080 **	−.085 **	−.109 **
		Total sample INT	−.021 **	−.049 **	−.129 **	−.086 **
		Total sample THT	−.073 **	−.087 **	−.137 **	−.118 **
		Subgroup EXT	−.078 **	−.090 **	−.135 **	−.122 **
		Subgroup INT	−.020 **	−.048 **	−.127 **	−.087 **
		Subgroup THT	−.077 **	−.089 **	−.136 **	−121 **
	3(ext)/3(ext)/1 (n = 9163)					
		Total sample EXT	−.076 **	−.043 **	−.014 **	−.046 **
		Total sample INT	.016	−.010	−.010	−.020
		Total sample THT	−.032 **	−.032 **	−.027 *	−.035 **
		Subgroup EXT	−.055 **	−.045 **	−.028 **	−.043 **
		Subgroup INT	.005	−.016	−.010	−.028 **
		Subgroup THT	−.031 **	−.032 **	−.027 *	−.034 **
	3(ext)/3(ext)/2 (n = 11,738)					
		Total sample EXT	−.097 **	−.048 **	−.032 **	−.028
		Total sample INT	.028 **	.021 *	−.026 **	−.012
		Total sample THT	−.047 **	−.014	−.041 **	−.043 **
		Subgroup EXT	−.101 **	−.048 **	−.036 **	−.033 *
		Subgroup INT	.031 **	.020 *	−.025 **	−.004
		Subgroup THT	−.050 **	−.015	−.041 **	−.044 **
	3(ext)/3(ext)/3 (n = 12,305)					
		Total sample EXT	−.162 **	−.117 **	−.067 **	−.125 **
		Total sample INT	−.055 **	−.071 **	−.125 **	−.139 **
		Total sample THT	−.118 **	−.110 **	−.135 **	−.170 **
		Subgroup EXT	−.157 **	−.127 **	−.114 **	−.167 **
		Subgroup INT	−.062 **	−.075 **	−.122 **	−.139 **
		Subgroup THT	−.121 **	−.111 **	−.135 **	−.171 **
	2(ext)/1(ext)/3 (n = 11,306)					

Model	Subgroup		VC	PR	WM	PS
		Total sample EXT	−.099 **	−.101 **	−.116 **	−.133 **
		Total sample INT	−.044 **	−.064 **	−.149 **	−.116 **
		Total sample THT	−.091 **	−.101 **	−.156 **	−.142 **
		Subgroup EXT	−.090 **	−.100 **	−.155 **	−.143 **
		Subgroup INT	−.047 **	−.067 **	−.148 **	−.120 **
		Subgroup THT	−.095 **	−.104 **	−.154 **	−.145 **
	1(tht)/X/1(tht) (n = 33,218)					
		Total sample EXT	−.098 **	−.065 **	−.022 **	−.029 **
		Total sample INT	.026 **	.002	−.027 **	−.020
		Total sample THT	−.039 **	−.040 **	−.040 **	−.043 **
		Subgroup EXT	−.099 **	−.065 **	−.023 **	−.039 **
		Subgroup INT	.025 **	.001	−.027 **	−.022 **
		Subgroup THT	−.038 **	−.039 **	−.040 **	−.042 **
	2(tht)/X/2 (n = 32,610)					
		Total sample EXT	−.106 **	−.053 **	−.025 **	−.027 **
		Total sample INT	.069 **	.036 **	−.016 **	−.002
		Total sample THT	−.027 **	−.009	−.039 **	−.031 **
		Subgroup EXT	−.106 **	−.054 **	−.024 **	−.026 **
		Subgroup INT	.089 **	.045 **	−.005	.008
		Subgroup THT	−.026 **	−.009	−.037 **	−.029 **
	3(tht)/X/3 (n = 33,255)					
		Total sample EXT	−.179 **	−.120 **	−.069 **	−.104 **
		Total sample INT	−.017 **	−.049 **	−.138 **	−.110 **
		Total sample THT	−.102 **	−.103 **	−.153 **	−.150 **
		Subgroup EXT	−.177 **	−.117 **	−.062 **	−.097 **
		Subgroup INT	−.004	−.040 **	−.134 **	−.103 **
		Subgroup THT	−.102 **	−.104 **	−.153 **	−.150 **
Revised Bifactor Model (C)						
	3(ext)/3(ext)/3 (n = 12,305)					
		Total sample EXT	−.130 **	−.082 **	−.004	−.053 **
		Total sample INT	.122 **	.077 **	.008	.046 **
		Total sample p	−.117 **	−.109 **	−.135 **	−.170 **
		Subgroup EXT	−.121 **	−.075 **	.011	−.034 **
		Subgroup INT	.076 **	.043 **	−.017	.003
		Subgroup p	−.120 **	−.111 **	−.135 **	−.171 **
	1(tht)/X/1(tht) (n = 33,219)					
		Total sample EXT	−.094 **	−.057 **	−.014 **	−.030 **
		Total sample INT	.086 **	.049 **	.005	.016 **
		Total sample p	−.039 **	−.041 **	−.040 **	−.043 **
		Subgroup EXT	−.093 **	−.057 **	−.012 *	−.029 **
		Subgroup INT	.082 **	.046 **	.003	.013 **
		Subgroup p	−.038 **	−.040 **	−.040 **	−.040 **
	2(tht)/X/2 (n = 32,610)					
		Total sample EXT	−.103 **	−.053 **	−.020 **	−.023 **
		Total sample INT	.101 **	.050 **	.006	.017 **
		Total sample p	−.027 **	−.009	−.039 **	−.030 **

Model	Subgroup		VC	PR	WM	PS
		Subgroup EXT	−.105 **	−.054 **	−.022 **	−.025 **
		Subgroup INT	.096 **	.048 **	.001	.013 *
		Subgroup *p*	−.027 **	−.009	−.037 **	−.029 **
	3(tht)/X/3 (n = 33,255)					
		Total sample EXT	−.145 **	−.080 **	−.002	−.037 **
		Total sample INT	.148 **	.087 **	−.011	.041 **
		Total sample *p*	−.102 **	−.103 **	−.153 **	−.150 **
		Subgroup EXT	−.153 **	−.088 **	−.011 *	−.049 **
		Subgroup INT	.145 **	.083 **	−.018 **	.035 **
		Subgroup *p*	−.102 **	−.104 **	−.153 **	−.150 **
Single-Factor Model (D)						
	2(ext)/1(ext)/3 (n = 11,306)					
		Total sample *p*	−.087 **	−.098 **	−.156 **	−.141 **
		Subgroup *p*	−.092 **	−.102 **	−.155 **	−.144 **

Note. VC = verbal comprehension; PR = perceptual reasoning; WM = working memory; PS = processing speed; EXT = externalizing; INT = internalizing; THT = thought disorder.
* = $p < .05$. ** = $p < .01$.

2.4. Discussion

The main objective of this study was to explore the extent to which popular structural models of psychopathology could be fit to subgroups of individuals within a sample and to explore the similarities and differences between characteristics of the factors within the subgroups. We generated a large sample of symptom and intelligence data that simulated that of empirical work published by Caspi et al. (2014) and explored the (a) applicability, (b) loading patterns, and (c) factor associations of four models of psychopathology on a subgroup level.

In analysing the fit of models to each of the 63 subgroups, we were able to determine the extent to which different structural models are successful in capturing the utility of the *p*-factor and the specific factors of psychopathology in representing symptoms. Of the 63 subgroups, each of which was fitted to four different models of psychopathology, only eight were found to fit one or more model well. Put differently, only 3.17% of the models tested showed a reliable fit. At first pass, these results suggest that when exploring the nature of symptoms in subgroups of a population, traditional structural models of psychopathology may be of low utility. This suggestion may seem to contrast with previous research. For example, structural models have been shown to fit well in specific circumstances (i.e., populations with a specific diagnosis), suggesting structural models are robust to mild symptom range limitation (Shevlin et al., 2016;

Xie et al., 2012). However, we explored different combinations of symptoms and factor loadings to better capture deviation amongst the general population.

Of the eight subgroups that were found to fit one or more models, all eight fit the correlated factors model, none fit the full bifactor model, only four fit the revised bifactor model, and a single subgroup fit the one-factor model. This may be unexpected, given that bifactor models are more accommodating and fit indices are biased towards bifactor models over correlated factor models (Greene et al., 2019). This suggests that within subgroups, including individuals with variability in lower-level symptom counts and higher-level factors scores (derived from the full sample), the *p*-factor is of little utility.

Our results, therefore, lend support to the suggestion that the *p*-factor is poor at representing variations in the number of symptoms displayed per disorder at the subgroup level, while the specific factors of psychopathology can better account for deviations. Future research should consider these findings when discussing the use of the *p*-factor in individualised treatment settings. Furthermore, it is important to note that the four subgroups fit the revised bifactor model rather than the standard bifactor model. Heinrich et al. (2020) explain that *p* in the revised bifactor model, which did not include the specific thought disorder factor, reflects the thought disorder factor and not general psychopathology. When Caspi et al. (2014) removed the thought disorder-specific factor from the model, it became an "S−1" bifactor model (see Burns et al., 2019; Eid, 2020; Heinrich et al., 2020). The *p*-factor came to represent thought disorder, and internalising and externalising reflected the variance of their indicators over and above the variance subsumed by the thought disorder-referenced *p*-factor (Heinrich et al., 2020). Therefore, as no subgroups fit the standard bifactor model well and only a single subgroup fit the one-factor model well, *p* as understood as a general factor of psychopathology did poorly at accounting for the symptoms of the subgroups. It is also important to acknowledge that the factor loadings and factor covariance for each of the subgroups differed markedly. This variability between different groups mirrors findings of Levin-Aspenson et al. (2020) and supports the proposition that the factors of psychopathology reflect different components within different groups. Ultimately, our results suggest that within subgroups of a population, the factors of psychopathology do poorly at accounting for psychopathological symptoms, especially the *p*-factor.

Associations between the total-sample- and subgroup-derived factor scores were generally very high. This suggests that even though the subgroup-derived factors of

psychopathology have different loading characteristics when compared with the full sample, for the small number of subgroups that fit a model well, the factors' scores derived from within the subgroup largely reflect that of the scores derived from the sample. However, there was one notable exception: the externalising factor scores derived from the subgroups for the correlated factors model, in some instances, differed substantially compared with the population derived scores. It is possible that the externalising factor from these models is more tolerant of deviation, while still fitting the assigned CFA model well. Future research should explore what differs between externalising, internalising, and thought disorder to allow for this tolerance.

The associations between the WAIS subscales and both the subgroup- and total-sample-derived factor scores were very similar. This parallels the finding that, generally, the subgroup- and sample-derived factor scores correlated highly for the subgroups. However, the associations between the WAIS subscales and the factor scores of the total sample differed markedly from many of the subgroups' factor scores. The variation of association strength between the subgroups' and the full samples' factor scores and the WAIS subscales has implications for the development of a substantive construct of the factors, including the *p*-factor. For example, if a substantive construct of the *p*-factor was to be developed, with prespecified constructs with which it must correlate within a certain range (e.g., Levin-Aspenson et al., 2020), it is likely this definition of *p* would only be applicable to a full population sample. Therefore, our results suggest that not only do the factors of psychopathology do poorly at accounting for the symptoms of subgroups, especially *p*, but also even for the small number of subgroups for which the factors do have utility, large variations in outcome correlates (in our case WAIS subscales) suggest that developing a universal substantive meaning of the factors would be challenging. Furthermore, these results also reinforce the importance of considering neurocognitive associations with psychopathology at the individual level (Haywood & Baughman, 2021; Haywood et al., 2021c).

Recently, authors have described the S−1 bifactor approach to the study of psychopathology as an alternative to the standard bifactor approach (Burns et al., 2019; Eid, 2020; Heinrich et al., 2020). The S−1 approach allows for the predefinition of the substantive construct of the general factor by loading theoretically important indicators representing a domain of interest directly on the general factor (Burns et al., 2019; Haywood et al., 2021a). Indeed, the S−1 approach combats many of the issues that we present here with regard to the

bifactor model. The approach allows for a clear understanding of what the general factor reflects and for a clear interpretation of what the specific variances, factor loadings, and covariance represent (Eid, 2020). The S−1 approach is clearly useful for exploring specific questions and the utility of a specific domain of interest in accounting for the common variance of psychopathological symptoms. However, exclusive use of the S−1 approach means abandoning the possibility of a general factor that may reflect the overall propensity towards psychopathology (e.g., Caspi et al., 2014). Furthermore, the S−1 approach may not be unaffected by some of the issues we raise here. For example, the correlated factors model also did poorly at accounting for the symptoms of our subgroups, and it also had substantial variability in the strengths of correlates. It is possible that the S−1 approach may be subject to similar issues. Further research should explore these possibilities.

2.4.1 Limitations and Directions for Future Research

This study, while accounting for the issues of statistical power when conducting this type of research, had some limitations. The simulation approach taken means that the results need to be interpreted with some caution. It is likely that, even though symptom and WAIS data were developed based on known sample characteristics (Caspi et al., 2014), the simulation data differed from the simulated sample. We based symptom distributions on the general skew of symptoms found in the population (Curran et al., 1996), and this has been successfully utilised in previous research (Greene et al., 2019). However, it is likely that the distributions from the base sample varied from our simulated data, which may have implications on the characteristics of the subgroups. Furthermore, we chose to develop a single dataset instead of using Monte Carlo simulations. The use of a single dataset had strengths that were important to this research, such as allowing for a nuanced assessment of the fit, loadings, and associations of each subgroup and the parallel with empirical research. However, Monte Carlo simulations in which the populations conditions may be varied systematically would offer further insight into the boundaries of structural model applicability. We also used continuous variables representing symptom counts. Once again, even though this has been used successfully in previous research (e.g., Greene et al., 2019), it differs from the ordinal data derived from the base sample. Further, the first level of categorisation we used for the subgroups was the symptom level. We used the threshold of at least one disorder representing each factor to be in the lower, middle, or upper third of the total sample. This approach was chosen to account for individuals with just one pervasive difficulty

(e.g., pervasive alcohol use while not displaying significant comorbidity). However, this also meant that individuals were often nested within more than one subgroup. Lastly, we tested the four models of psychopathology used by Caspi et al. (2014). These models were chosen based on their popularity in the literature. However, we did not test a higher-order factor model—a model where the specific factors load onto the p-factor rather than the *p*-factor taking variance directly from the disorder level (see Bornovalova et al., 2020)—or a more traditional bifactor model that does not have correlated specific factors. It is possible that the use of a higher-order factor model or a traditional bifactor model might have produced differential findings.

Future research should look to use Monte Carlo simulations to attempt to validate and extend these findings and also explore various different grouping methods to accommodate the nesting limitation mentioned above. However, even though simulation and computational methodologies are useful for providing a way to discretely and precisely examine a research question (see Haywood, Lawrence, et al., 2021), research using large samples of human data is needed to validate these findings. Future research may also measure a larger range of correlates and assess the applicability of the higher-order factor model and the traditional bifactor model at the subgroup level. Research should then explore the development and maintenance of the factor scores on the individual level. This could be done through computational approaches exploring how variations of symptoms over time impact an individual's factor scores at different time points. Lastly, future research should explore the availability, variability, and correlates of the factors of psychopathology in S−1 bifactor models and should determine if the limitations of the structural models that we present here are applicable to that approach.

2.5. Conclusions

Ultimately, our work suggests that the models of psychopathology we tested are poor at accounting for the symptoms of psychopathology within subgroups of the population. Furthermore, with respect to the utility of the *p*-factor, we showed that *p* has little utility in accounting for the symptoms of individuals within subgroups. Associations between the WAIS subscales and psychopathology factor scores within the subgroups had significant variability and often differed markedly from the associations in the total sample. Together, these findings not only suggest that bifactor models of psychopathology be of limited utility at the subgroup level, and therefore at the individual level, but also suggest that developing universal substantive constructs of the factors may be difficult. If universal substantive constructs of the *p*-factor and

the specific factors of psychopathology are to be developed, they may only be useful for describing a sample or population as a whole. The factors may have little utility in describing a subgroup's psychopathological symptom structure. This has implications for the use of the factors within the treatment setting.

Preface to Chapter 3

Chapter three explores the second direction of future research provided in chapter one of road forward in the exploration of psychopathology and neurocognitive abilities. This second direction of future research coincides with the second aim of this thesis. The following chapter also moves forward from the findings in chapter two in that a universal consistent *p*-factor for the exploration between psychopathology and neurocognitive abilities is unlikely to be developed. As with chapter three we use data simulation methods. However here we created datasets to provide three examples of the use of S-1 bifactor models, where the general factor (the *p*-factor in traditional psychopathology bifactor models) is referenced by neurocognitive abilities. This chapter ultimately provides examples and a discussion of how the use of S-1 bifactor models (with correlated factors models) offers distinct advantages over more traditional methods of assessment and mitigates the issue of an inconsistent and non-replicable general factor for the study of psychopathology and neurocognitive abilities.

Note: The following chapter has been published in International Journal of Environmental Research and Public Health in the special issue titled Diagnosis and Advances in Research on Human Behavior.

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2021). Going “up” to move forward: S-1 bifactor models and the study of neurocognitive abilities in psychopathology. *International Journal of Environmental Research and Public Health, 18*(14), 7413. https://doi.org/10.3390/ijerph18147413

Minor edits have been made to the following chapter, such as phrasing and Australian Spelling, to ensure consistency within the thesis. This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

Chapter 3 (Study 2): Going "Up" to Move Forward: S-1 Bifactor Models and the Study of Neurocognitive Abilities in Psychopathology

3.1 Introduction

In recent years, there has been a body of research that has moved away from the study of any single disorder (see Lahey et al., 2021a). This work is driven by issues of high comorbidity and low diagnostic stability within the traditional nosological approach to diagnosis (see Hovenkamp-Hermelink et al., 2016; Krueger & Eaton, 2015; Newman et al., 1998). In particular, the rise of dimensional structural models of psychopathology has led to explorations of the commonalities of disorders, as well as what may underpin these commonalities (Lahey et al., 2021a). One of the most prominent structural models of psychopathology, the bifactor model, revealed that a significant amount of variance from symptoms of a range of disorders could be accounted for by a single general factor at the "top" of the model, while specific disorder variance could be largely accounted for by a group of lower-order, or "specific" factors, such as externalising, internalising and thought disorder (e.g., Caspi et al., 2014). The general factor was termed the *p*-factor, likened to the *g* factor of intelligence, and said to be a normally distributed property across the population that determines an individual's propensity toward all common psychopathological symptoms (Caspi et al., 2014). A range of research, with little consensus, has attempted to uncover what is the substantive construct of *p*, or in other words, what *p* represents. For example, the *p*-factor has been claimed to reflect neuroticism (Brandes et al., 2019), disordered thought (Caspi & Moffitt, 2018), functional impairment (Smith et al., 2020), and impulsive responsivity to emotion (Carver et al., 2017). Furthermore, constructs, such as neurocognitive abilities, due to their reliable correlation with the general factor, have been claimed to be a key driver of the *p*-factor (Lahey et al., 2021a).

Higher-order neurocognitive abilities have long been theorized to be important components and processes underlying the development and maintenance of psychopathology (e.g., Beck & Rector, 2005; Cannon, 2015; Kéri & Janka, 2004; Trivedi, 2006). However, evidence of the contributions of higher-order neurocognitive abilities to psychopathology is often inconsistent (e.g., Bloemen et al., 2018; Geurts et al., 2014; Kofler et al., 2019; Raffard & Bayard, 2012). One possible reason for this heterogeneity may be the diagnostic instability and comorbidity present in research grounded in the nosological approach (e.g., see Trivedi, 2006).

Therefore, exploring how neurocognitive abilities may contribute to *p*, and the specific factors of psychopathology using the dimensional based structural approach, is appealing.

However, recently, there has been strong evidence against *p* as a substantive construct. Murray et al. (2016) and Snyder and Hankin (2017) explain that the *p*-factor, is a function of the sample from which it is derived. Levin-Aspenson et al. (2020) demonstrated that the *p*-factor derived from two different samples is a substantially different construct, and Fried et al. (2021) showed that, statistically, *p* is simply a representation of the combination of an individual's diagnosis. Furthermore, using simulation methodologies, Greene et al. (2019) showed that fit indices, often used to champion the bifactor model (with a *p*-factor) over a correlated factors model (without a *p*-factor), unfairly bias the more accommodating bifactor model. Correlations between specific factors in a bifactor model also often switch signs when compared to the specific factor associations in the correlated factors models (e.g., Caspi et al., 2014), and these changes do not have a strong theoretical explanation (Pettersson et al., 2021). Furthermore, recently, we have demonstrated the particular lack of applicability and consistency of the *p*-factor within subgroups of a population (Haywood et al., 2021b), limiting the possibility of a universal substantive *p*. Ultimately, Lahey et al. (2021a) explains that *p* is simply "…a "weighted average" of some aspects of all symptoms exhibited by each person at that point in time" (pp. 61), and it is unclear whether *p* can have any substantive, theoretical meaning.

As neurocognitive abilities are associated with a wide range of disorders, *p* as a substantive construct has promise for increasing our understanding of how neurocognitive abilities are involved in the development, maintenance, and treatment of psychopathology. However, without a theoretical consensus on what *p* is (See Fried et al., 2021; Levin-Aspenson et al., 2020; Watts et al., 2020a), it may not greatly enhance our understanding of the association between neurocognitive abilities and psychopathology. Relatedly, as *p* is inherently fluid, changes in the makeup of *p* also result in substantive changes to the specific factors of the models, further limiting our ability to consistently interpret the associations between neurocognitive abilities and psychopathology. To combat the statistical concerns of *p*, alternative bifactor models have been developed. In particular, there has been increasing interest in the use of the S-1 bifactor model in the study of psychopathology (Burns et al., 2019; Eid, 2020; Heinrich et al., 2020). The S-1 bifactor model includes a "reference domain" that acts to predefine the meaning of the general factor (see Burns et al., 2019; Eid, 2020; Heinrich et al.,

2020, for detailed explanations). The predefining of the general factor removes the issues presented in the traditional bifactor literature in that *p* is an undefined, flexible statistical construct (Eid, 2020). An S-1 bifactor model includes a reference domain with typically two or more indicators that load only onto the general factor, while the other indicators load onto the general factor as well as a specific factor (Burns et al., 2019). The general factor in a S-1 bifactor model therefore represents the reference domain, and the specific factors represent the "…true score variance in non-reference symptom facets that is not shared with the general reference factor" (Burns et al., 2019, pp. 885). Further, the correlations between the specific factors represent the shared variance between the two factors that is not common with the general factor. Heinrich et al. (2020) explain that when Caspi et al. (2014) removed the thought disorder factor from their bifactor model, due to a Heywood case (an indicator with negative variance), their model became an S-1 bifactor model and the thought disorder factor became the reference domain for the *p*. Effectively, *p* in Caspi et al.'s (2014) revised bifactor model *became* the thought disorder factor, rather than a general factor of psychopathology (Heinrich et al., 2020). This demonstrates the difficulties with an undefined general factor (e.g., the *p*-factor), because as *p* is a non-stable statistical weighted summary of symptoms, it is susceptible to changes in meaning in line with changes in model structure and indicators. Therefore, currently, knowledge stemming from associations between the *p*-factor and theoretically important constructs and processes, such as neurocognitive abilities, lack substantive meaning. In contrast, S-1 bifactor models allow us to predefine the meaning of the general factor with a theoretically outstanding candidate (Eid, 2020) and, as the general factor has substantive meaning, unexpected or novel findings, such as specific factors switching signs, could have clear theoretical interpretations and facilitate hypotheses development. Furthermore, a large limitation of traditional bifactor models is inconsistency. However, the inclusion of a reference domain means that the S-1 bifactor models are consistent and therefore replicable (Eid, 2020). Previously, S-1 bifactor models have predominantly been used with a symptom domain as the general factor. However, as Greene et al. (2021) states, any such etiological domain of interest could be modelled as the general factor in an S-1 bifactor model and facilitate the exploration of that domain and psychopathology. Further, Thone et al. (2022) have successfully used S-1 bifactor modelling in multi-trait multi-method models, suggesting that the utility of the S-1 bifactor modelling approach does not arise from common method variance.

Recently, we called for the use of S-1 bifactor models in the study of associations between neurocognitive abilities and psychopathology, by using neurocognitive abilities as the reference domain (Haywood et al., 2021b, 2021c). S-1 bifactor models account for many of the limitations of both the traditional nosological approach, and the traditional bifactor models. In the following sections, we use data simulation methods to illustrate how S-1 bifactor models could be used to examine the association between neurocognitive abilities and psychopathology and how, even when an S-1 bifactor model has unexpected results, it is able to facilitate a theoretical interpretation and hypotheses development.

3.2 Materials and Methods

3.2.1 Data Generation

3.2.1.1. Symptoms

Caspi et al. (2014) developed and tested models of psychopathology using data from the Dunedin Multidisciplinary Health and Development Study (total $N = 1037$; $N = 1000$) used by Caspi et al. (2014). The symptom data were gathered using the Diagnostic Interview Schedule (Robins et al., 1995) and comprised of the number of DSM-IV symptoms with which each individual presented for a range of common disorders (see Caspi et al., 2014). Caspi et al. (2014) also used a range of potential correlates of psychopathological symptoms from the Dunedin Multidisciplinary Health and Development Study, including measures of neurocognitive ability, to further examine their models. To develop our simulated data, we used a top-down approach from previous work (see Haywood et al., 2021b) that is similar to Greene et al.'s (2019) approach in order to develop a dataset comprising of 11 disorder variables for 10,000 participants. Specifically, we used the loadings of Caspi et al.'s (2014) revised bifactor model to develop 11 continuous variables representing the symptom counts of (1) alcohol use, (2) cannabis use, (3) hard drug use, (4) tobacco use, (5) conduct disorder, (6) fears and phobias, (7) major depressive episode, (8) generalised anxiety disorder, (9) obsessive compulsive disorder, (10) mania, and (11) schizophrenia, respectively. Like Greene et al. (2019), we then assigned a skew of approximately positive skew of 2.0 across the variables to represent the distributions of symptoms typically found in the general population (Curran et al., 1996). All data generation and analysis were conducted with RStudio using the Lavaan package (Rosseel, 2012).

3.2.1.2. Intelligence

Following the development of the symptom counts, we fitted the Caspi et al. (2014) revised bifactor model to the data (details are presented in the analysis section) and saved the factor loadings back to the dataset. We then developed three variables simulating Caspi et al.'s (2014) measures of the intelligence quotient (IQ) from (1) the Stanford–Binet Intelligence Scale (age 5), (2) the Wechsler Intelligence Scale for Children-Revised (WISC-R; ages 7–11), and (3) the Wechsler Adult Intelligence Scale-IV (WAIS-IV) full scale. The data for the three IQ measures were based on the correlations between each measure of IQ and externalising, internalising and the *p*-factor from Caspi et al.'s (2014) revised bifactor model, as well as the correlations between each of the IQ measures from longitudinal research (Kaufman & Van Hagen, 1977; Spector, 2013). The IQ variables were normally distributed and standardised to a mean of 100 and a standard deviation of 15.

3.2.1.3. Executive Functioning

We also developed, and added to the dataset, variables representing two of Caspi et al.'s (2014) measures of executive functioning, the Trail Making Test-B (TMT-B) and the Cambridge Neuropsychological Test Automated Battery - Rapid Visual Information Processing task (CANTAB: RVIP). We developed these data based on the correlations between the two executive functioning measures and externalising, internalising and the *p*-factor from Caspi et al.'s (2014) revised bifactor model, as well as correlations found within the literature between the two executive functioning variables, and adult measures of IQ (Ardila et al., 2000; Green et al., 2019; Smith et al., 2013). To illustrate how S-1 bifactor models can facilitate the interpretation of novel findings, we also developed a second set of data for the TMT-B and the CANTAB: RIVP. We developed the second set of data for these measures based on the unlikely scenario of the measures having an $r = 0.8$ correlation, an undefined association with *g*, no correlation with externalising or internalising, but a maximum possible association with *p* from the revised bifactor model (actual correlations differed slightly to the assigned correlations in the data producing code, due to random data generation factors and association compatibility constraints).

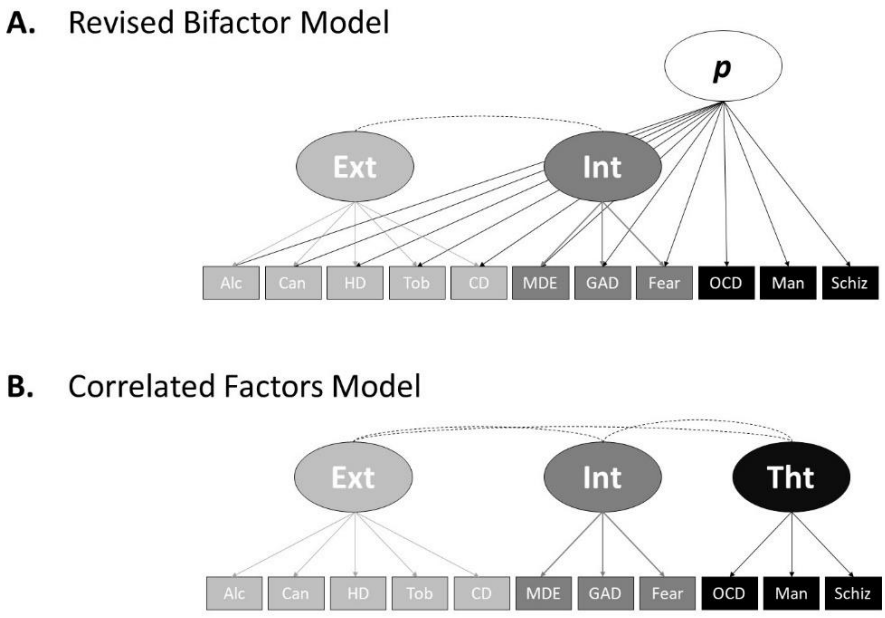
3.2.2 Analysis

First, to validate the simulated dataset, and to act as a comparison to the S-1 bifactor models, we tested the fit of two of Caspi et al.'s (2014) structural models (see Figure 3.1), (A)

the revised bifactor model that the simulated data was based on, and (B) the correlated factors model, a popular model in psychopathology research (e.g., see Caspi et al., 2014).

Figure 3.1

Caspi et al.'s (2014) Confirmatory Factor Analysis Models



Note. Depiction of the bifactor model and the correlated factors model adapted from Caspi et al. (2014). Panel A: Revised bifactor model. Panel B: Correlated factors model. Alc = alcohol. Can = cannabis. HD = Hard drugs. Tob = Tobacco, CD = Conduct disorder. MDE= Major depressive episode. GAD = Generalized anxiety disorder. Fear = Fears and phobias. OCD = obsessive compulsive disorder. Man = Mania. Schiz = Schizophrenia. Ext = Externalising. Int = Internalising. Tht = Thought Disorder.

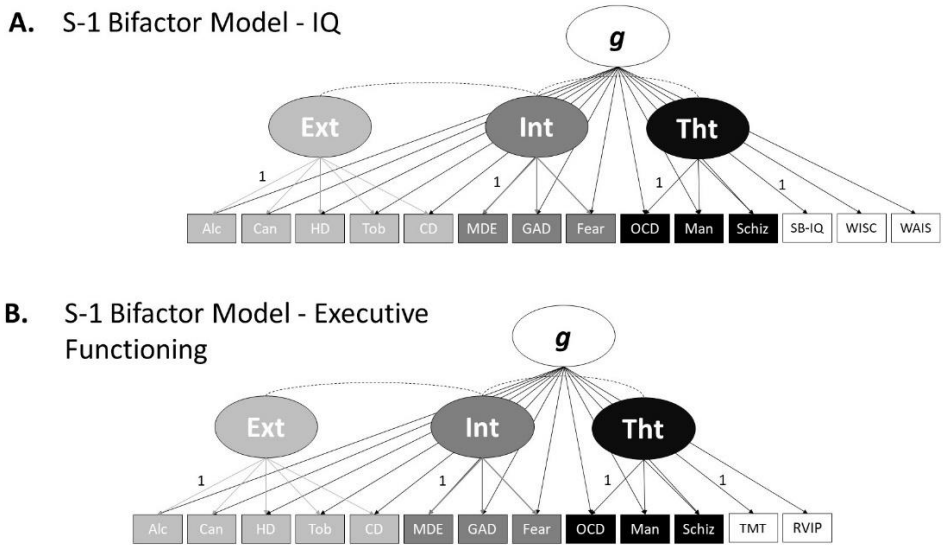
For both confirmatory factor analyses (CFAs), we used a maximum likelihood estimation with robust standard errors (MLR), and Pearson's correlations in RStudio. MLR is robust to deviations of normality, such as symptom count data, by correcting chi-square statistics and standard errors to compensate for skewed data. MLR is also widely used in psychopathology

research (see Greene et al., 2019) and is used for continuous indicators, such as our symptom count dimensions. We used the root mean square error of approximation (RMSEA; Steiger, 1990), the Tucker–Lewis Index (TLI) and the comparative fit index (CFI; Bentler, 1990) to determine model fit, while we also report standardised root mean square residual (SRMR; Hu & Bentler, 1995). A good-fitting model was determined by RMSEA values of <0.05 (Bollen & Curran, 2006) and CFI and TLI values of >.95. However, it is important to note that good model fit, is not a theoretically robust way to choose a model and is not the focus of this research (e.g., see Greene et al., 2019). Rather, the loading patterns and specific factor covariance will be the focus in this paper.

Next, we tested the fit of three S-1 bifactor models using the simulated datasets (see Figure 3.2). The first S-1 bifactor model (Figure 3.2A) used IQ over time, measured by (1) the Stanford–Binet Intelligence Scale (age 5), (2) the WISC-R (ages 7–11), and (3) the WAIS-IV full scale, as the reference domain for the general factor (in this model, the IQ factor), and externalising, internalising and thought disorder as specific factors. Each disorder loaded onto the general factor as well as one of the specific factors as per Caspi et al. (2014). Following the directions of Burns et al. (2019), Heinrich et al. (2020), and Eid (2020), the unstandardised loading of the first indicator of the reference factor, and each specific factor, was fixed to 1, and acted as a reference indicator for that factor. The fit of the S-1 bifactor models was determined using the same criteria as the revised bifactor and correlated factors models.

Figure 3.2.

S-1 Bifactor Confirmatory Factor Analysis Models



Note. Depiction of the two S-1 bifactor models used in this research. Panel A: S-1 bifactor model—IQ. Panel B: S-1 bifactor model executive functioning. IQ = Intelligence quotient. Alc = alcohol. Can = cannabis. HD = Hard drugs. Tob = Tobacco, CD = Conduct disorder. MDE = Major depressive episode. GAD = Generalized anxiety disorder. Fear = Fears and phobias. OCD = Obsessive compulsive disorder. Man = Mania. Schiz = Schizophrenia. SB-IQ = Stanford–Binet Intelligence Scale—intelligence quotient. WISC = Wechsler Intelligence Scale for Children—intelligence quotient. WAIS = Wechsler Adult Intelligence—intelligence quotient. TMT = Trail Making Test. RVIP = Rapid Visual Information Processing. Ext = Externalising. Int = Internalising. Tht = Thought Disorder. *g* = General Factor

The second S-1 bifactor model (see Figure 3.2B) used the same specifications as above. However, executive functioning, measured by the TMT-B and the CANTAB: RVIP, acted as the reference domain for the general factor (in this model, the executive functioning factor) and the first indicator of the factor (TMT-B) acted as the reference indicator. Externalising, internalising and thought disorder remained the specific factors and their first indicator, respectively, remained as the reference indicator.

The final S-1 bifactor model was used to illustrate that even when unexpected, or novel findings occur in S-1 bifactor models, they may have theoretical explanations and drive hypotheses. The S-1 bifactor model was identical to the executive functioning reference domain model above (Figure 3.2B). However, unlike the model above, which used simulated executive functioning data based on empirical research, this model used data for the TMT-B and CANTAB: RVIP, based on a very unlikely combination of correlations (see data generation section).

3.3. Results

The simulated data fit the Caspi et al. (Caspi et al., 2014) revised bifactor model well (Figure 2.3A) ($\chi 2(35, N = 10{,}000) = 44.78$, CFI = 1.00, TLI = 1.00, SRMR = 0.005, RMSEA = 0.004, 90% confidence interval (CI) = [0.000, 0.009]), as well as the correlated factors model (Figure 3.2B), $\chi 2(41, N = 10{,}000) = 432.27$, CFI = 0.990, TLI = 0.987, SRMR = 0.29, RMSEA = 0.031, 90% CI) = [0.028, 0.034]). Table 3.1 shows the loadings and association characteristics of the revised bifactor model, and Table 3.2 shows the loadings and association characteristics for the correlated factors model. The factor loadings and correlations with IQ and executive functioning do slightly differ to Caspi et al. (2014) due to random data generation factors, potential skew differences between Caspi et al.'s (2014) data and the simulation data, different estimators, correlation compatibility constraints, and the use of continuous instead of ordinal variables. However, our models' loading patterns and characteristics, as well as factor associations with IQ and executive functioning, closely resemble that of Caspi et al. (2014). Therefore, we conclude that our simulated data represent that of Caspi et al. (2014) well.

Table 3.1

Loadings and Associations of the Revised Bifactor Model.

Factors	Alc	Can	HD	Tob	CD	MDE	GAD	Fear	OCD	Mania	Schiz	SB-IQ~	WISC-IQ~	WAIS-IQ~	TMT~	RVIP~	Ext~Int
																	−0.366 **
p (Unstandardised)	0.284 (0.205)	0.311 (0.251)	0.336 (0.258)	0.391 (0.265)	0.387 (0.304)	0.609 (0.477)	0.589 (0.471)	0.472 (0.328)	0.695 (0.361)	0.969 (0.507)	0.804 (0.428)	−0.252 **	−0.129 **	−0.231 **	0.133 **	−0.181 **	
Ext (Unstandardised)	0.546 (0.394)	0.628 (0.508)	0.558 (0.428)	0.383 (0.260)	0.557 (0.437)	-	-	-	-	-	-	0.000	0.054 **	−0.042 **	−0.045 **	−0.026 **	
Int (Unstandardised)	-	-	-	-	-	0.248 (0.194)	0.394 (0.315)	0.334 (0.233)	-	-	-	−0.004	−0.056 **	0.027 *	0.061 **	0.004	

Note. The indication of significance appears below the corresponding association figure in the table. ~ = Correlation. Alc = Alcohol. Can = Cannabis. HD = Hard drugs. CD = Conduct disorder. MDE= Major depressive episode. GAD = Generalized anxiety disorder. Fear = Fears and phobias. OCD = Obsessive compulsive disorder. Schiz = Schizophrenia. SB-IQ = Stanford–Binet Intelligence Scale - intelligence quotient. WISC-IQ = Wechsler Intelligence Scale for Children - intelligence quotient. WAIS-IQ = Wechsler Adult Intelligence Scale - intelligence quotient. TMT = Trail Making Test. RVIP = Rapid Visual Information Processing. Ext = Externalising. Int = Internalising.

All loadings significant at $p < 0.01$. * = $p < 0.05$. ** = $p < 0.01$.

Table 3.2

Loadings and Associations of the Correlated Factors Model

Factors	Alc	Can	HD	Tob	CD	MDE	GAD	Fear	OCD	Mania	Schiz	SB-IQ~	WISC-IQ~	WAIS-IQ~	TMT~	RVIP~	Ext ~Int	Ext~Tht	Int~Tht
																	0.307 **	0.530 **	0.870 **
Ext (Unstandardised)	0.604 (0.436)	0.679 (0.549)	0.649 (0.498)	0.553 (0.375)	0.687 (0.540)	-	-	-	-	-	-	−0.145 **	−0.031 *	−0.166 **	0.041 *	−0.126 **			
Int (Unstandardised)	-	-	-	-	-	0.685 (0.537)	0.683 (0.546)	0.555 (0.386)	-	-	-	−0.237 **	-0.138 **	−0.206 **	0.144 **	−0.167 **			
Tht (Unstandardised)	-	-	-	-	-	-	-	-	0.695 (0.361)	0.969 (0.507)	0.803 (0.428)	−0.252 **	−0.129 **	−0.231 *	0.133 **	-0.181 **			

Note. The indication of significance appears below the corresponding association figure in the table. ~ = Correlation. Alc = Alcohol. Can = Cannabis. HD = Hard drugs. CD = Conduct disorder. MDE = Major depressive episode. GAD = Generalized anxiety disorder. Fear = Fears and phobias. OCD = Obsessive compulsive disorder. Schiz = Schizophrenia. SB-IQ = Stanford–Binet Intelligence Scale-intelligence quotient. WISC-IQ = Wechsler Intelligence Scale for Children Intelligence Quotient. WAIS-IQ = Wechsler Adult Intelligence Scale - intelligence quotient. TMT = Trail Making Test. RVIP = Rapid Visual Information Processing. Ext = Externalising. Int = Internalising. Tht = Thought Disorder.

All loadings significant at $p < 0.01$. * = significant at $p < 0.05$. ** = significant at $p < 0.01$.

Table 3.3 shows the loadings and association characteristics of the first S-1 bifactor model (Figure 3.2A). This S-1 bifactor model used IQ over time as the reference domain and the first indicator of each factor as the reference indicator. The data fit the model well ($\chi 2(63, N = 10,000) = 787.91$, CFI = 0.988 TLI = 0.982, SRMR = 0.030, RMSEA = 0.034, 90% CI) = [0.032, 0.036]). Largely, as the IQ general factor's loadings on the symptom indicators show, the IQ general factor did poorly at accounting for variance amongst the symptoms. However, on closer inspection, the IQ general factor accounted for notably more variance amongst the internalising and thought disorder indicators when compared to the externalising indicators. This mirrors the trend of bivariate correlations between the specific factors and the measures of IQ in the correlated factors model above. In the correlated factors model, associations between the specific factors and the measures of IQ represent the correlations between the variance of the items that load onto each specific factor, respectively. In the S-1 bifactor model, loadings of indicators on the general, predefined, factor represent the variance of each indicator that is accounted for by that factor. The S-1 bifactor model therefore not only answers a different research question (e.g., "what amount of variance of each symptom can be accounted for by the general predefined factor?") when compared to the correlated factors model, but also allows us to explore the partial associations between the specific factors after accounting for the general predefined factor. As Table 3.3 shows, the covariation between the specific factors fell slightly when compared to the correlated factors model (Table 3.2) and fell by a similar magnitude, indicating that the IQ general factor accounts for a small amount of the association between externalising, internalising and thought disorder.

Table 3.3

Loadings and Associations of the IQ S-1 Bifactor Model

Factors	SB-IQ	WISC-IQ	WAIS-IQ	Alc	Can	HD	Tob	CD	MDE	GAD	Fear	OCD	Mania	Schiz	Ext~Int	Ext~Tht	Int~Tht
															0.287 **	0.518 **	0.863 **
IQ *g* (Unstandardised)	0.936 (1.00)	0.790 (0.844)	0.826 (0.883)	−0.071 (−0.004)	−0.085 (−0.005)	−0.078 (−0.004)	−0.082 (−0.004)	−0.099 (−0.006)	−0.159 (−0.009)	−0.149 (−0.009)	−0.124 (−0.006)	−0.188 (−0.007)	−0.255 (−0.010)	−0.201 (−0.008)			
Ext (Unstandardised)	-	-	-	0.601 (1.00)	0.674 (1.26)	0.645 (1.14)	0.546 (0.856)	0.679 (1.23)									
Int (Unstandardised)				-	-	-	-	-	0.666 (1.00)	0.668 (1.02)	0.541 (0.722)	-	-	-			
Tht (Unstandardised)				-	-	-	-	-	-	-	-	0.669 (1.00)	0.935 (1.41)	0.778 (1.19)			

Note. The indication of significance appears below the corresponding association figure in the table. ~ = Correlation. IQ *g* = Intelligence quotient general factor. Alc = Alcohol. Can = Cannabis. HD = Hard drugs. CD = Conduct disorder. MDE = Major depressive episode. GAD = Generalized anxiety disorder. Fears = Fears and phobias. OCD = Obsessive compulsive disorder. Schiz = Schizophrenia. SB-IQ = Stanford–Binet Intelligence Scale - intelligence quotient. WISC-IQ = Wechsler Intelligence Scale for Children - intelligence quotient. WAIS-IQ = Wechsler Adult Intelligence Scale - intelligence quotient. Ext = Externalising. Int = Internalising. Tht = Thought Disorder.

All loadings significant at p <.01. * = $p < 0.05$. ** = $p < 0.01$.

Table 3.4. shows the loadings and association characteristics of the second S-1 bifactor model (Figure 3.2B). This S-1 bifactor model used executive functioning as the reference domain and the first indicator of each factor as the reference indicator. The data fit the model well ($\chi 2(51, N = 10{,}000) = 419.96$, CFI = 0.991, TLI = 0.987, SRMR = 0.023, RMSEA = 0.027, 90% CI) = [0.025, 0.029]). The loading of the CANTAB-RVIP was negative as lower TMT-B scores reflect better performance and TMT-B was the reference indicator for the general factor. The executive functioning referenced general factor did better than the IQ referenced general factor in accounting for variance amongst the symptoms. However, the executive functioning general factor accounted for notably less variance amongst the first three externalising indicators (alcohol, cannabis, and hard drugs use) when compared to the rest of the symptoms. The executive functioning general factor did better than the IQ general factor at accounting for tobacco use and conduct disorder symptoms, comparatively similar when accounting for internalising indicators, and better when accounting for thought disorder indicators. The finding that the executive functioning general factor did notably poorer at accounting for alcohol, cannabis and hard drugs use when compared to the other indicators might inform hypotheses regarding their aetiological interrelations and separability when compared to other symptomology. As Table 3.4 shows, like with the IQ referenced S-1 model, partial covariation between the specific factors fell slightly when compared to the correlated factors model (Table 3.2). However, association between specific factors fell by a greater magnitude when compared to the IQ referenced domain S-1 model. The association between the externalising and thought disorder factors, and the association between internalising and thought disorder factors fell, when compared to the correlated factors model by a similar magnitude (0.18 and 0.16, respectively). However, the association between externalising and internalising fell to a greater extent (0.34). This may inform hypotheses, such as executive functioning completely mediating the association between internalising symptoms (e.g., generalised anxiety) and externalising behaviours (e.g., substance use).

Table 3.4

Loadings and Associations of the Executive Function S-1 Bifactor Model 1

Factors	TMT	RVIP	Alc	Can	HD	Tob	CD	MDE	GAD	Fears	OCD	Mania	Schiz	Ext~Int	Ext~Tht	Int~Tht
														0.273 **	0.512 **	0.854 **
EF *g* (Unstandardised)	0.467 (1.00)	−0.552 (−1.22)	0.046 (0.069)	0.077 (0.130)	0.083 (0.132)	0.174 (0.132)	0.132 (0.217)	0.237 (0.387)	0.223 (0.372)	0.166 (0.241)	0.238 (0.258)	0.329 (0.359)	0.270 (0.300)			
Ext (Unstandardised)	-	-	0.608 (1.00)	0.678 (1.25)	0.645 (1.129)	0.529 (0.818)	0.672 (1.21)									
Int (Unstandardised)			-	-	-	-	-	0.642 (1.00)	0.647 (1.03)	0.531 (0.734)	-	-	-			
Tht (Unstandardised)			-	-	-	-	-	-	-	-	0.653 (1.00)	0.912 (1.41)	0.757 (1.20)			

Note. The indication of significance appears below the corresponding association figure in the table. ~ = Correlation. EF g = Executive function general factor. Alc = Alcohol. Cann = Cannabis. HD = Hard drugs. CD = Conduct disorder. MDE = Major depressive episode. GAD = Generalized anxiety disorder. Fears = Fears and phobias. OCD = Obsessive compulsive disorder. Schiz = Schizophrenia. TMT = Trail Making Test. RVIP = Rapid Visual Information Processing. Ext = Externalising. Int = Internalising. Tht = Thought Disorder

All loadings significant at p <.01. ** = $p < 0.01$. * = $p < 0.05$.

Lastly, to demonstrate how S-1 bifactor models facilitate hypothesis generation from unexpected or novel findings, we tested the same S-1 bifactor model above, but with executive functioning data with a highly unexpected correlation matrix (see Data Generation, section 3.2.1). The data fit the model well ($\chi2(51, N = 10{,}000) = 215.17$, CFI = 0.997, TLI = 0.996, SRMR = 0.12, RMSEA = 0.018, 90% CI) = [0.18, 0.020]). The loadings and associations of this model are displayed in Table 3.5. The general factor, with executive functioning as the reference domain, did well in accounting for variance amongst the symptoms. As expected, due to the measures' assigned correlation with *p* from the revised bifactor model, the thought disorder indicators loaded the highest on the executive functioning general factor in this S-1 bifactor model. Loadings of the externalising and internalising indicators on the executive functioning general factor were also comparatively high, due to the common variance between the specific factors, even though they were assigned not to correlate in the revised bifactor model during data generation. Due to the highly abnormal executive functioning data, partial associations between the specific factors differ extensively when compared to the correlated factors model and the previous S-1 bifactors models. The association between the externalising and internalising factors switched signs, and the factors had minimal relation (−0.044), while the covariation of the externalising and thought disorder factors, and the internalising and thought disorder factors, dropped substantially when compared to the correlated factors model (0.115 and 0.191). If we were to interpret these patterns of covariation, in particular the negative externalising and internalising factors association, in a standard bifactor model there would be little theoretical reason for the association to change signs, and substantive interpretation would be difficult due to the ambiguity of the *p*-factor (see Pettersson et al., 2021). However, in a S-1 bifactor model the general factor is defined *a*-priori. The knowledge of what the general factor represents, in this case executive functioning, allows us to make substantive interpretations of the changes in specific factor associations and develop hypotheses as a result. For example, the substantial reduction in the association between the externalising and internalising factors in this S-1 bifactor model, when compared to the correlated factors model, may have led to the hypothesis that executive functioning is a full mediator of the association between internalising symptoms, such as depression or anxiety, and externalising behaviours, such as substance use. This demonstrates the utility of using of a typical structural model, the correlated factors model, in conjunction with the S-1 bifactor model for data interpretation. Furthermore, these results may have suggested that, when executive functioning is accounted for, those with more internalising symptoms are conversely slightly less inclined to externalising behaviours. This hypothetical illustration shows that theory building, and testing is a useful characteristic of S-1 bifactor models.

Table 3.5

Loadings and Associations of the Executive Function S-1 Bifactor Model 2.

Factors	TMT-B	RVIP	Alc	Can	HD	Tob	CD	MDE	GAD	Fears	OCD	Mania	Schiz	Ext ~Int	Ext~Tht	Int~Tht
														−0.044 *	0.315 **	0.679 **
EF *g* (Unstandardised)	0.849 (1.00)	−0.980 (−1.06)	0.233 (0.140)	0.265 (0.178)	0.282 (0.180)	0.334 (0.189)	0.324 (0.189)	0.517 (0.338)	0.504 (0.336)	0.392 (0.228)	0.593 (0.257)	0.823 (0.359)	0.676 (0.300)			
Ext (Unstandardised)	-	-	0.568 (1.00)	0.640 (1.26)	0.587 (1.10)	0.434 (0.720)	0.599 (1.15)	-	-	-	-	-	-			
Int (Unstandardised)	-	-	-	-	-	-	-	0.437 (1.00)	0.466 (1.09)	0.407 (0.827)	-	-	-			
Tht (Unstandardised)	-	-	-	-	-	-	-	-	-	-	0.362 (1.00)	0.512 (1.43)	0.433 (1.23)			

Note. The indication of significance appears below the corresponding association figure in the table. ~ = Correlation. EF *g* = Executive function general Factor. Alc = Alcohol. Can = Cannabis. HD = Hard drugs. CD = Conduct disorder. MDE = Major depressive episode. GAD = Generalized anxiety disorder. Fears = Fears and phobias. OCD = Obsessive compulsive disorder. Schiz = Schizophrenia. TMT-B = Trail Making Test-B. RVIP = Rapid Visual Information Processing. Ext = Externalising. Int = Internalising. Tht = Thought Disorder.

All loadings significant at $p < 0.01$. * = $p < 0.05$. ** = $p < 0.01$.

3.4 Discussion

In this paper, we provided the case for the use of S-1 bifactor models in the exploration of neurocognitive abilities in psychopathology. We used simulation methodologies to show how no matter the results of a S-1 bifactor model, using neurocognitive abilities as a reference domain, due to the general factor reflecting a substantive construct, an interpretable hypothesis or theoretical explanation could emerge. S-1 bifactor models account for the issues of substantive and statistical inconsistency of an undefined general factor in psychopathology research (Burns et al., 2019; Eid, 2020; Heinrich et al., 2020). We provided three examples of how S-1 bifactor models could be used to further our understanding of the associations between neurocognitive abilities and psychopathology. In our first example, we used IQ over time as the reference domain for our general factor with externalising, internalising and thought disorder serving as the specific factors. The IQ general factor accounted for a small amount of variance among the symptoms, with the thought disorder indicators generally having the strongest loading on the IQ general factor. This example showed the utility of the S-1 bifactor approach over the sole use of the correlated factors model. Using the S-1 approach we could see the loading of each specific disorder on the IQ general factor, allowing us to examine the proportion of variance in each indicator that was accounted for by the IQ general factor, as well as the indicators loadings on the specific factors. For example, if we look to the associations between the measures of IQ and the externalising, internalising and thought disorder factors in the correlated factors model, we see generally consistent strengths of association (minus WISC-IQ and externalising). However, this only tells part of the story, as the factors reflect the common variance amongst their specific indicators, and not the common variance amongst the indicators after the general factor has been taken into account, as per the S-1 model. Therefore, while correlated factors models can show us the association between IQ and the common variance of indicators for each specific factor, the S-1 bifactor model can show us the common variance amongst the specific indicators once IQ has been taken into account, as well as the loadings of each specific indicator on IQ. Therefore, given the attributes of each approach, we suggest that the correlated factors model and the S-1 bifactor model should be used in parallel to answer different research

questions and provide a range of evidence assessing the association between neurocognitive abilities and psychopathology.

Our second example used executive functioning as the reference domain for the general factor. In this example, the executive functioning general factor typically accounted for more variance in psychopathology symptoms when compared to the IQ general factor. The externalising indicators had the lowest loadings on the executive functioning general factor when compared to the internalising or thought disorder indicators. However, as S-1 bifactor models allow us to examine the loadings of each indicator on the predefined general factor, we can see that alcohol and cannabis use had noticeably lower loadings on the executive functioning general factor when compared to the other indicators. This indicates that, in this instance, when compared to other disorders/symptoms, executive functioning did not have as much utility in accounting for alcohol and cannabis use. Results such as this, due to the knowledge of what the substantive construct of the general factor is, can drive hypotheses for future work. In this model, the associations between the specific factors all fell when compared to the correlated factors model. In particular, the associations between the externalising and internalising factors fell notably. As in S-1 bifactor models the associations between the specific factors are partial associations after accounting for the predefined general factor, it may be possible to hypothesise that executive functioning may be particularly important in the association between internalising symptoms (e.g., anxiety) and externalising behaviours (e.g., substance use).

Lastly, we used the same symptom data but developed data based on unrealistic patterns of associations for the two measures of executive functioning. In the data simulation code, the two measures were made to be highly correlated with each other and the *p*-factor from the revised bifactor model, but that had almost no correlation with the externalising and internalising factors from that model. These data were developed to illustrate that even when models produce novel results, when using a S-1 bifactor approach, the results can still have substantive interpretation due to the known “meaning” of the general factor. In this S-1 model, the general factor, referenced by executive functioning, did well in accounting for symptomology. Loadings on the executive functioning general factor were typically high when compared to the previous models

and the thought disorder indicators again had the strongest loadings on the executive functioning general factor. Here, our focus is on the covariation between the specific factors externalising, internalising and thought disorder. Traditional bifactor models often result in the associations between the specific factors changing signs and differing substantially from the correlated factors model (e.g., Caspi et al., 2014). Pettersson et al. (2021) point out that due to the unspecified nature of the *p*-factor in a traditional bifactor model, associations between the specific factors changing substantially are difficult to interpret and have no clear theoretical explanation. However, if this occurs in a S-1 bifactor model, due to the a priori specification of the general factor, if the associations between specific factors do change substantially, it can be clearly interpreted. In our example, the association between externalising and internalising fell substantially and the two factors were negatively associated. As we clearly understand what the associations between the specific factors in a S-1 bifactor model represent, we may use the results to develop hypotheses. In our example, we may, for instance, hypothesise that executive functioning is all important in the association between internalising symptoms and externalising behaviours, such that executive functioning is a full mediator. Then, as the general factor is predefined, further research could attempt to replicate and build upon this finding. This is impossible using standard bifactor approaches. This research was conducted to inform the use of the S-1 bifactor approach for the study of neurocognitive abilities in psychopathology in the research setting. However, recently, there have been suggestions that structural models such as these may inform a clinicians' practices in the treatment setting. A detailed discussion regarding the utility of these approaches in a treatment setting is beyond the scope of this research. Ruggero et al. (2019) provides information and a case illustration of how structural approaches can guide clinical practice. For example, a clinician taking a dimensional structural approach might, instead of viewing a clients' symptoms as representing a certain diagnosis, view the symptoms as dimensional indicators that share commonality and relations at different hierarchical levels (see Ruggero et al., 2019 for a detailed demonstration) for a detailed demonstration). With regard to the clinical usefulness of the exploration of neurocognitive abilities using the S-1 bifactor approach, it is possible to use the patterns of loadings typically presented between the general

neurocognitive factor and a range of symptomology to inform how the specific collection of symptoms (and their severity) an individual client is experiencing may be functionally associated with their neurocognitive performance and direct their treatment accordingly.

3.4.1 Limitations of the S-1 Bifactor Approach in the Study of Neurocognitive Abilities and Psychopathology

It is important to acknowledge the potential limitations of the S-1 bifactor approach for studying neurocognitive abilities in psychopathology. There are three primary limitations for using the S-1 bifactor approach: first, the sole use of the S-1 approach would remove the ability to examine the associations of particular neurocognitive abilities and psychopathological symptoms or factors. However, the combined use of the S-1 approach and the correlated factors model (with bivariate correlations) mitigates this issue. Second, similar to other structural models (see Haywood et al., 2021b), it is likely that, the S-1 bifactor model would have limited utility subgroups of a population. However, as the general factor is predefined S-1 bifactor models would likely have better consistency when compared to other structural models. This means that the S-1 bifactor model may only be useful for a general population sample, or a sample with large variability in symptoms and limited symptom heterogeneity (see Haywood et al., 2021b). Third, given that having neurocognitive abilities modelled as the general factor results in symptoms loading directly on the factor, it limits our ability to explore nuanced proposals (i.e., the multidimensional hypothesis Haywood & Baughman, 2021) of the heterogeneity of neurocognition within psychopathology on the individual level. However, again mitigating this issue by using the S-1 approach in conjunction with the correlated factors model approach, it may be possible to explore the heterogeneity of neurocognitive abilities within the factors derived from a correlated factors model on the individual level (Haywood et al., 2021c)

3.4.2 Limitations of the Research and Directions for Future Research

This study, while demonstrating the usefulness of the S-1 bifactor model to explore neurocognitive abilities and psychopathology, did have some limitations. First, all data used were simulated from Caspi et al. (2014). The simulated data approach allowed for useful control over the data to facilitate the demonstration of different modelling circumstances (e.g., the data based

on a very unlikely combination of executive functioning correlations). However, the results should be considered as a demonstration of the use of the S-1 bifactor model in this context, rather than used to elucidate any important contributions of neurocognitive abilities toward psychopathology. In line with this, we developed continuous data for 11 disorder categories that summarised Caspi et al.'s (2014) longitudinal data. This, while providing a neater dataset to demonstrate the usefulness of the S-1 bifactor model, did diverge from the base ordinal data that were gathered from five different time points in adulthood. Furthermore, to facilitate a neat demonstration, we did not use Monte-Carlo simulations that are often used in simulation research of this kind, and while we based the positive skew of psychopathological symptoms on empirical data, it is likely that it differed to the skews of individual symptoms from Caspi et al. (2014). We therefore encourage further research to use human data with a number of neurocognitive and symptom measures within a S-1 bifactor approach.

Recently, there have been calls for CFA structural models of psychopathology to be developed, validated and crosschecked with exploratory factor analytic (EFA) approaches to mitigate the issues such as collapsing specific factors and over extraction (Greene et al., 2021). Future S-1 bifactor models may be synergistically developed through the use of both EFA and CFA. Further, even although we advocate for the use of the correlated factors model and the S-1 bifactor model, future research may also continue to explore the traditional bifactor approach, and the possibility of a universal substantive meaning of the *p*-factor. The S-1 bifactor model allows us to examine specific theoretically important variables (e.g., neurocognitive abilities) within a dimensional psychopathology framework (Eid, 2020). However, the traditional bifactor approach facilitates a useful description and, in the future, possible explanations of psychopathological symptoms at the population level. Therefore, we welcome future research examining a theoretical conceptualisation of *p* built on top of its statistical make up (e.g., Fried et al., 2021; Watts et al., 2020a)).

3.5. Conclusions

In this paper, we showed the utility of the S-1 bifactor approach to the study of neurocognitive abilities and psychopathology. We demonstrated the distinct advantages that the

S-1 bifactor model has over the traditional bifactor model for examining the potential contribution of neurocognitive abilities towards psychopathology. Specifically, we show how S-1 bifactor models, using neurocognitive abilities as the reference domain for the general factor, allow for the assessment of each individual indicator's loadings on the neurocognitive ability referenced general factor, and how those factor loadings and the associations between the specific factors, even if unexpected, can inform hypotheses and theoretical understandings. We also suggest that the correlated factors model and the S-1 bifactor model can be used in parallel to explore associations of neurocognitive abilities and psychopathology due to their distinct ability to answer different research questions and facilitate data interpretation through comparison. Lastly, even though we argue for the benefits of the S-1 bifactor model over a traditional bifactor model for the exploration of neurocognitive abilities in psychopathology, we welcome the possibility of the development of a theoretical, substantive conceptualisation of *p* that is useful on the individual and subgroup level (Haywood et al., 2021c), and that can be replicated and is falsifiable.

Preface to Chapter 4

Chapter four explores the third direction of future research in presented chapter one. The third direction of future research coincides with the third aim of this thesis. The following chapter also builds upon the previous chapters by exploring whether each of the factors of psychopathology can be partly explained by discrete association patterns of neurocognitive performance. If discreet association patters were to emerge this would suggest there is utility in further understanding linear trends of neurocognitive deficits relating to each of the factors of psychopathology. Unlike chapters two and three that employed simulation methods, in this chapter we collected data from human participants to examine our aim. This chapter builds upon the previous chapter by using analysis techniques that provide additional information that is distinct from that provided by S-1 bifactor modelling approaches. We collected symptom, substance use, and neurocognitive ability data from a representative community sample. We used the data collected to build and test structural models of psychopathology and examine the neurocognitive correlates of the models' factors. This chapter ultimately provides an empirical assessment and discussion of neurocognitive performance associated with each of the factors of psychopathology to determine if there were clear patterns of differentiation between the factors.

Note: The following chapter has been published in Brain Sciences in the special issue titled Diagnosis and Advances in Research on Human Behavior.

Haywood, D., Baughman, F.D., Mullan, B.A., and Heslop, K.R. (2022). What accounts for the factors of psychopathology? An investigation of the neurocognitive correlates of internalising, externalising, and the *p*-factor. *Brain Sciences, 12*(4). 421. https://doi.org/10.3390/brainsci12040421

Minor edits have been made to the following chapter, such as phasing and Australian Spelling, to ensure consistency within the thesis. This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

Chapter 4 (Study 3): What accounts for the factors of psychopathology? An investigation of the neurocognitive correlates of internalising, externalising, and the *p*-factor

4.1 Introduction

Neurocognitive abilities refer to cognitive capabilities grounded in particular neurological properties or systems, and include both higher and lower level cognitive processes (Lezak et al., 2004). Higher level neurocognitive processes include executive functioning that is responsible for the control of mental abilities, including the control of working memory (i.e., updating), attention (i.e., shifting), and predominant responses (i.e., inhibition) (Miyake, Friedman, et al., 2000). While lower-level neurocognitive processes, such as general information processing (i.e., speed of pro-cessing), are more basic to the system (Salthouse, 1996). The proper functioning of neurocognitive abilities, at both higher and lower levels, govern the ability to conduct goal-oriented activity, respond to environmental demands in a timely and appropriate way, and are fundamental to the successful completion of many everyday activities (Lezak et al., 2004). It is therefore understandable that deficits in neurocognitive performance may result in adverse cognitive and behavioural experiences.

Neurocognitive deficits have been consistently associated with a wide range of psychopathological disorders (McTeague et al., 2016). Neurocognitive deficits have been proposed to not only be a consequence of the development of psychopathology but also directly involved in the aetiology of psychopathology (e.g., Beck & Rector, 2005; Romer & Pizzagalli, 2021). It has been suggested that humans are often exposed to novel and conflicting information. To effectively deal with this information, humans need to stop, reflect, and choose the most appropriate behaviours (Cunningham et al., 2007; Romer & Pizzagalli, 2021; Zelazo, 2015). Cunningham et al. (2007), Romer and Pizzagalli (2021), and Zelazo (2015) suggested that the reflection and selection actions require proper neurocognitive performance, including the updating of the contents of working memory, switching between mental sets, inhibiting a predominant response, and effectively processing information. Romer and Pizzagalli (2021) also proposed that deficits in neurocognitive processes may therefore result in the selection of inappropriate behaviours, poor adaptive ability, and poor conflict resolution, all of which are

features of many psychopathologies. Further evidence for neurocognitive abilities being an aetiological feature of psychopathology comes from recent longitudinal research that has found that executive functioning deficits were typically present prior to the development of psychopathology in adolescence, and that deficits in executive functioning predicted change in psychopathological symptoms over the following two years (Romer & Pizzagalli, 2021). Given the proposed importance of neurocognitive abilities in psychopathology, there have been vast amounts of research attempting to uncover what particular neurocognitive abilities contribute to each specific diagnosis. This line of research has had little success. Even though neurocognitive deficits are common in most disorders, the nature of neurocognitive deficits within disorders is extensively heterogeneous (Haywood & Baughman, 2021). For example, Martino et al. (2008) and Raffard and Bayard (2012) found extensively heterogeneous combinations of neurocognitive deficits within their samples of people diagnosed with bipolar disorder and schizophrenia, respectively. Furthermore, even when comparing different disorders, particular neurocognitive deficits cannot differentiate diagnoses (Moritz et al., 2002).

One possible reason for the extensive heterogeneity of the associations between neurocognitive performance and psychopathology is the predominant use of the traditional nosological approach to diagnosis (Haywood et al., 2021c). Traditional nosological approaches to the diagnosis of mental disorder, which use tools such as the DSM, have resulted in high levels of comorbidity and poor diagnostic stability (e.g., Hovenkamp-Hermelink et al., 2016; Newman et al., 1998), making the study of any single psychopathological disorder difficult (Haywood et al., 2021c). Further, the overlapping symptoms present between different disorders, as well as the ability for two people to be diagnosed with the same disorder and having no, or very few, common symptoms (Fried, 2021; Fried et al., 2020), means that finding particular collections of neurocognitive deficits fundamental to any particular disorder is unlikely (Haywood, et al., 2021a; Haywood & Baughman, 2021; Haywood et al., 2021c).

In recent years, to mitigate the issues of comorbidity and diagnostic stability of the traditional nosological approach, there have been calls to move towards dimensional approaches to describing and explaining psychopathology (Kotov et al., 2017). Rather than classifying

collections of symptoms into categories known as diagnoses, dimensional approaches typically assess symptoms and organise them into dimensional structures of psychopathology using factor analytic approaches (Kotov et al., 2017; Kotov et al., 2021). Many such structural models of psychopathology exist. Two structures, the correlated factors model and the bifactor model, have gained the most interest within the literature. The correlated factors model contains a range of symptoms, serving as indicators, and a smaller collection of specific factors (such as internalising, externalising, and thought disorder) that account for the common variance of closely related symptoms. The bifactor model contains the same fundamental components as the correlated factors model but also incorporates a single higher-order general factor (called the *p*-factor) that has been claimed to represent general psychopathology or the propensity toward all psychopathological symptoms (Caspi et al., 2014). Other common structures include the single-factor model that incorporates the symptom indicators and the p-factor, but no specific factors (Caspi et al., 2014).

As structural models of psychopathology do not create diagnostic categories and in-stead measure dimensionally, it has been suggested that structural models will improve our ability to find more reliable patterns of risk factors and outcomes associated with psychopathology (Haywood et al., 2021a; Haywood et al., 2021c; Kotov et al., 2021). While it is important to note that there is a lack of consensus on the substantive interpretation of the factors of psychopathology, in particular the p-factor, and their applications to subgroups of a population (for further detail see Fried et al., 2021; Haywood et al., 2021a; Haywood et al., 2021b), structural models of psychopathology offer a useful framework for examining the associations between neurocognitive abilities and psychopathology (Haywood et al., 2021c). Previously, we suggested that it may be possible to find, at a population level, patterns of association between neurocognition and the factors of psychopathology that help explain the differentiation between the factors. That is, factors from structural models such as internalising, externalising, and thought disorder might have discrete patterns of neurocognitive ability associations that differentiate the factors (Haywood et al., 2021c). This finding may provide insight into what neurocognitive abilities are particularly salient for the common variance of collections of

different types of symptoms. This knowledge may then inform the starting point for assessment and treatment decisions on the individual level and direct longitudinal work exploring the specific neurocognitive risk factors for collections of psychopathological symptoms. However, there has been a lack of detailed examination of the associations between specific neurocognitive processes and psychopathology factors from different types of structural models. Previous studies of the association between neurocognitive abilities and structural models of psychopathology have typically only reported the bivariate correlations between the factors of psychopathology and neurocognitive tasks and composite scores (Caspi et al., 2014) or used a single neurocognitive ability score (Martel et al., 2017). Further, other work has modelled neurocognitive abilities as a factor within the structural models of psychopathology (Eadeh et al., 2021; Haywood et al., 2021a). Modelling neurocognition within models of psychopathology, while having several unique strengths (see Haywood et al., 2021a), does not allow for the exploration of the patterns of association between discreet neurocognitive abilities and the different factors of psychopathology. For example, we called for the use of S-1 bifactor models, with neurocognitive abilities modelled as the general factor, to explore neurocognitive abilities associated with psychopathology (Haywood et al., 2021a). We describe how S-1 bifactor models, with neurocognition modelled as the general factor, offer the unique opportunity of mitigating the issue of the un-known substantive meaning of the p-factor. However, we also described how the S-1 bi-factor approach is limited as neurocognitive abilities may only be explored as a single factor, and therefore the sole use of this approach means it is not possible to examine the associations of particular neurocognitive abilities with each of the factors of psychopathology [18], an area that is particularly lacking within the literature. Due to this limitation of the S-1 approach, and the complementary information that may be obtained, we suggested that using other modelling approaches (e.g., the correlated factors model) to examine neurocognition in psychopathology remains important. Ultimately, to provide a starting point for assessment and treatment decisions, as well as to inform the future assessment of neurocognitive risk factors of psychopathology, it is important to gain a de-tailed understanding of the specific relations between various neurocognitive abilities and the different factors of psychopathology found in

the literature. Yet, to date, the degree to which psychopathology factors may differ regarding patterns of neurocognitive ability associations is unknown.

In this paper, our first aim was to (1) develop and test the fit of three of the most prominent models of psychopathology within a community sample; (a) the correlated factors model, (b) the bifactor model, and (c) the single factor model, using dimensional symptom measures. Our second aim (2) was to explore the degree to which tasks measuring four prominent neurocognitive components, (a) working memory, (b) shifting, (c) inhibition, and (d) speed of processing, are associated with, and can account for, the fac-tors of psychopathology regarding each model.

4.2 Materials and Methods

4.2.1 Participants

Through Prolific, we collected data online from a representative community sample (based on simplified census data on age, gender, and ethnicity) of 425 participants in the USA. Exclusion criteria were (1) any condition or injury that could impact their motor movements, thereby interfering with the participants' ability to complete the cognitive tasks, and (2) colour blindness or colour perception issues. This study was approved by the Curtin University Human Research Ethics Committee (HRE2021-0105).

4.2.2 Procedure

After providing consent, participants provided demographic information (i.e., age, gender), psychiatric history information (i.e., diagnosis, psychiatric hospital admissions), and information of any psychotropic medication use. Participants then completed measures of substance use (ASSIST V3.1; WHO, 2002), psychiatric symptoms (the Brief Symptom Inventory; Derogatis & Melisaratos, 1983), and completed eight neurocognitive tasks. Participants were instructed to complete the survey and tasks in an environment as free from distractions as possible. Online neurocognitive data collection does not allow for controlling the participants' testing environment or hardware. However, a myriad of research supports the validity and quality of online, crowd-sourced, neurocognitive data and has found participants' performance comparable to laboratory-based studies (Crump et al., 2013; Johnson et al., 2021;

Merz et al., 2020; Sauter et al., 2020). Furthermore, Prolific recently has been shown to obtain behavioural task data practically indistinguishable from in-person lab testing, far outperforming similar crowd-sourcing platforms concerning quality and comparability (Uittenhove et al., 2022).

4.2.3 Materials

4.2.3.1. Substance Use and Symptomology

To measure substance use and personal, social, and legal issues related to that use, we used the Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST) V3.1 (WHO, 2002). The ASSIST assesses the use of tobacco products, alcoholic beverages, cannabis, cocaine, amphetamine-type stimulants, inhalants, sedatives or sleeping pills, hallucinogens, opioids, and other substances. Each participant indicates frequency of use, desire or urge to use, and frequency of health, social, legal, and financial issues related to the use of each substance used within the last three months. The ASSIST generates a substance involvement score for each substance assessed. The ASSIST has shown strong reliability and validity in general community samples (WHO, 2002).

Psychiatric symptoms were measured via the Brief Symptom Inventory-53 (BSI) [27]. The BSI is a psychiatric symptom measure that assesses symptoms over the past seven days, and is valid and reliable in both clinical and community samples (Akhavan Abiri & Shairi, 2020). The measure is comprised of 53-items, measured on a five-point Likert-type scale, that jointly assess nine symptom dimensions based on the original factor structure. The nine symptom dimensions are somatisation, obsession-compulsion, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism.

4.2.3.2. Neurocognitive Abilities

Participants completed eight neurocognitive tasks presented in randomised order. The performance metric for each task that involved both speed of response and accuracy was calculated using the Rate-correct Score (RCS) method (Vandierendonck, 2017), in which the number of correct responses is divided by total reaction time (in milliseconds) to provide a metric number of correct responses per millisecond. For the tasks that did not require a speed of

response aspect, we used only a metric of accuracy. All tasks were developed in JavaScript. The details of each task are provided below.

4.2.3.2.1. Working Memory

Working memory was assessed via (1) a digit span task, and (2) a visual working memory task designed after the visual array task described in Cowan et al. (2006). In the digit span task, participants were presented with number sequences, in which each number remained on screen for 1000 ms. Following the presentation of the last number in each sequence, participants were prompted to enter the previously shown sequence, in order, using an on-screen keypad and their PC mouse. The task was designed so that trials increased in difficulty, starting with a 3-digit sequence and progressing to the most challenging 15-digit sequence. Sequences across trials either increased by one digit for each correct response or decreased in length by one digit for every two consecutive errors made. Participants completed 12 trials of the digit span task, with a maximum digit span possible of 15. The outcome variable used was the maximum digit span across the 12 trials.

In the visual memory task, participants completed 84 trials showing sets of either 4, 6, 8, or 10 coloured dots. The initial presentation of dots remained on-screen for 300ms, followed by a brief interstimulus interval of 1000 ms before a second set of dots was presented. Participants were instructed to indicate whether a circled dot in the second presentation was different in colour to the initial presentation. Performance was assessed via accuracy on the number of correct responses across all 84 trials.

4.2.3.2.2. Shifting

Shifting was assessed using a Shape-Number switching task and the Inferring Relevance shifting task (Wilson & Niv, 2012). The Shape-Number task was adapted from the Letter-Number task (Kimberg et al., 2000). It consisted of participants completing 96 trials in which, following familiarization blocks, they were required to respond to either the number (i.e., 2 vs. 3 dots) or shape (i.e., square or diamond) as stimuli were presented in a 2 × 2 grid. Stimuli appeared sequentially and in a clockwise pattern, and participants used either the Z or M key on their keyboard to respond. For stimuli that appeared in the top row, participants responded based

on their shape. When stimuli appeared on the bottom row, participants responded based on the number of dots. Our outcome variable for the Shape-Number task was the number of correct responses divided by total reaction time. Our outcome variable for the Shape-Number task was the number of correct responses divided by total reaction time.

The Inferring Relevance task (Wilson & Niv, 2012) was derived from the Wisconsin Card Sorting task (WCST; Heaton et al., 1981) and the Intra-Dimensional/Extra-Dimensional Shifts task (Mackintosh, 1965). This task required that participants use their PC mouse to select one of three different on-screen stimuli, depending on what they believed to be the dimension-to-match, in a given trial. Participants completed 200 trials whereby the dimension-to-match was either 'shape' (i.e., squares, triangles, and circles), 'colour' (i.e., shapes outlines were either red, green, or yellow), or 'pattern' (i.e., within each shape was either grid lines, dots, or waves). The correct dimension-to-match changed after 15–25 consecutive trials of one dimension. As per the WCST, correctly identifying the dimension-to-match occurs initially via trial and error and feedback presentation. However, to increase task difficulty, participants' certainty of response was interfered with by providing incorrect feedback on 25% of trials (Wilson & Niv, 2012). The primary outcome measure was the number of correct responses divided by total reaction time.

4.2.3.2.3. Inhibition

Inhibition was assessed via computerized versions of (1) the Stroop Task [41] and (2) the Go/NoGo task (Nosek & Banaji, 2001). The Stroop task comprised a total of 48 trials, with 16 trials each for neutral (four "X"s appeared in one of three colours: blue, red, or green), congruent (words "BLUE", "RED", or "GREEN" appeared in colours that matched the meaning of the word presented; i.e., the word "BLUE" appeared in the colour blue), and incongruent (words "BLUE", "RED", or "GREEN" appeared in colours that did not match the meaning of the word presented; i.e., the word "RED" appeared in the colour green) conditions. In all trials, participants were required to indicate the colour of letters presented on screen, using their mouse to select one of three corresponding buttons on-screen ("Blue", "Red", or "Green"). Participants were asked to select, as quickly as possible, the box that corresponded to the colour of the text presented. Therefore, in the incongruent condition, participants had to inhibit selecting the box that

corresponded to the text rather than the colour (MacLeod, 1992). The primary outcome variable was the number of correct responses for congruent stimuli divided by the total reaction time for those stimuli, subtracted from the number of correct responses for incongruent stimuli divided by the total reaction time for those stimuli.

The Go/NoGo task (Nosek & Banaji, 2001) used consisted of 120 trials. One of two stimuli, either an "M" or a "W", was presented on-screen, and participants were instructed to press the space bar as quickly as possible when presented with the "M" (the "Go" stimuli) but not to press the space bar when presented with the "W" (the "NoGo" stimuli). Out of the 120 stimuli, the "Go" stimuli accounted for 80%, while the "NoGo" accounted for 20%. This weighting of "Go and "NoGo" stimuli has been shown to provide adequate variability of errors (Wilson et al., 2016). Stimuli were presented between 1000 ms and 1550 ms apart, and participants were given 1200 ms to respond. As 80% of the stimuli were "Go" stimuli, when presented with a "NoGo" stimuli, participants were required to actively inhibit the predominant response of pressing the space bar. Our primary outcome variable of the Go/NoGo was the number of correct NoGo omissions divided by the total reaction time of responses.

4.2.3.2.4. Speed of Processing

Speed of processing was assessed via two tasks, (1) a simple reaction time task and (2) the Inspection Time (IT) task (Anderson et al., 2001). For the simple reaction time task, participants were instructed to respond to the on-screen presentation of a blue "circle" by pressing the space bar on their keyboard as quickly as possible. Participants completed a total of 40 trials, with each trial separated by an interval of between 1000 ms and 1750 ms (this was to avoid participants preempting responses). The outcome variable for the simple reaction time task was the number of correct responses divided by the total reaction time.

On the IT task, participants were presented with images depicting an alien with two antennae (Anderson et al., 2001). Four variations of this stimulus were used, showing (1) both short antennae, (2) both long antennae, (3) the left antennae being longer than the right antennae, and (4) the right antennae being longer than the left. The exposure duration of stimuli was manipulated so that stimuli were tested at 4 ms increments, between 6 ms and 62 ms a total of 4

times each, thus comprising a total of 60 trials. After each presentation, a mask was presented on screen, and participants were required to indicate whether the previously shown antennae were the same (via pressing the "Z" key) or were different (via pressing "M") in length. Our outcome variable was $a + b$ where: a = the lowest exposure duration for two consecutive blocks where accuracy was at 75% or higher. And b = a growing sum of exposure duration blocks with greater than 75% accuracy, divided by the number of blocks over 75%. Lower scores, therefore, reflected better performance.

4.2.1 Analysis

The analysis of this data occurred in multiple steps. Step one was to confirm the factor structure of the Brief Symptom Inventory (BSI). We used confirmatory factor analysis (CFA) to test two structures of the BSI from the literature. (1) the original Derogatis and Melisaratos (1983) nine-factor/49-item structure, and (2) a more recent six-factor, 40-item structure found by Schwannauer and Chetwynd (2007). Step two of the analysis was used to create the subscale scores for the choice of BSI factor structure, and to examine the bivariate correlations between the demographic, BSI, and ASSIST variables. Step two of the analysis used exploratory factor analysis (EFA), among the BSI subscales and the ASSIST variables, to support the development of the models of psychopathology. Step three consisted of choosing the specific psychopathology factors by devising correlated factors models based on the EFA and conceptual interpretation and using CFA to test the models' fit. In step four, the four structural models of neurocognition were tested; a correlated factors model, two versions of a bifactor model, and a single-factor model. Finally, step five consisted of assessing partial bivariate correlations (accounting for covariates) between the neurocognition and the factors of psychopathology, and a multivariate multiple regression analysis to examine the degree to which the participants' performance on the neurocognitive tasks could account for the factors of psychopathology after accounting for covariates and the common variance of the tasks.

For our CFIs, we applied less stringent rules of thumb to indicate a good fitting model and used these rules in combination with conceptual interpretation when choosing a model from alternatives. This approach was taken due to a smaller sample size with an initial large number of

observed variables (Shi et al., 2018) and Greene et al. (2021) suggesting the use of conceptual interpretation and the minimisation of the reliance on fit measures when choosing models. For our CFAs, an RMSEA of <0.05 indicated a good fit, <0.08 indicated a reasonably good fit, and <0.10 indicated a mediocre fit (Browne & Cudeck, 1992). An SRMR of <0.09 indicated a good fit (Hu & Bentler, 1999), while the earlier convention of the TFI and the CFI of =>0.9 was used to indicate a good fit, rather than using the later convention of =>0.95 due to the tendency of a =>0.95 cut off to over reject true-population models with smaller sample sizes (<N = 500; Hu & Bentler, 1995; Hu & Bentler, 1999). All factor loadings were required to be significant at the alpha level of <0.05. For all models, the MLR estimator with robust test statistics was used.

Regarding the EFA, an oblique (GeominQ) rotation was used, with an ML estimator, and a model was chosen from alternatives based on information derived from the EFAs, as well as the subsequent CFAs, in combination with theoretical and conceptual interpretation. Therefore, models were chosen based upon an exploratory-confirmatory continuum (Greene et al., 2021), incorporating the importance of conceptual interpretation of the models.

4.3. Results

After cleaning the data, 25 of the 425 participants were removed due to incomplete data for one or more of the neurocognitive tasks, leaving a final sample of N = 400. The demographic and clinical variables for our final sample can be found in Table 4.1.

Table 4.1

Participant Characteristics.

Variable		Mean (SD/%)/Count	Min	Max
Age		44.47 (16.35)	18	83
Gender				
	Male	194 (48.5%)	-	-
	Female	206 (51.5%)	-	-
Diagnosis (Yes/No)				
	Yes	114 (28.5%)	-	-
	No	286 (71.5%)	-	-
Diagnoses [a]				
	Depression	66 (16.5%)	-	-
	Generalised Anxiety	57 (14.2%)	-	-
	Agoraphobia	2 (0.5%)	-	-
	Social Anxiety	7 (1.8%)	-	-
	Panic Disorder	4 (1.0%)	-	-
	Schizoaffective	1 (0.3%)	-	-

Variable		Mean (SD/%)/Count	Min	Max
	Psychosis	2 (0.5%)	-	-
	Eating Disorder	1 (0.3%)	-	-
	Cyclothymia	1 (0.3%)	-	-
	Bipolar	17 (4.3%)	-	-
	OCD	3 (0.8%)	-	-
	Impulse Control	1 (0.3%)	-	-
	BPD	3 (0.8%)	-	-
	PTSD	19 (4.8%)	-	-
	Substance Use	3 (0.8%)	-	-
	Trichotillomania	1 (0.3%)	-	-
Year of First Diagnosis		2007.69 (10.71)	1980	2021
Admitted to a Mental Health Facility (Yes/No)				
	Yes	26 (6.5%)	-	-
	No	374 (93.5%)	-	
Year of First Admission		2003.31 (13.14)	1980	2020
Using Psychotropic Medication (Yes/No)				
	Yes	60 (15.0%)	-	-
	No	340 (85.0%)	-	-

4.3.1 Step-One

First, we confirmed the structure of the BSI by testing the original nine-factor model [27] and the newer six-factor model (Schwannauer & Chetwynd, 2007). The original, nine-factor, BSI structure (Derogatis & Melisaratos, 1983) did not fit the data well, with the CFI and the TLI not meeting the criteria for a good fit (χ^2 (1091, N = 400) = 2350.07, CFI = 0.867, TLI = 0.857, SRMR = 0.064, RMSEA = 0.066, 90% CI = [0.051, 0.056]), and had multiple non-positive definite identification issues. This suggested that there were multiple redundant items within the factor structure. However, the Schwannauer and Chetwynd (2007) six-factor structure provided a "reasonably good fit", with regards to RMSEA, a "good fit" regarding the SRMR, and bordering on a good fit for the TFI and the CFI (χ^2 (725, N = 400) = 1518.41, CFI = 0.891, TLI = 0.885, SRMR = 0.058, RMSEA = 0.064, 90% CI = [0.049, 0.055]). The six-factor structure also had no identification issues and very good-to-excellent internal consistency (Cronbach's Alpha's ranging from 0.858 to 0.940). Therefore, we concluded that, overall, the six-factor structure provided an adequate fit for the data while offering clearly conceptually interpretable factors. The six-factors, with names devised from examining the contents of each factor, their associated BSI item numbers, the original factor they were placed within the nine-factor solution, and their

Cronbach's Alpha's, can be found in Table 4.2. Although the factor named "mental fog" contained only items from the BSI aimed at measuring distress related to obsessive-compulsive symptoms, it was named as such due to the subset of items retained reflecting perceived mental performance in daily life, just one aspect of the obsessive-compulsive phenotype. Example items included "Having to check and double-check what you do" and "Your mind going blank" (Derogatis & Melisaratos, 1983).

Table 4.2

The Six-Factor Model

Factor Name	Item Numbers	Original Factor	Cronbach's Alpha
Depression			0.940
	17	Depression	
	18	Depression	
	16	Depression	
	14	Psychoticism	
	35	Depression	
	50	Depression	
	44	Anxiety	
Agoraphobia			0.865
	8	Phobic Anxiety	
	43	Phobic Anxiety	
	28	Phobic Anxiety	
	31	Phobic Anxiety	
	45	Anxiety	
Hostility			0.832
	13	Hostility	
	46	Hostility	
	41	Hostility	
	40	Hostility	
	6	Hostility	
Mental Fog			0.909
	36	Obsessive-Compulsive	
	5	Obsessive-Compulsive	
	26	Obsessive-Compulsive	
	32	Obsessive-Compulsive	
	27	Obsessive-Compulsive	
	15	Obsessive-Compulsive	
Interpersonal Anxiety			0.904
	21	Interpersonal Sensitivity	
	22	Interpersonal Sensitivity	
	51	Paranoid Ideation	
	20	Interpersonal Sensitivity	
	42	Interpersonal Sensitivity	
	48	Somatisation	

Factor Name	Item Numbers	Original Factor	Cronbach's Alpha
	24	Paranoid Ideation	
	4	Paranoid Ideation	
	10	Paranoid Ideation	
Somatisation			0.858
	7	Somatisation	
	30	Somatisation	
	33	Somatisation	
	29	Somatisation	
	23	Somatisation	
	2	Somatisation	
	37	Somatisation	
	1	Anxiety	

4.3.2 Step-Two

Following the choice of the six-factor BSI solution, scores for each of the six factors were created from the relevant BSI items using the original scoring procedure. An "other substances" ASSIST variable was also created by adding together scores from the cocaine, amphetamine, inhalants, sedatives, and hallucinogens categories, as there was little variation within these substances. The combination of less commonly used substances is standard amongst the literature (e.g., Caspi et al., 2014). The bivariate associations between the six BSI and the four ASSIST variables were explored. This was done to test for the appropriateness of using each BSI and ASSIST variable in developing our models of psychopathology. The bivariate correlations can be found in Table 4.3. All BSI and ASSIST variables had significant positive correlations, except for tobacco, which was only significantly associated with one of the six BSI variables (Somatisation). As the development of structural models of psychopathology is grounded in significant positive associations between the variables, tobacco use was not included in the development of the models of psychopathology or any other subsequent analyses.

Table 4.3

Symptom and Substance Bivariate Correlations

	Dep	Agor	Host	Fog	Inter. Anx	Somat	Tob	Alc	Cann	Other
Depression	1	0.627 **	0.620 **	0.761 **	0.808 **	0.690 **	0.040	0.132 **	0.262 **	0.185 **
Agoraphobia	0.627 **	1	0.534 **	0.621 **	0.667 **	0.718 **	0.070	0.133 **	0.244 **	0.211 **
Hostility	0.620 **	0.534 **	1	0.668 **	0.698 **	0.655 **	0.022	0.206 **	0.193 **	0.227 **
Mental Fog	0.761 **	0.621 **	0.668 **	1	0.765 **	0.745 **	0.070	0.138 **	0.275 **	0.223 **
Inter. Anxiety	0.808 **	0.667**	0.698 **	0.765 **	1	0.723 **	0.076	0.150 **	00.257 **	0.233 **
Somatisation	0.690 **	0.718 **	0.655 **	0.745 **	0.723 **	1	0.124 *	0.195 **	0.302 **	0.262 **
Tobacco	0.040	0.070	0.022	0.070	0.076	0.124 *	1	0.303 **	0.257 **	0.279 **
Alcohol	0.132 **	0.133 **	0.206 **	0.138 **	00.150 **	0.195 **	0.303 **	1	0.279 **	0.201 **
Cannabis	0.262 **	0.244 **	0.193 **	0.275 **	0.257 **	0.302 **	0.257 **	0.279 **	1	0.349 **
Other Drugs	0.185 **	0.211 **	0.227 **	0.223 **	0.233 **	0.262 **	0.279 **	0.201 **	0.349 **	1

Note. *. Correlation is significant at the 0.05 level (two-tailed). **. Correlation is significant at the 0.01 level (two-tailed). Dep = Depression. Agor = Agoraphobia. Host = Hostility. Fog = Mental Fog. Inter. Anx = Interpersonal Anxiety. Somat = Somatisation. Tob = Tobacco. Alc = Alcohol. Cann = Cannabis. Other = Other Substances

4.3.2 Step-Three

In step three, we used EFA to inform the development of the specific, second-order factors of psychopathology. Given we had nine observed variables, six BSI variables, and three ASSIST variables, we started by examining a four-factor structure, which is the largest structure with the possibility of at least two observed variables loading onto each factor. The EFAs can be found in table 4.4. For the four-factor EFA, a factor emerged consisting of depression, mental fog, and interpersonal anxiety. This factor also showed a cross-loading between factor two for agoraphobia. Furthermore, a second factor emerged consisting of the cross-loaded agoraphobia variable and somatisation, and a third factor consisting of a single loading > 0.3 in hostility. Finally, a fourth factor emerged consisting of the three substance use variables. Next, we assessed a 3-factor structure. The three-factor structure revealed similar results when compared to the four-factor structure. A factor still emerged consisting of depression, hostility, mental fog, and interpersonal anxiety, but now also included hostility, which was moved from its own factor. Factor two emerged still consisting of somatisation and the agoraphobia cross-loading with factor 1. The third factor contained the three substance use variables. Finally, we tested a two-factor solution. The two-factor solution consisted of a factor accounting for the BSI items and for the ASSIST items. The results of the EFAs are presented in Table 4.4.

Table 4.4

EFA Factor Loadings

Number of Factors	Item	Factor 1	Factor 2	Factor 3	Factor 4
4					
	Depression	0.936			
	Agoraphobia	0.398	0.418		
	Hostility			0.960	
	Mental Fog	0.617			
	Interpersonal Anxiety	0.797			
	Somatisation		0.975		
	Alcohol				0.376

Number of Factors	Item	Factor 1	Factor 2	Factor 3	Factor 4
	Cannabis				0.712
	Other Substances				0.450
3					
	Depression	0.903			-
	Agoraphobia	0.396	0.400		-
	Hostility	0.604			-
	Mental Fog	0.699			-
	Interpersonal Anxiety	0.931			-
	Somatisation		0.981		-
	Alcohol			0.401	-
	Cannabis			0.685	-
	Other Substances			0.489	-
2					
	Depression	0.900		-	-
	Agoraphobia	0.711		-	-
	Hostility	0.739		-	-
	Mental Fog	0.862		-	-
	Interpersonal Anxiety	0.920		-	-
	Somatisation	0.772		-	-
	Alcohol		0.429	-	-
	Cannabis		0.595	-	-
	Other Substances		0.524	-	-

Note. Factor loadings < 0.3 are hidden.

The three and two-factor models provided the most parsimonious solutions, and were chosen to be further explored using CFAs. First, we tested two different three-factor models and two different two-factor models. The first three-factor model tested (a), following the exact structure as the three-factor EFA, and loading agoraphobia onto factor two, due to its slightly stronger loading, and its conceptual relationship to somatisation. The second three-factor CFA tested (b) was the same as the first. However, hostility was loaded onto factor three, with the substance use variables. All of the three factors were allowed to correlate. The first two-factor solution tested (c) was derived directly from the two-factor EFA, but the second two-factor model tested (d), like the three-factor model (b), had hostility loaded on as a factor with the substance use variables. We tested the alternative two and three-factor models for two reasons; the four-factor solution showed hostility loading on a separate factor, not on factor one, and hostility or conduct issues are primarily conceptualized with substance use as an "externalising" factor within the literature (e.g., Caspi et al., 2014). Furthermore, regarding the alternative two-factor solution, by having hostility loading onto a factor with substance use, we tested a model with "Internalising" and "Externalising" factors. These factors have been repeatedly validated and received a great amount of interest throughout the literature (Caspi et al., 2014; Caspi & Moffitt, 2018; Haywood et al., 2021b, 2021c; Lahey et al., 2012). For all models, the factors were allowed to correlate.

The two three-factor solutions showed to be a "reasonably good" and "mediocre" fit, respectively. The first model (a) (χ^2 (24, N = 400) = 34.00, CFI = 0.993, TLI = 0.989, SRMR = 0.020, RMSEA = 0.032, 90% CI = [0.000, 0.052]), with hostility loaded onto factor one provided a marginally better fit than the second model (b) (χ^2 (24, N = 400) = 73.34, CFI = 0.963, TLI = 0.945, SRMR = 0.057, RMSEA = 0.072, 90% CI = [0.057, 0.087]), with hostility loaded onto factor three with the substance use variables.

Next, we tested the fit of the two variations of the two-factor model. The first two-factor model tested (c), with hostility loading onto factor one, was a "reasonable" fit for the data with regards to the RMSEA, and a good fit for the CFI, TLI, and SRMR (χ^2 (26, N = 400) = 96.73, CFI = 0.975, TLI = 0.966, SRMR = 0.027, RMSEA = 0.057, 90% CI = [0.042, 0.073]). The

alternative two-factor solution tested (d) was a “mediocre”-to-“reasonable” fit for the data with regards to the RMSEA, and a good fit for the CFI, TLI, and the SRMR (χ^2 (26, N = 400) = 96.73, CFI = 0.942, TLI = 0.926, SRMR = 0.059, RMSEA = 0.081, 90% CI = [0.068, 0.097]).

Given that all of the four CFAs tested provided a fit for the data, each model may have been acceptable to select. However, given that a two-factor “Internalising” and “Externalising” model fitted the data and that there is a large amount of conceptual and empirical evidence supporting the use of these factors, we selected this model as our correlated factors model (Caspi et al., 2014; Caspi & Moffitt, 2018; Haywood et al., 2021b, 2021c; Lahey et al., 2012).

Next, after developing the choice of the correlated-factors model, we tested the fit of two different bifactor models. Each model tested consisted of the same observed variables and the same specific factors (Internalising and Externalising) as in the correlated factors model, but included a higher-order *p*-factor. Each of the nine observed variables loaded onto the *p*-factor as well as either Internalising or Externalising. What differentiated the models was whether the specific factors were allowed to correlate. In the first model tested (a), the specific factors were not allowed to correlate, but in the second model (b), the specific factors were allowed to correlate. We tested both of these versions of the bifactor model as previous research has applied both types successfully (Caspi et al., 2014; Watts et al., 2019)

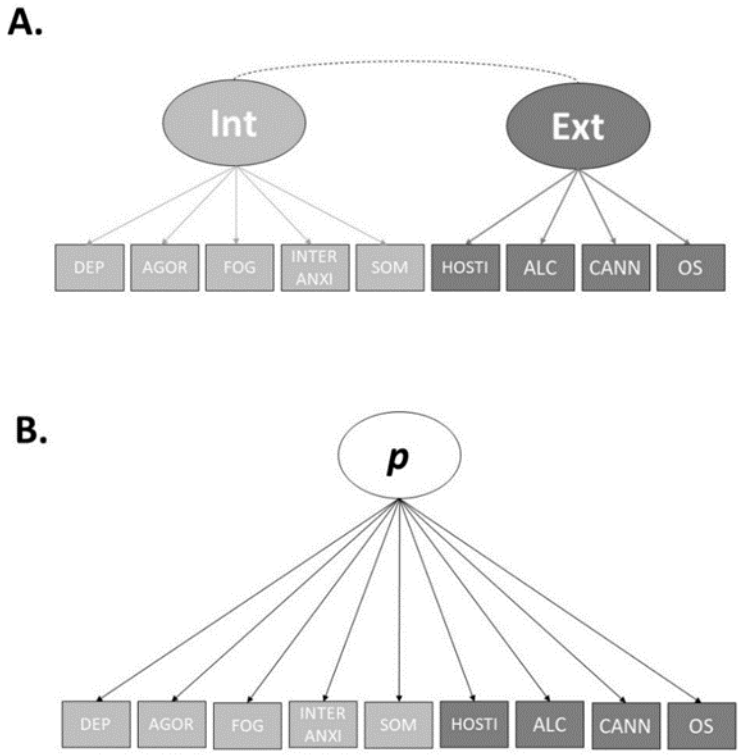
The first bifactor model tested (a), without correlated specific factors fit the data well (χ^2 (18, N = 400) = 23.82, CFI = 0.996, TLI = 0.992, SRMR = 0.020, RMSEA = 0.029, 90% CI = [0.000, 0.053]). However, none of the three observed variables retained significant loadings on the Internalising specific factor, and hostility did not retain its significant loading on the Externalising factor. Finally, there was also a Heywood case, an observed variable with negative variance (somatisation). These findings are thought to be due to the higher-order *p*-factor subsumed the Internalising specific factor, as well as the variance in hostility accounted for by the Externalising factor. The second bifactor model tested (b), that contained correlated specific factors, also fit the data well (χ^2 (17, N = 400) = 21.18, CFI = 0.997, TLI = 0.993, SRMR = 0.017, RMSEA = 0.025, 90% CI = [0.000, 0.051]). However, the second model (b) shared many of the same issues as the first (a). For model two (b), none of the observed variables retained

significant loadings on Internalising. Hostility also did not retain its significant loading on Externalising. Furthermore, somatisation was also a Heywood case within this model. Overall, for both bifactor models, the *p*-factor subsumes the Internalizing factor. A specific factor being subsumed is relatively common in bifactor models of psychopathology, and previous research has removed the factor subsumed (Caspi et al., 2014). However, this is now known to be poor practice, as if the subsumed factor is removed, the *p*-factor becomes defined by that removed factor, changing its interpretation (Eid, 2020; Haywood et al., 2021a). Therefore, the bifactor model is not appropriate to explore further within this data. The results do, however, suggest a single-factor model may be a good fit for the data.

Lastly, we tested the fit of the single-factor model of psychopathology within our sample. The single-factor model consists of the same nine observed variables used in the other models, however, containing one higher-order *p*-factor and no specific factors. The single factor provided a "mediocre"-to-"reasonably" good fit for with regards to the RMSEA, and a good fit for the CFI, TLI, and SRMR (χ^2 (27, N = 400) = 98.12, CFI = 0.946, TLI = 0.928, SRMR = 0.062, RMSEA = 0.081, 90% CI = [0.067, 0.095]). All of the nine-observed variables loaded significantly of the *p*-factor. Therefore, we decided to use the (A) correlated factors model and (B) the single-factor model for our examination of the utility of neurocognitive abilities in accounting for the factors of psychopathology. Figure 4.1 displays two final models.

Figure 4.1

Final Structural Models of Psychopathology.



Note. Depiction of the two final structural models used models used in this chapter. Panel A: Correlated factors model. Panel B: Single-factor model. DEP = Depression. AGOR = Agoraphobia, FOG = Mental Fog, INTER ANXI = Interpersonal Anxiety. SOM = Somatisation, HOSTI = Hostility. ALC = Alcohol. CANN = Cannabis. OS = Other Substances.

The factor loadings for both the final correlated factors model and the single-factor model can be found in Table 4.5. As specified by Caspi et al. (2014), we standardised the *p*-factor scores to a mean of 100 and a standard deviation of 15. The internalising and externalising factors were mildly-to-moderately correlated (r = 0.743), while the correlations between the *p*-factor in the single factor model and specific factors in the correlated factors model were strong (*p* and Internalising, r = 0.996; *p* and Externalising, r = 0.799). The *p*-factor and Internalising correlated almost perfectly, indicating the *p*-factor in the single-factor model largely represented Internalising symptoms.

Table 4.5

CFA for the Final Two Models

Model	Factor	Depr	Agor	Fog	Int. Anx.	Soma	Host	Alc	Cann	Other	Int~Ext
Correlated Factors											0.743 **
	Internalising	0.862	0.750	0.867	0.896	0.842					
	Externalising						0.806	0.227	0.328	0.300	
Single-Factor											
	p	0.861	0.750	0.867	0.896	0.842	0.758	0.192	0.316	0.272	

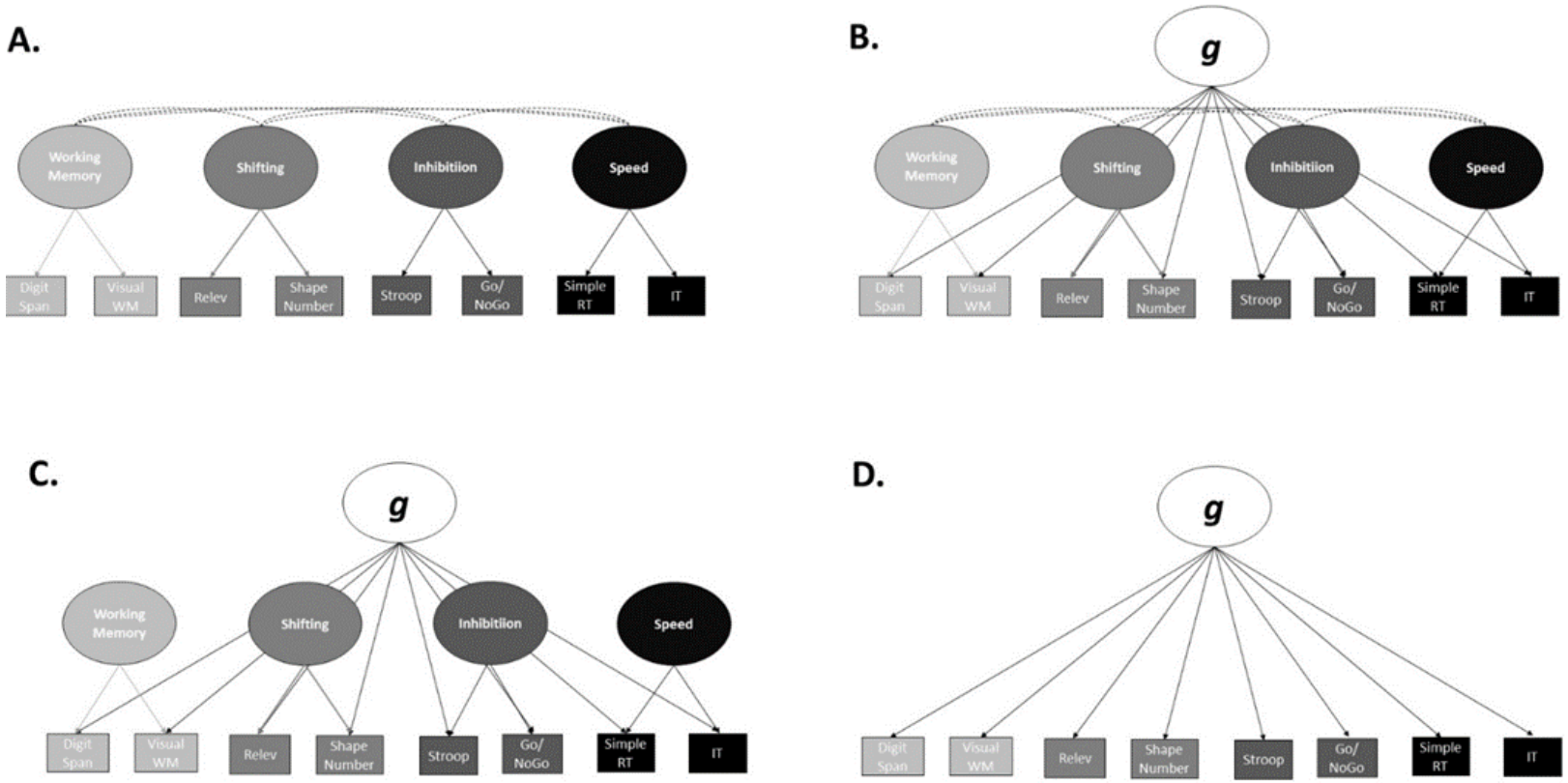
Note. Depr = Depression. Agor = Agoraphobia. Fog = Mental fog. Int. Anx. = Interpersonal Anxiety. Soma = Somatisation. Host = Hostility. Alc = Alcohol. Cann = Cannabis Other = Other Substances. Int = Internalising. Ext = Externalising. ~ = correlation.

3.3.3 Step-Four

After The choice of structural models of psychopathology, we examined the fit of three different structural models of neurocognition; (a) a correlated factors model, (b) a bifactor model with correlated specific factors, (c) a bifactor model without correlated specific factors, and (d) a single factor model. Figure 4.2 depicts the four models. Unlike our approach to developing the models of psychopathology, we did not precede the confirmatory with exploratory factor analyses. This is because, unlike the components from our measure of psychopathology, we actively chose two specific tasks to measure each theoretically driven neurocognitive component. Therefore, it would be inappropriate to conduct exploratory factor analyses as any alternative structures would forgo the conceptual interpretation and theoretical foundations of the neurocognitive components.

Figure 4.2

Neurocognitive Structural Models Tested



Note. Depiction of the neurocognitive structural models tested. Panel A: Correlated factors model. Panel B: bifactor model with specific factors. Panel C: bifactor model with higher-order factors. Panel D: single-factor model. Rele = Inferring Relevance task. WM = Working Memory. RT = Reaction Time. IT = Inspection Time.

All of the four tested models failed to converge. To explore if the different units of measurement between our data (e.g., RCS units Vs number of trails correct) caused the failure of convergence, we normalised the data and attempted to re-fit the models. All models, apart from the single-factor model, still failed to converge. The single-factor model, although it did converge with the normalised data, did not provide a good fit for data ($\chi 2$= 40.152, CFI = .876, TLI = .826, RMSEA = .05, 90% CI = [.027, .073]. SRMR = .05), and three (Digit Span, Simple Reaction Time, and IT) out of the eight variables did not significantly load onto the general factor This may be expected based on the generally low correlations amongst the neurocognitive components. These results suggested it would be most appropriate to examine each neurocognitive test independently within our remaining analyses. Descriptive statistics for the neurocognitive tests are presented in Table 4.6.

Table 4.6

Neurocognitive Task Descriptive Statistics.

	Minimum	Maximum	Mean	Std. Deviation
Digit Span	4	14	7.82	1.79
Visual WM	27	74	57.09	8.76
Inferring Relevance	0.0002761	0.0019141	0.0008937	0.0002978
Shape-Number	0.0000799	0.0016144	0.0006979	0.0002565
Stroop	−0.0002305	0.0008982	0.0002402	0.0001444
Go/NoGo	0.0001582	0.0005083	0.0003417	0.0000592
Simple RT	0.0002084	0.0003973	0.0003921	0.0000206
IT	28.67	112.00	67.53	22.08

Note. Values presented for the Inferring Relevance, Shaper-Number, Stroop, Go/NoGo, and Simple Reaction Time relate to the number of correct responses per millisecond.

4.3.4 Step-Five

In Step Five we examined the partial (controlling for age and gender) correlations between the neurocognitive tasks and internalising, externalising and the *p*-factor. We controlled for age and gender as both demographic variables were significantly associated with one or more of the factors of psychopathology. Higher age being associated with lower internalising,

externalising and *p*-factor scores (internalising, $r = -0.422$, $p < 0.001$; externalising, $r = -0.348$, $p < 0.001$; *p*-factor, $r = -0.424$, $p < 0.001$), and females (males = 1, females = 2) tended to have higher scores on internalising and the *p*-factor each factor (internalising, $r = 0.201$, $p < 0.001$; externalising, $r = 0.006$, $p = 0.910$; *p*-factor, $r = 0.182$, $p < 0.001$). Table 4.7. shows the bivariate correlations between the neurocognitive tasks and the factors of psychopathology after accounting for age and gender

Table 4.7

Partial Bivariate Correlations Between Neurocognition and Psychopathology.

Control Variables		Digit Span	Vis WM	Infer. Rel.	Shape-Num	Stroop	Go/No Go	Simple RT	IT	Int	Ext	*p*
	Digit Span	1.000	0.045	0.020	0.020	−0.020	0.050	0.060	−0.042	−0.060	0.038	−0.048
	Visual WM	0.045	1.000	0.187 **	0.160 **	0.064	0.122 *	0.203 **	−0.051	−0.053	−0.066	−0.056
	Infer. Rel.	0.020	0.187 **	1.000	0.359 **	0.177 **	0.077	0.014	0.003	−0.016	−0.005	−0.015
	Shape-Number	0.020	0.160 **	0.359 **	1.000	0.100*	0.107*	0.078	0.021	−0.016	0.010	−0.013
	Stroop	−0.020	0.064	0.177 **	0.100 *	1.000	0.013	−0.017	−0.044	−0.032	0.031	−0.024
Age & Gender	Go/NoGo	0.050	0.122 *	0.077	0.107 *	0.013	1.000	0.223 **	−0.048	−0.088	−0.061	−0.087
	Simple RT	0.060	0.203 **	0.014	0.078	−0.017	0.223 **	1.000	−0.023	−0.130 **	−0.226 **	−0.148 **
	IT	−0.042	−0.051	0.003	0.021	−0.044	−0.048	−0.023	1.000	0.112 *	0.089	0.113 *
	Internalising	−0.060	−0.053	−0.016	−0.016	−0.032	−0.088	−0.130 **	0.112 *	1.000	0.719 **	0.995 **
	Externalising	0.038	−0.066	−0.005	0.010	0.031	−0.061	−0.226 **	0.089	0.719 **	1.000	0.782 **
	p-Factor	−0.048	−0.056	−0.015	−0.013	−0.024	−0.087	−0.148 **	0.113 *	0.995 **	0.782 **	1.000

Note. **. Correlation is significant at the 0.01 level (two-tailed). *. Correlation is significant at the 0.05 level (two-tailed). Vis WM = Visual Working Memory. Infer. Rel. = Inferring Relevance. Shape-Num = Shape-Number RT = Reaction Time. IT = Inspection Time. Int = Internalising. Ext = Externalising. *p* = *p*-factor

Of the eight neurocognitive tasks, after accounting for age and gender, only the two tasks designed to measure the speed of processing were significantly associated with one or more of the factors of psychopathology. Specifically, performance on the simple reaction time task was significantly negatively associated with internalising, externalising, and the *p*-factor. This finding indicates that better performance on the simple reaction time task is significantly associated with lower internalising and externalising symptoms, as well as the *p*-factor score. The Inspection Time task was significantly positively associated with internalising and the *p*-factor, indicating that better performance on the Inspection Time task was associated with lower internalising symptoms and lower *p*-factor scores. Combined, these results indicate that within our data, speed of processing is the primary neurocognitive correlate with higher-order psychopathology.

Next, we used a multivariate multiple regression analysis to examine the degree to which each neurocognitive task could account for unique variance in the psychopathology factors, accounting for age and gender, as well as the common variance amongst the tasks. The model accounted for a significant 23.8% of variance in internalising ($F(10, 389) = 12.17$, $p < 0.001$, $R^2 = 0.238$), a significant 15.6% of variance in externalising ($F(10, 389) = 8.37$, $p < 0.001$, $R^2 = 0.156$), and a significant 23.6% of variance in the *p*-factor ($F(10, 389) = 12.05$, $p < 0.001$, $R^2 = 0.236$). Table 4.8 provides the results of the regression analysis.

Table 4.8

Multivariate Multiple Regression Analysis

Predictors	Internalising					Externalising					*p*-Factor				
	B	β	*p*	*Partial*	*Sr* ^{2}	*B*	β	*p*	*Partial*	*Sr* ^{2}	*B*	β	*p*	*Partial*	*Sr* ^{2}
Age	−0.027	−0.433	<0.001 **	−0.404	0.148	−0.026	−0.346	<0.001 **	−0.321	0.095	−0.398	−0.434	<0.001 **	−0.404	0.149
Gender	0.360	0.174	<0.001 **	0.188	0.028	0.007	0.003	0.951	0.003	<0.001	4.68	0.156	0.001 **	0.170	0.023
Digit Span	−0.024	−0.041	0.360	−0.046	0.002	0.036	0.054	0.249	0.058	0.003	−0.251	−0.030	0.505	−0.034	0.001
Vis WM	−0.002	−0.013	0.783	−0.014	0.001	−0.003	−0.022	0.669	−0.022	<0.001	−0.026	−0.015	0.758	−0.016	<0.001
Infer. Rel.	13.03	0.003	0.950	0.003	<0.001	149.99	0.032	0.555	0.030	0.001	415.57	0.007	0.891	0.006	<0.001
Shape-Num	−9.37	−0.003	0.958	−0.003	<0.001	−61.02	−0.015	0.778	−0.014	<0.001	−214.33	−0.004	0.934	−0.004	<0.001
Stroop	−182.03	−0.025	0.578	−0.028	0.001	264.25	0.031	0.664	0.034	0.001	−1916.14	−0.018	0.686	−0.020	<0.001
Go/NoGo	−840.70	−0.048	0.297	−0.053	0.002	−203.95	−0.010	0.835	−0.011	<0.001	−11229.51	−0.044	0.336	−0.049	0.002
Simple RT	−4973.28	−0.099	0.034 *	−0.107	0.009	−12291.50	−0.209	<0.001 **	−0.214	0.040	−84882.76	−0.117	0.012 *	−0.126	0.012
IT	0.004	0.094	0.040 *	0.104	0.008	0.005	0.082	0.083	0.088	0.006	0.064	0.095	0.037 *	0.105	0.009

Note. *. is significant at the 0.05 level. **. is significant at the 0.01 level. Vis WM = Visual Working Memory. Infer. Rel. = Inferring Relevance. Shape-Num = Shape-Number RT = Reaction Time. IT = Inspection Time.

Regarding internalising, the simple reaction time task and the Inspection Time task remained significant predictors after accounting for the variance of age and gender, as well as the common variance of the neurocognitive tasks. Simple reaction time performance uniquely accounted for 0.9%, and the inspection time task accounted for 0.8% of the variance in internalising, respectively. This indicates that our tasks assessing the speed of processing are not only significantly associated with internalising after accounting for age and gender but can also account for a significant amount of unique variance in internalising after accounting for age and gender in addition to the common variance from the neurocognitive tasks. However, it is important to acknowledge that combined the unique variance in internalising accounted for by the speed of processing tasks was just 1.7%.

Regarding externalising, simple reaction time performance was a significant predictor of externalising in our model after accounting for age, gender, and the common variance of the remaining neurocognitive tasks. Simple reaction time accounted for a significant 4.0% of unique variance in externalising that could not be explained by age and gender or the common variance of the remaining neurocognitive tasks. However, unlike internalising, our other measure of the speed of processing, the inspection time task, did not account for a significant amount of unique variance in externalising.

Even though the internalising and the *p*-factor were highly correlated, to ensure a full investigation of the study aims and psychopathology factors, it was still important to examine the relations between neurocognitive performance and the general factor. Further, as the internalising and *p*-factor are highly, but not perfectly, correlated, the analyses remained important. Both simple reaction time and inspection time task performance accounted for a significant amount of unique variance in the *p*-factor over and above age, gender, and the common variance of the neurocognitive tasks. Simple reaction time performance accounted for a significant 1.2% of unique variance in the *p*-factor, while performance on the inspection time task accounted for a significant 0.9% of unique variance in the *p*-factor. Overall, our findings suggest that the tasks measuring the speed of processing were the most efficacious when compared to tasks measuring

working memory, shifting, and inhibition, in accounting for higher-order psychopathology within our sample.

4.4 Discussion

Strong evidence suggests that deficits in neurocognitive abilities play a role in the aetiology of psychopathology (e.g., Romer & Pizzagalli, 2021). In recent years, evidence has grown for the utility of dimensional structural models of psychopathology as an alternative to traditional nosological diagnostic approaches. However, there is a lack of understanding of how neurocognitive abilities are associated with factors of psychopathology derived from structural models. The aim of this paper was to (1) develop and test the fit of three popular models of psychopathology within a community sample; (a) the correlated factors model, (b) the bifactor model, and (c) the single factor model, using dimensional symptom measures. Our second aim (2) was to explore the degree to which tasks measuring four prominent neurocognitive components, (a) working memory, (b) shifting, (c) inhibition, and (d) speed of processing, are associated with, and can account for, the factors of psychopathology from each model.

Within our sample, only the correlated-factors model and the single factors model fit our data well. The correlated factors model consisted of an internalising factor and an externalising factor. The internalising factor had loadings from depression, agoraphobia, mental fog, interpersonal anxiety, and somatisation, while the externalising factor had loadings from hostility, alcohol use, cannabis use, and other drug use. Our correlated factors model parallels many other correlated factors models found within the literature (e.g., Lahey et al., 2012). However, we did not find a third factor, namely “thought disorder”, that is commonly found within the literature. Thought disorder is commonly defined by psychotic symptoms (Caspi et al., 2014), and the absence of thought disorder factors from our models may be explained by the use of the six-factor BSI model over the original nine-factor model. The original nine-factor BSI model, which included a psychoticism factor, did not fit our data well, so the alternative six-factor model, which only included a single psychosis item amongst its factors, was used. Therefore, our six observed variables used to develop the models did not include strong indicators of psychoticism. The bifactor models fit our data well, although they had several non-significant factor loadings

and a Heywood case. The good fit of the bifactor models is not surprising given that fit indices bias bifactor models over correlated factors models (Greene et al., 2019). However, the single-factor model also fit the data well. The *p*-factor from the single factor model, however, was almost perfectly correlated with the internalising factor from the correlated factors model. This suggests, along with the Heywood cases, that the bifactor models were a poor structure for the data because the *p*-factor primarily represented internalising. The issue of the *p*-factor being malleable and primarily representing a specific factor has been discussed previously in the literature (e.g., Haywood et al., 2021a; Haywood et al., 2021b; Haywood et al., 2021c; Heinrich et al., 2020). This represents a limitation of developing an understanding of the substantive meaning of *p.*

We also attempted to fit different structural models of neurocognitive abilities. None of the models fit our data. This was unexpected as previous research generally finds similar models to be a good fit (e.g., Karr et al., 2018). However, we chose neurocognitive tasks that assessed different aspects of each neurocognitive domain. For example, to measure the speed of processing, we used a simple reaction time task that required participants to respond to a stimulus as quickly and as accurately as possible, as well as an Inspection Time task that did not involve any response speed but instead involved high-speed image processing. Furthermore, to measure shifting, we used a more traditional switching task, the shape-number task, that required participants to switch mental set in response to a known, defined rule (top or bottom of the grid), as well as a less traditional Inferring Relevance switching task (Wilson & Niv, 2012), that required participants to switch mental set in a probabilistic, more real-world, context. Therefore, given that we measured the breadth of each neurocognitive domain, it is understandable that performance heterogeneity resulted in the models of neurocognition not being a good fit.

We found that, after controlling for age and gender, simple reaction time performance was significantly associated with the internalising and externalising factors concerning the correlated factors model, as well as the *p*-factor regarding the single factors model. We also found additionally, after controlling for age and gender, the IT task performance was significantly associated with internalising and the *p*-factor. However, tasks that measure working memory,

shifting, and inhibition were not significantly associated with any of the factors of psychopathology. Furthermore, after accounting for age, gender, and the common variance of the neurocognitive tasks, the simple reaction time task accounted for a significant 0.9%, 4.0%, and 1.2% of the variance in internalising, externalising, and the *p*-factor, respectively. After accounting for age, gender, and the common variance of the neurocognitive tasks, the IT task accounted for a unique 0.8% and 0.9% variance in internalising and the *p*-factor. This suggests that, in our data, tasks that measured speed of processing had the greatest predictive utility, although limited to a combined predictive utility of 1.7%. The lack of predictive utility of working memory, shifting, and inhibition tasks regarding the factors of psychopathology both conflicts and supports findings from the limited research in this area. Caspi et al. (2014) found working memory to be significantly associated with internalising and externalising in their correlated factors model as well as to the *p*-factor in their bifactor model. However, paralleling our findings, Caspi et al. (2014)also found that a shifting task (i.e., the Trail-Making-Test-B) was not significantly associated with externalising within their correlated factors model but did find it was significantly associated with internalising. Finally, our findings also parallel previous findings (Caspi et al., 2014) in that speed of processing was significantly associated with both internalising and externalising, as well as the *p*-factor. Previously, we suggested that factors from structural models such as internalising, externalising, and the *p*-factor may have discrete patterns of neurocognitive ability associations that differentiate the factors (Haywood et al., 2021c). However, our results suggest that internalising, externalising, and the *p*-factor may not be clearly differentiated by neurocognitive performance, and that processing speed is a common correlate.

The importance of processing speed has been primarily studied in relation to ageing (Albinet et al., 2012; Kail, 1991). However, there has been growing interest in the role speed processing in psychopathology plays in internalising and externalising disorders and symptoms. For example, a recent systematic review has found that people with major depressive disorder typically have processing speed deficits, and provided evidence that, to compensate for this deficit, people with major depressive disorder are required to use greater cognitive effort to perform daily tasks (Nuño et al., 2021). Nuño et al. (2021) also suggest that if a task requires a

high cognitive demand, deficits in the speed of processing cannot be compensated for by higher cognitive effort, and therefore task performance is poor. Deficits in the speed of processing in depression, therefore, have been suggested to negatively impact occupational performance (Nuño et al., 2021), and this may perpetuate depressive symptoms. Further evidence for the importance of speed of processing in psychopathology is that there is evidence for speed of processing being a reliable cognitive endophenotype for bipolar disorder, with not only people with bipolar disorder experiencing speed of processing deficits, but also significant proportions of relatives of those with bipolar disorder experiencing deficits in speed of processing (Daban et al., 2012). Furthermore, it has been found that an intervention designed to train speed of processing in the elderly resulted in a significant reduction in the risk of experiencing depressive symptoms 1 and 5 years post-intervention, while training in perceptual reasoning, or working memory, had no impact (Wolinsky et al., 2009). Regarding externalising behaviours, there is a reliable association between alcohol use disorder and speed of information processing deficits (Galandra et al., 2021; Paolillo et al., 2019), and speed of processing may not only be a consequence of externalising behaviours but may also be involved in the aetiology of those behaviours. Durazzo et al. (2008) found that processing speed deficits significantly predicted relapse in people treated for alcohol dependence after accounting for demographic, psychiatric, metabolic, and clinical covariates. Furthermore, there is evidence that deficits in speed of processing are related to an earlier onset of conduct disorder (Johnson et al., 2015). Our findings, therefore, parallel previous research proposing the importance of speed of processing within psychopathology. However, our results extend the literature by employing dimensional structures of psychopathology in a representative community sample, showing that deficits in speed of processing are not only related to nosologically defined disorders but also statistically derived dimensions in the general population.

4.4.1 Limitations of the Research and Directions for Future Research

The data for this study was collected online through Prolific. Therefore, we had little experimental control over the context in which participants completed the survey and tasks and the devices and settings used. However, the range of evidence suggests that the quality of task

data collected online through crowd-sourcing platforms, such as prolific, is comparable to in-lab studies (Crump et al., 2013; Johnson et al., 2021; Merz et al., 2020; Sauter et al., 2020), and we used the most valid crowd-sourcing platform for this context (Uittenhove et al., 2022). Further, we only explored neurocognitive ability's associations with the factors of psychopathology on the sample and not the individual level. Previously we provided evidence for compensatory neurocognitive profiles on the individual level that can explain the heterogeneous findings between specific neurocognitive abilities and psychopathology (Haywood & Baughman, 2021; Haywood et al., 2021c). Even though at the sample level, measures of working memory, inhibition, and shifting were not significantly associated with the factors of psychopathology, at the individual level, explanatory heterogeneous profiles of neurocognitive performance may exist. For example, there are two individuals with the same high *p*-factor score of 140. Individual One may have a pervasive deficit in working memory while having good shifting, inhibition, and speed of processing ability. However, Individual Two may have a good working memory, shifting, and inhibition ability, but a pervasive deficit in speed of processing. For each individual, their neurocognitive strength and weakness profile may explain their high level of *p*, however, on the sample level (N = 2), no associations would exist between any neurocognitive ability and their level of *p* due to the heterogeneity.

Future research should validate our findings in a laboratory setting to limit potential confounding variables. Future research should also examine the associations between neurocognitive abilities and psychopathology factors on the individual level, exploring potential compensatory neurocognitive profiles and dynamic multidimensional explanations.

4.5 Conclusions

In this paper, we explored the associations between neurocognitive abilities and structural models of psychopathology. We found a correlated factors model and single factor model to best fit our psychopathology data. We found tasks measuring speed of processing had the most predictive utility for internalising, externalising, and the *p*-factor. Specifically, poorer performance on the simple reaction time task was significantly associated with higher scores of internalising, externalising, and the *p*-factor, and poorer performance on the Inspection Time

Task was significantly associated with higher scores of internalising and the *p*-factor. Tasks that measured working memory, shifting, and inhibition were not significantly associated with psychopathology factors. We found neurocognitive abilities were not differentially associated, but that speed of processing was a common correlate of psychopathology factors.

Preface to Chapter 5

Chapter five investigates the fourth and final direction of future research presented in chapter one. The fourth direction of future research coincides with the fourth aim of this thesis. This chapter builds upon the previous chapter by directly comparing the linear models developed in chapter four to our alternative dynamic multidimensional conceptualisation and interprets the findings in relation to each of the previous chapters. Specifically, the following chapter examines if the factors of psychopathology (e.g., internalising, externalising, and the *p*-factor) are usefully explained by non-linear multidimensional interactions between neurocognitive components, and if this non-linear multidimensional conceptualisation is superior to the traditional linear conceptualisation of neurocognitive abilities at predicting both (a) lower-level and (b) higher-level psychopathology. In chapter five we use the same data as in chapter four. This chapter ultimately provides a critical comparison of the traditional linear versus non-linear multidimensional interactive conceptualisations of neurocognition and their utility in enhancing our understanding of psychopathology.

Note: The following chapter has been published in Brain Sciences in the special issue titled Diagnosis and Advances in Research on Human Behavior.

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2022). Neurocognitive Artificial Neural Network Models Are Superior to Linear Models at Accounting for Dimensional Psychopathology. *Brain Sciences*, *12*(8), 1060. https://doi.org/10.3390/brainsci12081060

Minor edits have been made to the following chapter, such as to phrasing and Australian Spelling, to ensure consistency within the thesis. This research is supported by an Australian Government Research Training
Program (RTP) Scholarship.

Chapter 5 (Study 4): Neurocognitive Artificial Neural Network Models Are Superior to Linear Models at Accounting for Dimensional Psychopathology

5.1 Introduction

It has been contended that deficits in neurocognitive abilities are an aetiological feature of psychopathology(e.g., Beck & Rector, 2005; Romer & Pizzagalli, 2021). Not only do those with psychopathology typically have neurocognitive deficits, but these deficits in neurocognitive performance are seen to precede the development of psychopathology (Romer & Pizzagalli, 2021). However, few, if any, deficits to underlying neurocognitive abilities appear to be deterministic in the study of psychopathology. Thus, the search for one-to-one correspondence between deficits and disorders has yielded little knowledge that can be constantly applied. Instead, evidence suggests that the associations between neurocognitive abilities and psychopathology are extensively heterogeneous (e.g., Carruthers, Gurvich, et al., 2019; Carruthers, Van Rheenen, et al., 2019; Haywood & Baughman, 2021; Malcolm et al., 2021; Martino et al., 2008; Moritz et al., 2002; Tan et al., 2021). For example, previous research has found that for people with bipolar disorder approximately 22% displayed deficits in three to four neurocognitive components, 40% showed deficits in one or two components, and 38% did not display any deficits (Martino et al., 2008). Of note is the fact that no consistent deficit could be isolated, in any single neurocognitive component. Multi-disorder research corroborates these findings as there is no evidence for specific, single neurocognitive deficits that reliably discriminate disorders(e.g., Moritz et al., 2002).

One possible explanation for the extensive heterogeneity of the associations between neurocognitive abilities and psychopathology may stem from the use of traditional nosological approaches to diagnosis. Traditional approaches to diagnosis, using tools such as the DSM, have resulted in extensive comorbidity, and poor diagnostic stability (Hovenkamp-Hermelink et al., 2016; Newman et al., 1998). The high levels of comorbidity and poor diagnostic stability, along with emerging aetiological evidence (Burdick et al., 2006; Craddock et al., 2006; Lichtenstein et al., 2009; Smucny et al., 2018), suggests that psychopathology may not be best represented as discrete diagnostic categories (Cuthbert, 2020; Kotov et al., 2021), and if this is the case, finding

associations between any particular neurocognitive component and a specific diagnostic category is unlikely. In recent years, the issues with the traditional nosological approach to diagnosis has led to the development of a range of dimensional, statistical models of psychopathology (Caspi et al., 2014; Lahey et al., 2012; Lahey et al., 2021b). These models of psychopathology do not categorise disorder, but rather represent symptoms dimensionally on a collection of higher-order statistically derived components of psychopathology. The models of psychopathology that have gained the most interest are the correlated factors model, the bifactor model, and the single-factor model. The correlated factors model contains a range of lower-level symptom indicators, such as depression, anxiety, and hostility, and two or more higher-level correlated dimensional factors, such as internalising and externalising (Caspi & Moffitt, 2018). The bifactor model is similar in structure, however it includes a single factor, named the *p*-factor, at the highest level of the structure that also receives its loadings from the lower-level symptom indicators (Caspi et al., 2014). The single-factor model contains the same lower-level symptom indicators, but only the higher-level *p*-factor (Caspi et al., 2014). Previously, we suggested that by using these dimensional statistical models it may be possible to find clear specific associations between different neurocognitive abilities and the factors of psychopathology that discriminate the factors (Haywood et al., 2021b). However, more recently, only a common deficit in speed of processing was found to be related to higher scores of internalising, externalising and *p*, providing evidence that there may not be discrete neurocognitive associations among the factors (Haywood et al., 2022).

The challenges faced with isolating neurocognitive deficits associated with specific disorders, and the issues concerning the heterogeneity of behavioural symptoms in disorders, point to the need for approaches that can examine the effects of dynamical interplay in the underlying processes. Towards this end, we recently used computational models of the Wisconsin Card Sorting Task (WCST) to explore an alternative conceptualisation of the relation between neurocognitive abilities and psychopathology, termed the multidimensional hypothesis (Haywood & Baughman, 2021; Haywood et al., 2021c). Ultimately, we claimed that to understand the functional associations between neurocognition and psychopathology consideration of the non-

linear interactions between neurocognitive components is of critical importance. This conceptualisation was inspired by the notion of Multiple Realisation from the study of the philosophy of mind; referring to the idea that any given state may be equally determined, or realised, by a number of different causes (Putnam, 1988). Rather than attempting to test which single process or ability related most to a given disorder, we tested the combined effects of various profiles of neurocognitive abilities on models' performance on the WCST. The results of this work revealed that a range of manipulations to the processes pertaining to neurocognitive abilities of updating, shifting, and inhibition were equivalent in simulating performance on the WCST in people with schizophrenia, their healthy-first degree relatives, and controls. These findings, we argue, highlight the advantages of using dynamic multidimensional approaches in the study of psychopathology (Haywood et al., 2021c).

Artificial neural networks (ANN) have been successfully used across levels, from psychological to genetic to help understand psychopathological and behavioural phenomena (Baughman & Thomas, 2008; Bosia et al., 2019; Dolce et al., 2020; Haywood, Lawrence, et al., 2021; Simeoli et al., 2021; Thomas et al., 2009; Tryon et al., 2017; Wei et al., 2022) and potentially provides an even richer methodology for studying the relations between neurocognitive abilities and psychopathology. For example, in a standard 3-layer feed-forward network, for which a problem may be specified and for which the desired outcome is known, input units are provided with a representation of the problem and the output of each unit is fed forward to all units it is connected to within a hidden layer (comprised of a number of processing units). The hidden units in turn feed forward to the output layer that represents the solution. Throughout the model, each connection partially determines (via its strength of connection, or weight) the final value, and the degree of error in the model's solution is then used to alter weights within the model with the goal of achieving a more accurate outcome on the next cycle. Traditionally, research in the domain of neurocognition and psychopathology has relied on linear explanations, often using popular correlational techniques. For example, multiple linear regression allows one to determine the unique and common contributions for any number of independent variables on an outcome. Whilst multiple linear regression is particularly accessible

and easy to perform, using any of a variety of modern statistical analysis packages, it allows only for additive combinations of a linear form (Vetter & Schober, 2018). The issue with this is that for many psychological phenomena, it appears that rather than contributing linearly to such outcomes, that instead factors interact in more dynamic ways in producing effects.

Towards the other end of the complexity spectrum, in terms of analytical techniques, are machine learning techniques, or artificial neural networks (ANN). The potential advantage of such approaches is that they allow highly complex, non-linear patterns of relations to be found between any number of variables and an outcome. However, studies examining the difference between standard analytical techniques, such as MLR, and machine learning approaches are lacking. Knowing what the potential benefits are, of one approach over another, offers clear advantages for elucidating the true role factors play in influencing specific outcomes, and the degree to which dynamic multidimensionality holds. The central objective of this study is to compare multiple linear regression models (MLR) to artificial neural network models (ANN), in order to determine the degree to which each are able to predict specific psychopathological outcomes. In each instance, to facilitate comparison, the models we develop represents the more accessible of techniques that exist with the respective approaches.

5.1.2 Aims and Hypotheses

The aim of this research is to compare the accuracy of linear models versus non-linear artificial neural network models with regard to how well they each predict (a) lower-level and (b) higher-level psychopathology.

Hypothesis One. The average correlations between the actual lower-level psychopathology scores and the models' predicted psychopathology scores will be significantly stronger for the ANN model when compared to the linear model.

Hypothesis Two. The correlations between actual and model predicted (a) internalising, (b) externalising, and (c) general psychopathology (the *p*-factor) scores will be significantly stronger for the ANN models when compared to the linear models.

5.2 Materials and Methods

5.2.1 Participants

In a large-scale study (Haywood et al., 2022), 425 people from a representative community sample from the USA were recruited through Prolific (Palan & Schitter, 2018). Participants completed a demographics and clinical characteristics survey (Derogatis & Melisaratos, 1983), a substance use measure (WHO, 2002), and eight neurocognitive tasks. After data cleaning 400 participants were retained. The mean age of the sample was 44.47 (SD = 16.35), 51.5% were female, and 28.5% reported having a previous or current mental health diagnosis. The detailed demographic and clinical characteristics of the sample can be found in Haywood et al. (2022).

5.2.2 Procedure

After providing consent, participants completed the demographic and clinical characteristics questions, and then completed measures on substance use, mental health symptomology, and then each of the eight neurocognitive tasks (see Haywood et al., 2022 for further information) for further information. This research was approved by the Curtin Human Research Ethics Committee (HRE2021-0105).

5.2.3 Measures

In this study, we used a subset of variables collected in the larger study (Haywood et al., 2022). We used structural models of psychopathology developed in the larger study derived from data collected using the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) V3.1 (WHO, 2002), and the 53 item Brief Symptom Inventory (BSI-53) (Derogatis & Melisaratos, 1983). The ASSIST is the gold-standard measure for substance involvement across tobacco products, alcoholic beverages, cannabis, cocaine, amphetamine-type stimulants, inhalants, sedatives or sleeping pills, hallucinogens, opioids, and other substances (WHO, 2002). The BSI is a 53-item psychiatric symptom measure that assesses degree of distress associated with a wide-range of psychiatric symptoms over the previous seven days (Derogatis & Melisaratos, 1983).

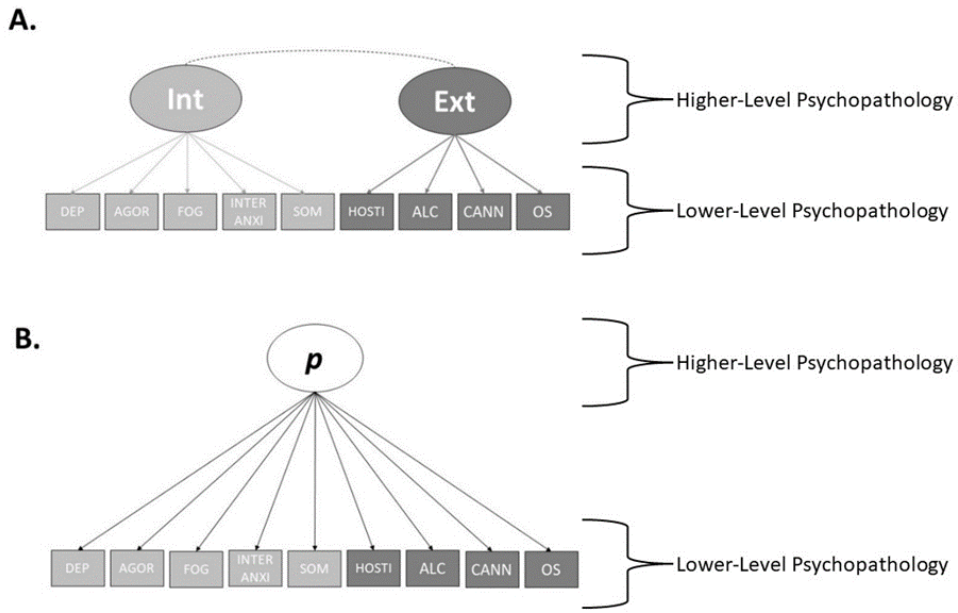
Data from eight computerised neurocognitive tasks were also collected. To measure working memory we used the Digit Span task, and a visual array task based on Cowen (2016). To measure shifting we used the Shape-Number task, based on the Letter-Number task (Kimberg et al., 2000), and the Inferring Relevance Task (Wilson & Niv, 2012). To measure inhibition, we used a computerised version of the Stroop task (MacLeod, 1992) and the Go/NoGo task (Nosek & Banaji, 2001). Lastly, to measure speed of processing we used the Simple Reaction Time task, and the Inspection Time (IT) task (Anderson et al., 2001). The Rate-Corrected Score (RCS) method was used for tasks that required both speed and accuracy to measure performance. Haywood et al. (2022) provides further detail on the tasks used and the metrics assessed.

5.2.3 Analysis

In this study, we used structural models of psychopathology that had previously been developed (Haywood et al., 2022). These structural models were developed and tested through confirmatory factor analysis in line with structural and hierarchical conceptual interpretations of psychopathology (e.g., Caspi et al., 2014). These models used a six-factor BSI model (Schwannauer & Chetwynd, 2007), with the six domains being Depression, Agoraphobia, Hostility, Mental Fog, Interpersonal Anxiety, and Somatisation, and three domains of substance use derived from the ASSIST V3.1, namely alcohol use, cannabis use, and other substance use. These nine domains were included as ‘lower-level’ indicators in our structural models. Regarding the structural models, we used the correlated factors model, with internalising and externalising specific factors, and the single factors model, developed previously (Haywood et al., 2022). However, the bifactor model was not used as it had a Heywood case (a variable with a negative variance estimate). The models included the BSI domains, derived from the Schwannauer and Chetwynd (2007) factor structure, and the ASSIST components as the observed variables (see Haywood et al., 2022 for further detail). All models were developed and tested in RStudio using the MLR estimator with robust test statistics, and the final models were chosen from the alternatives based on a combination of model fit, factor loadings, and conceptual interpretation see (see Haywood et al., 2022). The final correlated-factors model and the single factor model are depicted in Figure 5.1

Figure 5.1

Correlated Factors Model and the Single Factor Model



Note. Adapted from Haywood et al. (2022), final structural models of psychopathology used in this research. Pictured is the Correlated Factors Model (A) and Bifactor Model (B). DEP = Depression. AGOR = Agoraphobia, FOG = Mental Fog, INTER ANXI = Interpersonal Anxiety. SOM = Somatisation, HOSTI = Hostility. ALC = Alcohol. CANN = Cannabis. OS = Other Substances.

Factor scores for internalising, externalising, and the *p*-factor were extracted for each participant. These scores were the 'higher-level' psychopathology variables that the linear and ANNs models were to predict to test hypothesis two. Further, we used the scores for each of the six BSI variables, and the three ASSIST variables, as the 'lower-level' psychopathology scores that the two types of models were to predict in order to test hypothesis one.

5.2.3.1. Linear Models

Multivariate multiple regression analyses were used as the linear method to predict psychopathology from neurocognitive abilities. The models included the eight neurocognitive tasks, as well as age and gender as predictors. The outcome variables for the lower-level model were the six BSI domains; Depression, Agoraphobia, Hostility, Mental Fog, Interpersonal

Anxiety, and Somatisation, and the three ASSIST variables; Alcohol, Cannabis, and other drug use. The higher-level psychopathology model included the same predictors, but the outcome variables were internalising, externalising, and p-factor scores.

5.2.3.2. Artificial Neural Network Models

Multivariate multiple regression analyses were used as the linear method to predict psychopathology from neurocognitive abilities. The models included the eight neurocognitive tasks, as well as age and gender as predictors. The outcome variables for the lower-level model were the six BSI domains; Depression, Agoraphobia, Hostility, Mental Fog, Interpersonal Anxiety, and Somatisation, and the three ASSIST variables; Alcohol, Cannabis, and other drug use. The higher-level psychopathology model included the same predictors, but the outcome variables were internalising, externalising, and *p*-factor scores.

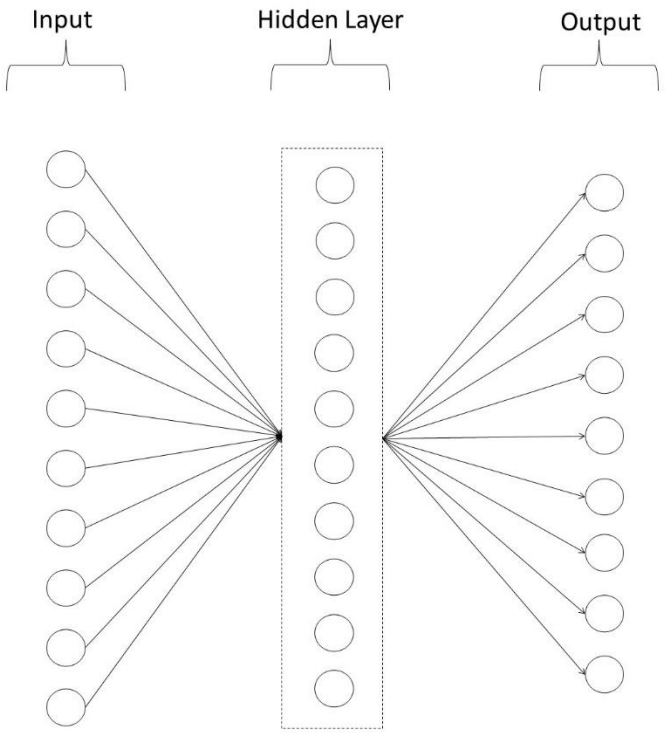
We developed two ANN models, one for lower-level psychopathology, and one for higher-level psychopathology. Both models were 3-layer feedforward connectionist networks consisting of an input layer of 10 units (representing age, gender, and performance on each of the eight cognitive tasks) a hidden layer of 10 units, and an output layer of either 9 units (lower-level psychopathology model) or 3 units (higher-level psychopathology model). In the lower-level model, the output layer comprised of 9 units, representing depression, agoraphobia, mental fog, interpersonal anxiety, somatisation, hostility, alcohol, cannabis, and other substances, while in the higher-level model the output layer consisted of 3 units, representing internalising, externalising and the *p*-factor. We used sigmoidal activation functions for units and the model was trained randomly, with replacement, on 100 of the 400 cases using back-propagation for 1000 epochs, with a learning rate of 0.03, and with the initial weights for all units randomised between ±0.5. The model was tested against the full set of 400 cases. To safeguard against possible under, or over-fitting our data, we examined the effect of varying the learning rate (0.01 to 0.5), and the number of hidden units (5, 10, 15, 20). These manipulations to the model's parameters did not alter the outcome or pattern of results, although greater differentiation was noted for some extremes. For example, by the end of training, a higher learning rate (0.5) had little effect in reducing error in the model with 5 units in the hidden layer. In contrast, in those models with

greater than 5 units in the hidden layer (i.e., 10, 15, 20) by the end of training, error was considerably smaller.

We did not explore the effects of using different activation function nor did we examine the effect of increasing the number of hidden layers in the model. These variations potentially may be of interest to us for future work. However, overall, and given the purpose of this study, to compare linear models to ANN models, the model described here offers a useful starting framework. The models were developed in MatLab. Figure 5.2 depicts the lower-level psychopathology ANN model, while Figure 5.3 depicts the higher-level psychopathology ANN model.

Figure 5.2

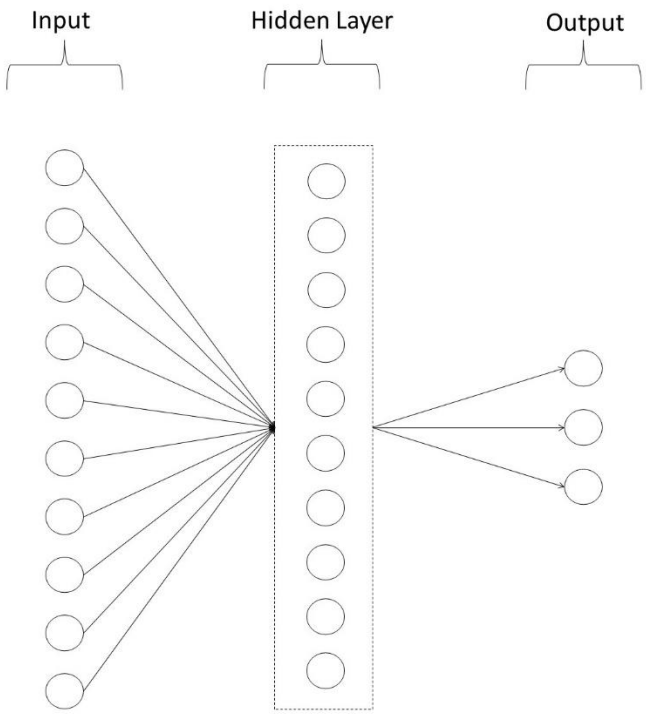
Lower-level Psychopathology Artificial Neural Network Model



Note. Depiction of the lower-level psychopathology artificial neural network model used in this research. The first layer of the model contains 10 input units consisting of age, gender, and the eight neurocognitive tasks. The second layer is the hidden layer consisting of 10 units. The final layer is the output layer consisting of nine output units, namely depression, agoraphobia, mental fog, interpersonal anxiety, somatisation, hostility, alcohol, cannabis, and other substances.

Figure 5.3

Higher-level Psychopathology Artificial Neural Network Model



Note. Depiction of the higher-level psychopathology artificial neural network model used in this research. The first layer of the model contains 10 input units consisting of age, gender, and the eight neurocognitive tasks. The second layer is the hidden layer consisting of 10 units. The final layer is the output layer consisting of three output units, namely internalising, externalising, and the *p*-factor.

5.2.3.3. Model Comparison

The predictive accuracy of the linear models and the ANN models was assessed by statistically comparing the correlations between the respective models' predicted outcome variable scores, and the actual outcome variable scores. The correlations between the models' predicted and actual scores for Depression, Agoraphobia, Hostility, Mental Fog, Interpersonal Anxiety, Somatisation, Alcohol, Cannabis, and Other Drugs were averaged to provide an overall indication of the predictive accuracy of the lower-level psychopathology models. The overall correlation for the linear and the ANN was compared using the Daniel Soper calculator (Soper,

2022), that applies a Fisher transformation (Hauke & Kossowski, 2011) to compare two correlations. Similarly, the correlations between the predicted and actual (a) internalising, (b) externalising, and (c) *p*-factor scores were statistically compared for the linear and the ANN model. Superior predictive accuracy of the ANN over the linear models at both the lower-level (BSI and ASSIST variables) and higher-level of psychopathology (internalising, externalising and *p*-factor), would evidence the existence of non-linear interactive relationships between the predictors (neurocognition, age, and gender) and the outcomes (psychopathology) (Warner & Misra, 1996).

5.3 Results

5.3.1 Linear Models

5.3.1.1. Lower-Level Psychopathology

Regarding lower-level psychopathology, the linear model with the predictor variables of the eight neurocognitive variables and age and gender were able to account for a significant amount of variance in each of the nine symptom domains. The model accounted for 18.9% of depression ($F(10, 389) = 9.08$, $p < 0.001$, $R^2 = 0.189$), 10.9% of agoraphobia ($F(10, 389) = 4.77$, $p < 0.001$, $R^2 = 0.109$), 17.1% of hostility ($F(10, 389) = 8.03$, $p < 0.001$, $R^2 = 0.171$), 22.7% of Mental Fog ($F(10, 389) = 11.40$, $p < 0.001$, $R^2 = 0.227$), 20.1% of interpersonal anxiety ($F(10, 389) = 9.80$, $p < 0.001$, $R^2 = 0.201$), 22.0% of somatisation ($F(10, 389) = 10.98$, $p < 0.001$, $R^2 = 0.220$), 7.2% of alcohol use ($F(10, 389) = 3.00$, $p = 0.001$, $R^2 = 0.072$), 5.5% of cannabis use ($F(10, 389) = 2.28$, $p = 0.013$, $R^2 = 0.055$), and 5.9% of other substance use ($F(10, 389) = 2.43$, $p = 0.008$, $R^2 = 0.059$). Table 5.1 provides the utility of the individual predictors in the model.

Table 5.1

Lower-Level Psychopathology Linear Model Outcomes

Dependent Variable	Parameter	B	B	Sig.
Depression	Age	−0.024	−0.396	<0.001 **
	Gender	0.275	0.137	0.007 **
	Digit Span	−0.051	−0.091	0.067
	Visual WM	0.002	0.014	0.790
	Inferring Relevance	117.25	0.030	0.598
	Shape-Number	155.76	0.046	0.410
	Stroop	−337.60	−0.049	0.334

Dependent Variable	Parameter	B	B	Sig.
	Go/NoGo	−512.58	−0.030	0.550
	Simple RT	−2866.43	−0.059	0.250
	IT	0.003	0.068	0.177
Agoraphobia	Age	−0.015	−0.241	<0.001 **
	Gender	0.221	0.111	0.007 **
	Digit Span	−0.002	−0.004	0.913
	Visual WM	−0.002	−0.018	0.673
	Inferring Relevance	−65.82	−0.017	0.714
	Shape-Number	−84.80	−0.025	0.578
	Stroop	100.15	0.014	0.722
	Go/NoGo	−709.85	−0.042	0.305
	Simple RT	−3372.44	−0.070	0.094
	IT	0.002	0.038	0.342
Hostility	Age	−0.015	−0.238	<0.001 **
	Gender	0.036	0.018	0.566
	Digit Span	0.018	0.032	0.306
	Visual WM	0.000	−0.004	0.905
	Inferring Relevance	61.31	0.016	0.659
	Shape-Number	−39.83	−0.012	0.736
	Stroop	96.77	0.014	0.657
	Go/NoGo	−293.64	−0.017	0.584
	Simple RT	−5504.99	−0.114	<0.001 **
	IT	0.002	0.050	0.112
Mental Fog	Age	−0.025	−0.410	<0.001 **
	Gender	0.348	0.174	<0.001 **
	Digit Span	−0.029	−0.051	0.245
	Visual WM	−0.002	−0.017	0.718
	Inferring Relevance	35.30	0.009	0.859
	Shape-Number	−37.12	−0.011	0.825
	Stroop	22.51	0.003	0.942
	Go/NoGo	−1498.37	−0.089	0.050
	Simple RT	−1794.27	−0.037	0.419
	IT	0.004	0.081	0.071
Interpersonal Anxiety	Age	−0.022	−0.358	<0.001 **
	Gender	0.214	0.107	0.007 **
	Digit Span	−0.007	−0.012	0.747
	Visual WM	−0.003	−0.027	0.522
	Inferring Relevance	40.08	0.010	0.817
	Shape-Number	−143.92	−0.043	0.329
	Stroop	−271.37	−0.039	0.319
	Go/NoGo	53.71	0.003	0.936
	Simple RT	−4043.57	−0.083	0.038 *
	IT	0.004	0.081	0.038 *
Somatisation	Age	−0.014	−0.228	<0.001 **
	Gender	0.272	0.136	<0.001 **
	Digit Span	−0.005	−0.008	0.776
	Visual WM	0.000	0.002	0.941
	Inferring Relevance	−89.74	−0.023	0.501
	Shape-Number	107.08	0.032	0.345
	Stroop	−57.52	−0.008	0.783
	Go/NoGo	−766.38	−0.045	0.136
	Simple RT	−5405.81	−0.112	<0.001 **

Dependent Variable	Parameter	B	B	Sig.
	IT	0.003	0.070	0.020 *
Alcohol	Age	−0.029	−0.468	0.223
	Gender	−1.62	−0.811	0.022 **
	Digit Span	0.240	0.430	0.214
	Visual WM	−0.041	−0.356	0.343
	Inferring Relevance	−297.26	−0.076	0.848
	Shape-Number	−393.96	−0.117	0.765
	Stroop	−682.54	−0.099	0.779
	Go/NoGo	8663.47	0.513	0.149
	Simple RT	−69,091.17	−10.426	<0.001 **
	IT	0.021	0.461	0.188
Cannabis	Age	−0.052	−0.857	0.005 **
	Gender	−0.646	−0.323	0.252
	Digit Span	−0.183	−0.328	0.234
	Visual WM	0.012	0.106	0.722
	Inferring Relevance	1064.69	0.273	0.389
	Shape-Number	−295.90	−0.088	0.778
	Stroop	861.75	0.124	0.657
	Go/NoGo	4069.70	0.241	0.394
	Simple RT	−34,946.23	−0.721	0.012 *
	IT	0.002	0.046	0.868
Other Substances	Age	0.014	0.223	0.648
	Gender	−0.332	−0.166	0.712
	Digit Span	0.533	0.954	0.031 *
	Visual WM	−0.112	−0.980	0.040 *
	Inferring Relevance	1196.78	0.307	0.545
	Shape-Number	1123.99	0.335	0.503
	Stroop	5111.74	0.738	0.100
	Go/NoGo	475.96	0.028	0.950
	Simple RT	−61098.64	−1.261	0.006 **
	IT	0.024	0.525	0.238

Note. * is significant at the 0.05 level. ** is significant at the 0.01 level. WM = Working Memory. RT = Reaction Time. IT = Inception Time.

The neurocognitive performance tasks failed to account for any unique variance in depression, agoraphobia, and mental fog. The speed of processing tasks provided unique predictive utility for the remaining six lower-level psychopathology domains. The working memory tasks were also able to account for unique variance in other substance use. No other neurocognitive tasks offered unique predictive utility for any of the symptom domains.

5.3.1.2. Higher-Level Psychopathology

As reported in Haywood et al. (2022), our multivariate multiple regression analyses revealed that our eight neurocognitive tasks in addition to age and gender accounted for a significant amount of variance in each internalising, externalising and the *p*-factor. The model

accounted for 23.8% of the variance in internalising ($F(10, 389) = 12.17$, $p < 0.001$, $R^2 = 0.238$), 15.6% of the variance in externalising ($F(10, 389) = 8.37$, $p < 0.001$, $R^2 = 0.156$), and 23.6% of the variance in the *p*-factor ($F(10, 389) = 12.05$, $p < 0.001$, $R^2 = 0.236$). Table 5.2 provides the results of the regression analyses as reported in Haywood et al. (2022)

Table 5.2

Higher-Level Psychopathology Linear Model Outcomes

Predictors	Internalising			Externalising			*p*-Factor		
	B	β	*p*	*B*	β	*p*	*B*	β	*p*
Age	−0.027	−0.433	<0.001 **	−0.026	−0.346	<0.001 **	−0.398	−0.434	<0.001 **
Gender	0.360	0.174	<0.001 **	0.007	0.003	0.951	4.68	0.156	0.001 **
Digit Span	−0.024	−0.041	0.360	0.036	0.054	0.249	−0.251	−0.030	0.505
Vis WM	−0.002	−0.013	0.783	−0.003	−0.022	0.669	−0.026	−0.015	0.758
Infer. Rel.	13.03	0.003	0.950	149.99	0.032	0.555	415.57	0.007	0.891
Shape-Num	−9.37	−0.003	0.958	−61.02	−0.015	0.778	−214.33	−0.004	0.934
Stroop	−182.03	−0.025	0.578	264.25	0.031	0.664	−1916.14	−0.018	0.686
Go/NoGo	−840.70	−0.048	0.297	−203.95	−0.010	0.835	−11,229.5	−0.044	0.336
Simple RT	−4973.28	−0.099	0.034 *	−12291.5	−0.209	<0.001 **	−84882.8	−0.117	0.012 *
IT	0.004	0.094	0.040 *	0.005	0.082	0.083	0.064	0.095	0.037 *

Note. * is significant at the 0.05 level. ** is significant at the 0.01 level. Vis WM = Visual Working Memory. Infer. Rel. = Inferring Relevance. Shape-Num = Shape-Number. RT = Reaction Time. IT = Inspection Time.

Only the neurocognitive tasks measuring speed of processing accounted for significant unique variance in higher-level psychopathology. Simple reaction time and Inspection Time were significant predictors of internalising and the *p*-factor, while simple reaction time was the sole significant predictor of externalising, bar age and gender. Tasks that measured working memory, shifting, or inhibition did not provide any unique predictive utility for the higher-level psychopathology factors. For further detail of these results see Haywood et al. (2022).

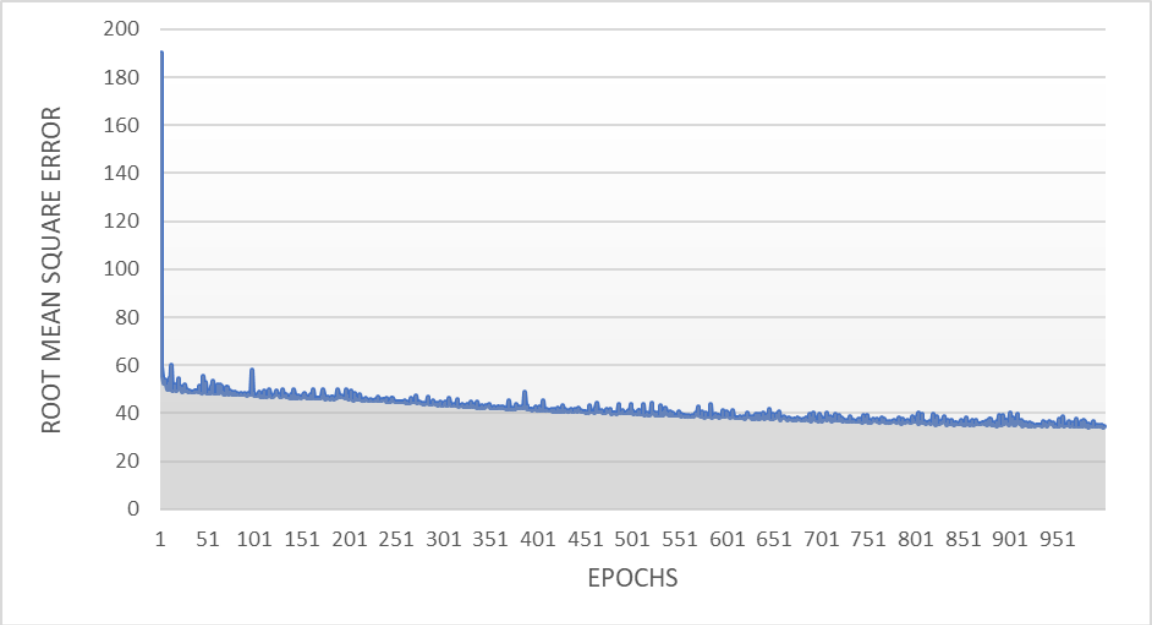
5.3.2 Artificial Neural Network Models

5.3.2.1. Lower-Level Psychopathology

Over the 1000 epochs of the basic backwards propagation, the lower-level psychopathology ANN model provided a final summed squared error of 34.76 and root mean squared error (RMSE) of 0.29. The model performed well with the relatively small number of

hidden units and a single hidden unit layer and learned very efficiently. For example, the summed squared error dropped from 190.27 following the first epoch to just 53.99 following the sixth epoch, and then learned steadily to end at a summed squared error of 34.76 on the 1000th epoch. The summed squared error to epochs for the lower-level psychopathology ANN are depicted in **Figure 5.4**

Root Mean Squared Error to Epochs for the Lower-level Psychopathology Artificial Neural Network Model



Note. Depiction of the accuracy and learning rate of the higher-level psychopathology artificial neural network model used in this research. The grey area under the blue line represents the summed squared error of the model that after particular number of epochs.

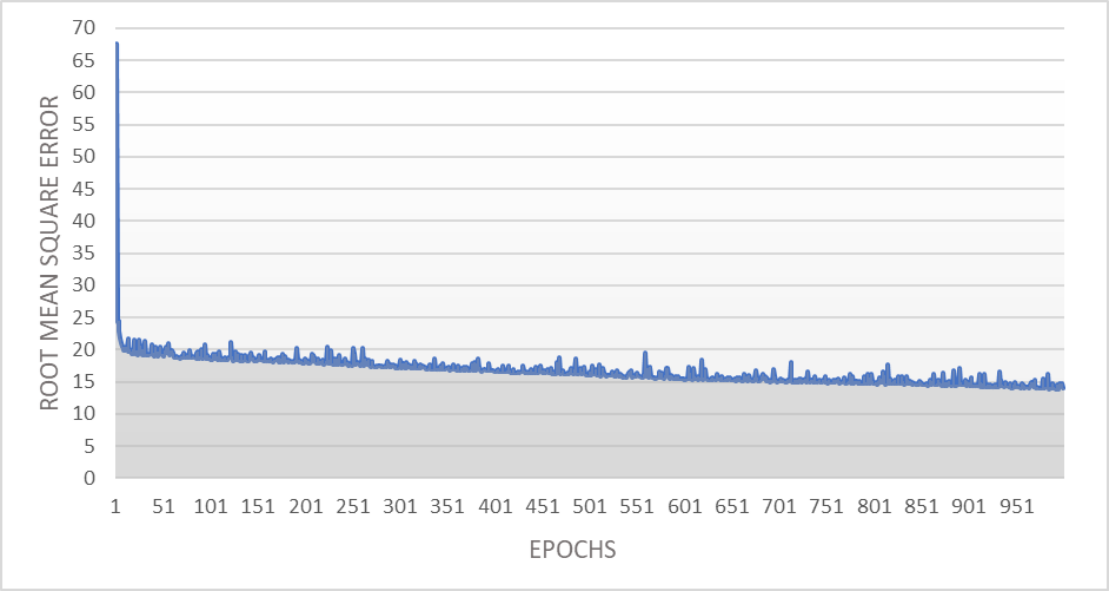
5.3.2.1. Higher-Level Psychopathology

Over the 1000 epochs the higher-level psychopathology ANN model provided a final summed squared error of 14.02 and a RMSE of 0.19. The higher-level psychopathology ANN performed better than the lower-level psychopathology ANN model (had a lower RMSE), however this may be attributed to the lower-level model having twice the number of output units. Again, even though the model was basic with a relatively small number of hidden units, and a single hidden unit layer, it learned efficiently. The summed squared error dropped from 67.55

following the first epoch to just 21.71 following the fifth epoch and learned progressively to end at a summed squared error of 14.02 on the 1000th epoch. The summed squared error to epochs for the higher-level psychopathology ANN are depicted in Figure 5.5.

Figure 5.5

Root Mean Squared Error to Epochs for the Higher-Level Psychopathology Artificial Neural Network Model.



Note. Depiction of the accuracy and learning rate of the higher-level psychopathology artificial neural network model used in this research. The grey area under the blue line represents the summed squared error of the model that after particular number of epochs.

5.3.3 Model Comparison

5.3.3.1. Lower-Level Psychopathology

The bivariate correlations between each lower-level psychopathology domain scores and the linear model and ANN model predicted scores are presented in Table 5.3. To allow easier comparisons to be made between linear and ANN approaches, Table 5.3 shows the results for linear and ANN models next to one another. For instance, the table shows the correlation between the observed depression scores and that predicted by the linear model (LM-Dep) is $r = 0.435$, versus $r = 0.648$ in the neural network model (ANN-Dep).

Table 5.3

Correlations Between Predicted and Actual Lower-Level Psychopathology Scores

	LM-Dep	ANN-Dep	LM-Agor	ANN-Agor	LM-Host	ANN-Host	LM-Fog	ANN-Fog	LM-Int. Anx	ANN-Int. Anx	LM-Soma	ANN-Soma	LM-Alc	ANN-Alc	LM-Cann	ANN-Cann	LM-Other	ANN-Other
Dep	0.435	0.648																
Agor			0.331	0.577														
Host					0.414	0.655												
Fog							0.476	0.711										
Int. Anx									0.449	0.675								
Soma											0.469	0.710						
Alc													0.268	0.338				
Cann															0.235	0.413		
Other																	0.243	0.552

Note. All correlations significant at $p < 0.01$ (one-tailed). LM = Linear Model. ANN = Artificial Neural Network Model. Dep = Depression. Agor = Agoraphobia. Fog = Mental Fog. Int. Anx = Interpersonal Anxiety. Soma = Somatisation. Alc = Alcohol use. Cann = Cannabis Use. Other = Other Substance use.

For each of the nine lower-level symptom domains the predicted values of the ANN model had a stronger correlation with the actual values when compared to the linear model. The correlations between the linear model's predicted values and the actual symptom values ranged between 0.243 and 0.476, while the correlations between the predicted values of the ANN model and the actual symptom values ranged between 0.338 and 0.711. The average correlation between the linear model's predicted values and the actual values was 0.369, while the average correlation between the ANN's predicted values and the actual values was 0.587. The difference between the linear and ANN models' average correlations with the actual values amongst the lower-level psychopathology domains was significant at a Bonferroni adjusted alpha level of 0.0125 ($Z = -4.027$. $p < 0.001$). Therefore, supporting hypothesis one, the ANN model performed significantly better than the linear model at predicting lower-level psychopathology.

5.3.3.2. Higher-Level Psychopathology

The bivariate correlations between each higher-level psychopathology factor scores and the linear model and ANN predicted scores are presented in Table 5.4.

Table 5.4

Correlations Between Predicted and Actual Higher-Level Psychopathology Scores

	LM-Int	ANN-Int	LM-Ext	ANN-Ext	LM-*p*	ANN-*p*
Internalising	0.488	0.661				
Externalising			0.421	0.619		
p-factor					0.486	0.666

Note. All correlations significant at p < 0.01 (one-tailed). LM = Linear Model. ANN = Artificial Neural Network Model. Int = Internalising. Ext -= Externalising. p = p-factor

Once again, for each of the three higher-level symptom domains the ANN model's predicted values had a stronger correlation with the actual values when compared to the linear model. The correlations between the linear model's predicted values and the actual symptom values ranged between 0.421 and 0.488, while the correlations between the ANN models' predicted values and the actual symptom values ranged between 0.619 and 0.666. The difference between the linear and ANN models' correlations with the actual values for internalising,

externalising, and the *p*-factor was significant at a Bonferroni adjusted alpha level of 0.0125. The ANN model was more accurate than the linear model at predicting internalising ($Z = -3.679$. $p < 0.001$), externalising ($Z = -3.867$. $p < 0.001$), and *p*-factor scores ($Z = -3.842$. $p < 0.001$). Therefore, supporting hypothesis two, the ANN model performed significantly better than the linear model at predicting lower-level psychopathology.

5.4 Discussion

The aim of this research is to compare the accuracy of linear models versus non-linear artificial neural network models with regard to how well they each predict (a) lower-level and (b) higher-level psychopathology. Overall, we found support for non-linear interactive relationships between the neurocognitive predictors and psychopathology. The ANN models were significantly more accurate than the linear models at predicting both lower-level and higher-level psychopathology. There is consensus that there is a high level of heterogeneity of neurocognition within psychopathology (Martino et al., 2008; Moritz et al., 2002), however understanding of the variability has been limited primarily by the use of descriptive or linear approaches and the use of DSM diagnostic categories. Previously, through computational modelling, we found that multiple different executive functioning profiles were able to account for the general neurocognitive performance of people with schizophrenia (Haywood & Baughman, 2021). This finding provided initial support for the multidimensional hypothesis, however, was limited by using a DSM defined disorder category that ignores that dimensionality and comorbidity of psychopathology. Using a dimensional approach, we find that the non-linear multidimensional conceptualisation is superior to traditional linear conceptualisations of the associations and functionality between neurocognition and psychopathology. Given that it is claimed that neurocognition is an aetiological feature of psychopathology (e.g., Beck & Rector, 2005; Romer & Pizzagalli, 2021), an accurate functional conceptualisation is fundamental to improving our understanding of psychopathology.

Previously, the search for a primary deficit of neurocognition within psychopathology has dominated the literature (Haywood et al., 2021c). While an understanding of a general trend of dysfunction across a specific population may be useful as a starting point to a fuller

understanding, our findings suggest further assessment of the within individual functionality of neurocognition is required. To illustrate, we recently found that measures of speed of processing, but not working memory, shifting, or inhibition, could significantly account for higher-level psychopathology linearly (Haywood et al., 2022). However, as in the present research the ANN models were superior in accuracy to the linear models, it suggests that working memory, shifting, and/or inhibition likely still play an important role in understanding the associations between neurocognition and psychopathology. Ultimately, as per the multidimensional hypothesis, the dynamic interactions between neurocognitive processes seem integral to a detailed understanding of the associations and functionality between neurocognition and psychopathology.

The use of dimensional, rather than categorical, conceptualisations of psychopathology in the present research has multiple strengths, including mitigating or accounting for the issues of comorbidity and diagnostic stability of the nosological approach (Kotov et al., 2021). However, examining the dynamic multidimensionality of neurocognition with regard to statistically derived higher-level factors of psychopathology does have conceptual considerations. While the lower-level scores of dimensional psychopathology (e.g., depression, hostility, etc.) were not factorised, scores of higher-level factors of psychopathology are intrinsically influenced by the scores of the population from which they were derived. For example, Lahey et al. (2021b) suggests that the *p*-factor is a "weighted average" (p. 61) of the sample's symptoms. Therefore, the *p*-factor (and internalising and externalising) scores on the individual level are dependent on the factors loadings of the indicators included in the sample model. Indeed, we have previously found that the underlying weightings of different lower-level psychopathology domains vary considerably between different samples (Haywood et al., 2021a). Findings such as these have led to the understanding that higher-level psychopathology factors may not have a universal substantive meaning (Haywood et al., 2021a; Levin-Aspenson et al., 2020). Considering the substantive interpretation difficulties of higher-level psychopathology, lower-level dimensional psychopathology may be better suited to enhance our understanding of the dynamics of neurocognition and psychopathology on the individual level.

An individual approach to neurocognition within developmental conditions, such as intellectual disability and autism spectrum disorder, is common in case conceptualisations and treatment approaches (Danielsson et al., 2012; Kleinhans et al., 2005). However, even though neurocognitive deficits are highly prevalent, albeit to generally a lesser severity, in psychopathology, this level of assessment and understanding is not commonplace (Egger et al., 2007). Our findings indicate that the dynamic multidimensionality, rather than general deficits, of neurocognition may be important to consider when understanding an individual's psychopathology. Further, our results imply that, beyond just strengths and weakness assessment common amongst developmental conditions' case conceptualisation, a consideration of the interactions between different neurocognitive domains' performance on the individual level may be important to understanding a person's psychological experience.

5.4.1 Limitations of the Research and Directions for Future Research

This research has four primary limitations. First, the data was collected online through Prolific (Palan & Schitter, 2018). Therefore, we had little control over the conditions under which data were obtained. However, there is evidence that the quality of task data collected through online platforms, in particular Prolific, is comparable to in-lab data (Crump et al., 2013; Johnson et al., 2021; Merz et al., 2020; Uittenhove et al., 2022). Second, age and gender were required to be predictors in both the linear and ANN models due to their associations between both neurocognition and psychopathology. While the role of age and gender in the linear models is easy to interpret, due to the structure and function of the ANN models the role age and gender played in these models is difficult to parse. Third, the comparisons between the linear models and the ANN models were able to provide evidence that the non-linear multidimensional conceptualisation of neurocognitive abilities in psychopathology is superior to the linear conceptualisation. However, our approach to the assessment of the ANN models was unable to provide the necessary information to detail the nuance of the multidimensional functionality. For example, we were not able to provide results for what neurocognitive profiles existed in the data, the specific interaction functionality, and what, if any, compensatory profiles existed. Nonetheless, the current research establishes the importance of considering dynamic

multidimensional explanations and provides future research with a platform for which to build upon. Lastly, in this study the type of ANN and linear models we developed were among the more accessible of techniques in their respective domains. More complex regression techniques, as well as more complex machine learning techniques exist. Examining how well some of these more complex techniques compare to one another, remains of interest to us for future work. Related to this last point, we also did not test a range of other architectures, activation functions, or use a larger data set. Though clearly each of these offer possible avenues for further study.

Future research should use tightly controlled lab-based data collection to explore non-linear multidimensional conceptualisations. Future research should also attempt to map the neurocognitive profiles that exist amongst the population, the functional dynamics of the neurocognitive domains, and their associations to dimensional psychopathology. More complex regression techniques and more complex machine learning techniques should be also examined and compared by future research. This knowledge may be used to inform aetiological theories of neurocognition and psychopathology and inform case conceptualisations on the individual level. Future research that uses a combination of computational modelling approaches(e.g., Haywood & Baughman, 2021), ANN approaches, and descriptive approaches may extend our knowledge of the non-linear multidimensionality.

5.5 Conclusion

In this research, we examined if neurocognitive ANN models were superior to linear models at predicting dimensional lower-level and higher-level psychopathology. We found support for the non-linear multidimensionality of neurocognition in psychopathology as the ANN models were significantly more accurate than the linear models at predicting both lower-level and higher-level psychopathology. Neurocognition was only able to account for a modest amount of variance in psychopathology, and our modelling approaches could not parse the variance in psychopathology accounted for by neurocognition, and age and gender. However, we still suggest that a non-linear multidimensional conceptualisation of neurocognition within psychopathology may be useful for aetiological examination and case conceptualisations. We also suggest that, due to the difficulties in interpreting the substantive meaning of higher-level factors of

psychopathology, the utility of examining the multidimensional functionality of neurocognition and psychopathology is greatest at the lower levels of psychopathology using dimensional measures.

6.0 General Discussion

In this, the final chapter of this thesis, I provide a summary of the key findings of this thesis related to each aim. I also provide a summary of the primary theoretical and practical implications for each study. I then conclude by providing an examination of the limitations of this work and directions for future research and an overall conclusion to the thesis.

6.1 Summary of Findings; Aim One

The first aim of this thesis was to examine if a universal substantive *p*-factor (and specific factors) could be developed by the assessment of the utility and consistency of structural models of psychopathology in subgroups. To address aim one, we used data simulation methods to create a dataset that closely mirrored the properties of that dataset used by Caspi et al. (2014) but had 100,000 rather than 1,000 cases. We then separated the dataset into 63 samples with heterogeneous symptom profiles, but adequate dimensional variability and attempted to fit four of the most popular structural models of psychopathology to each subgroup. We found that only eight out of the 64 subgroups fit one or more of the structural models. Of those eight subgroups they all fit the correlated factors model, none fit the bifactor model, four fit the revised bifactor model, and one subgroup fit the single factor model. Furthermore, the factor loadings and the neurocognitive correlates of subgroups fitted to the same model was highly variable.

6.1.1 Theoretical Implications

Overall, the structural models did poorly at accounting for the psychopathological symptoms of the subgroups. Simulation work has shown that fit indices bias bifactor models over correlated factors models (Greene et al., 2019), suggesting that the bifactors models were particularly poor at accounting for symptoms in our subgroups. Surprisingly, the correlated factors model fit more subgroups than the bifactors model. These findings have important theoretical implications, one of which is that structural models may be best suited to understanding the symptoms of the population as a whole, rather than subgroups with fairly homogenous symptom profiles. It is however, important to note that other research, has found structural models to account for symptoms in populations with a specific diagnosis or ailment (Shevlin et al., 2016; Xie et al., 2012). Therefore, in the context of the available literature,

structural models may be useful to describe some specific populations' symptoms, as long as the population in question generally has strong within group psychopathological symptom variability.

A second important implication is related to our finding that, for different subgroups that fit the same model (e.g., the bifactor model), the underlying factor loadings as well as the strength of the neurocognitive correlates, varied substantially. This suggests that the factors of psychopathology, for example the *p*-factor, have a different underlying substantive meaning between the subgroups. Therefore, a major implication of these finding is that it is unlikely that a universal substantive meaning of *p* (Fried et al., 2021; Levin-Aspenson et al., 2020) (and the other higher-order factors) could be developed. We did however find that for the eight subgroups that fit at least one model, their factors loadings were very highly correlated with the total population model's factor loadings. This may suggest that if a universal substantive meaning of the factors was to be developed (e.g., with specific components with which it must correlate within a specified range) it may only be found over the population as a whole.

The results of the exploration of aim one resulted in an improved understanding of the meaning and applicability of the factors of psychopathology; generating fundamental knowledge for the assessment of how neurocognition is associated with psychopathology. These results directed the methodological choices, and the interpretation of results in the remaining chapters.

6.1.2 Practical Implications

The potential for the development of a universal substantive *p,* and specific factors, has important implications for etiological and treatment domains of psychopathology (Fried et al., 2021; Lahey et al., 2021b). Ronald (2019) explains that treatment of psychopathology could be structured as specific (i.e., aimed at lower-level indicators, such as hostility), and general (i.e., aimed at the *p*-factor). Therefore, research exploring the underpinnings of the factors of psychopathology may lead to targeted (i.e., aimed to lessen specific lower-level psychopathology) and transdiagnostic (i.e., aimed to lessen the severity of general, *p,* experience) interventions (Ronald, 2019). However, for *p* and other factors of psychopathology to be useful in

the treatment setting, they need to be substantive constructs that can be applied to every individual consistently.

An important practical implication of our findings regarding aim one is that *p* and the other factors of psychopathology are fundamentally flexible entities that reflect a different construct depending upon (a) factor structure and components, and (b) the individual level data of the group being assessed. Therefore, applying research findings to treatment approaches that target *p* or other higher-order factors may lack transferability into the treatment setting more broadly. Our results suggest that more research into the nature, methodology, and substantive interpretation of structural models of psychopathology is required to optimise their applicability to treatment sectors.

6.2 Summary of Findings; Aim Two

The second aim of this thesis was to use and detail the S-1 bifactor model methodology in an exploration of neurocognitive ability and psychopathology. Driven by the findings of study one, that the fluidity of the underlying meaning of the *p*-factor limits clear interpretation of findings of the association between neurocognition and psychopathology (Haywood et al., 2021a), we examined the utility of the S-1 bifactor approach in understanding the association. Using simulated data, we found that that neurocognition could be successfully modelled as the general factor in a structural model of psychopathology. We found that within the simulated data a general executive functioning general factor was superior to an IQ factor in accounting for dimensional psychopathology. The association between internalising and externalising in the EF referenced S-1 bifactor model fell substantially when compared to the correlated factors model using the same data.

6.2.1 Theoretical Implications

The major theoretical implication of these findings is the demonstration of the use of neurocognition to pre-specify the substantive meaning of the general factor, mitigating the fluidity of an unspecified *p*-factor. We, in study one, along with work by other authors (e.g., Fried et al., 2021; Greene et al., 2021; Levin-Aspenson et al., 2020), have demonstrated that the *p*-factor lacks substantive consistency across models, samples, and methods. Therefore, even

though the *p*-factor has been claimed to be a substantive entity representative of peoples' overall propensity toward psychopathology, *p* has little consistency in its statistical make up (e.g., Fried et al., 2021; Greene et al., 2021; Levin-Aspenson et al., 2020), making the interpretation of neurocognitive findings relating to *p* difficult.

The utility of using S-1 bifactor models, with neurocognition modelled as the general factor, offers a unique way to assess and interpret the multifactorial associations between lower and higher-level psychopathology that mitigates the theoretical issues associated with having a fluid general factor. However, it is important to acknowledge three important limitations of the S-1 bifactor approach for these purposes. First, the use of a neurocognitive S-1 bifactor model relies on each of the neurocognitive measures having a relatively strong correlation. Secondly, solely using the S-1 bifactor model means that the investigation of conceptualisations between individual domains of neurocognitive and psychopathology (i.e., the multidimensional hypothesis) cannot be made. Thirdly, this approach, as with other structural modelling approaches (Haywood et al., 2021b), may only be applicable to the population as a whole, and not subgroups. Therefore, in addition to the S-1 approach, other methodologies are required to develop a full understanding of the true relation between neurocognition and psychopathology.

6.2.2 Practical Implications

The primary practical implication, of our S-1 bifactor model findings, is related to the dynamics between neurocognition, internalising and externalising. In the executive functioning referenced S-1 bifactor model, the correlation between internalising and externalising fell substantially when compared to the correlated factors model, suggesting that executive functioning may be particularly important in the functional dynamic between internalising symptoms (e.g., anxiety) and externalising behaviours (e.g., substance use). This suggests, across this sample, that neurocognitive abilities accounts for most of the comorbid dimensional symptoms across internalising and externalising. Therefore, clinical utility of the S-1 bifactor approach to examine neurocognitive abilities and psychopathology is that findings, in the way of patterns of factor loadings, may point towards specific collections of symptoms that people experience and the severity of those symptoms, and how they may be functionally related to their

neurocognitive abilities. This can then be used to inform case conceptualisations, and treatment and rehabilitation approaches (Reser et al., 2019). The lack of clear and consistent substantive interpretation of the *p*-factor in traditional structural models, demonstrated in study one (Haywood et al., 2021b), means the clinical utility of associations between neurocognition and *p* is unclear. However, as demonstrated in study two (Haywood et al., 2021a), the S-1 bifactor model, with a general neurocognitive factor, can lead to important, clinically transferable, findings.

6.3 Summary of Findings; Aim Three

The third aim of this thesis was to examine if, at the population level, each of the factors of psychopathology can be partly explained by discrete association patterns of neurocognitive component abilities, with little variability. As explored in detail in chapters one, two and three, the higher-level factors of psychopathology are proposed to be separate but related, substantive constructs (Haywood et al., 2021a, 2021b, 2021c). In study three we explored if there are discrete patterns of neurocognitive associations between the factors to inform our understanding of their substance, and the utility of neurocognition in accounting for the factors. Using human, rather than simulated data as in previous chapters, we found that a correlated factors model, with internalising and externalising higher-order factors, and a single factor model, provided good fits for our psychopathology data. However, the internalising factors in the correlated factors model was almost perfectly correlated with the *p*-factor from the single factor model. Using these models, we found that tasks that measured speed of processing, but not tasks that measured working memory, shifting or inhibition, had a significant association with the three factors. Each factor, internalising, externalising and *p* therefore, had a common negative association with speed of processing, but not any other measured neurocognitive components. Although, speed of processing significantly accounted for each of the factors of psychopathology, it is important to acknowledge the modest percentage of unique variance explained by the speed of processing tasks (1.7 - 4%). We believe that this modest percentage of unique variance is still useful to inform theoretical developments and practical applications, each of which is discussed the following sections.

6.3.1 Theoretical Implications

There are several major theoretical implications to the findings related to aim three. First, the *p*-factor from the single factor model represented internalising, rather than general psychopathology. This is evidenced by the almost perfect correlation between internalising and *p*. This finding illustrates the issues presented in chapters one, two and three regarding the lack of a constant substantive meaning of *p* (Haywood et al., 2021a, 2021b, 2021c). This finding is not exclusive. For example, Caspi et al. (2014)'s *p*-factor was shown to represent thought disorder rather than general psychopathology (Eid, 2020; Haywood et al., 2021a; Junghänel et al., 2020), leading them to later suggest that the substantive meaning of *p* is thought disorder (Caspi & Moffitt, 2018). Our results further demonstrate the lack of a consistent interpretation of *p*.

The second major theoretical implication of our findings is that, within our data, the factors of psychopathology could not be discerned by discrete patterns of neurocognitive abilities. Therefore, within this data and at the population level, the separate theoretical substantive nature of each of the factors cannot be explained via neurocognitive profiles. The general negative association between speed of processing measuring and each factor of psychopathology, across both the correlated factors model, and the single factor model, suggests that poor speed of processing is associated with higher general psychopathology. Therefore, on the population level using linear methods, speed of processing is useful at accounting for general, but not more specific psychopathology, while other forms of neurocognition do not add significant predictive utility. This finding does add support for a theoretical conceptualisation of depression by Nuño et al. (2021). Nuño et al. (2021) suggests that people with high levels of depression symptomology have speed of processing deficits and compensate for this by using greater cognitive effort. However, if the task at hand is of a high level of difficulty therefore requiring high levels of cognitive effort, the speed of processing deficits cannot be compensated for, and poor performance emerges. Nuño et al. (2021) states that, in this way, speed of processing may partly explain poorer occupational performance of people with depression. Given that in our data internalising (where depressive symptoms is located) came to also represent *p*, it lends support to

this theoretical conceptualisation of functionality, while speed of processing has a known consistent functional association with externalising symptoms (e.g., Durazzo et al., 2008).

The final theoretical implication of these findings is that a dynamic multidimensional approach to exploring neurocognition and psychopathology may be superior to the linear approach use in this study. Only speed of processing was significantly associated with each of the factors at the population level using linear methods, however working memory, shifting, and inhibition each had theoretical grounding functionally linking them to psychopathology (Haywood et al., 2021c; Romer & Pizzagalli, 2021). These findings further drove our explorations in study 4 that explored if the dynamic multidimensional conceptualisation of neurocognitive performance would be superior to the linear approach.

6.3.2 Practical Implications

The findings of study three have two main practical implications. The first implication is, once again, the potential for the *p* factor to drive treatment and rehabilitation approaches (Ronald, 2019). In this study, *p* in the single factor model mirrored internalising in the correlated factors model. If, taken on face value, or the associations between the different models' factors were not explored, *p* might have (as theoretically suggested) been interpreted as representing general psychopathology. Therefore, this research shows that the practical application of the *p*-factor from research to the treatment and rehabilitation setting is fraught by the inconstant nature of the *p*-factor. This research provides further knowledge to clinicians and researchers alike to exercise caution in interpreting research on a fluid *p*-factor in context of treatment and rehabilitation.

The second main practical implication is the general negative association between speed of processing and the factors of psychopathology. Evidence for the functional association between speed of processing and psychopathology is growing (Nuño et al., 2021), and the findings of study three reinforce the potential importance of speed of processing in psychopathology across domains. Speed of processing has been primarily studied, and considered by clinicians, in the domain of ageing (Albinet et al., 2012; Kail, 1991), however our results, along with emerging evidence, suggests that across the population people with higher levels of psychopathology generally have deficits in speed of processing. Speed of processing deficits may

then have a negative impact on daily life and exacerbate their psychopathology (Nuño et al., 2021). Clinicians may look to assess speed of processing in their patients and understand the impact speed of processing may have on their daily life and psychopathology. Clinicians and researchers may also look to implement, and further trial, training cognitive training specifically for speed of processing which has primarily been used in ageing and dementia (e.g., Verghese et al., 2021)

6.4 Summary of Findings; Aim Four

The fourth aim of this thesis was to assess if each factor of psychopathology (e.g., internalising, externalising, and the *p*-factor) is usefully explained by multidimensional interactive components of neurocognition. In each of the chapters thus far we discussed the potential for an alternative multidimensional conceptualisation of neurocognitions association with psychopathology. This conceptualisation proposed that instead of general trends of deficits in neurocognition in psychopathology across the population, the association may be best conceptualised by acknowledging the potential dynamic multidimensional interactions between different neurocognitive domains and dimensional psychopathology. In study four we explored if neurocognitive artificial neural network models, that learn multidimensional patterns in the data, were superior to linear neurocognitive models in predicting both lower-level and higher-level psychopathology. We used the data and models of psychopathology developed in study three to examine this research question. The artificial neural network models were significantly more accurate than the linear models at predicting both lower-level and higher-level psychopathology.

6.4.1 Theoretical Implications

The findings relating to aim four supported our dynamic multidimensional interactive conceptualisation of neurocognition in psychopathology and provided important theoretical implications. As discussed in chapter one, historically the dominance of the application of a common cause model to psychiatry has meant that theoretical conceptualisations of neurocognitive dysfunction in psychopathology have been linear or dichotomous (Haywood et al., 2021c). The association between neurocognition and psychopathology is now understood to

be extensively heterogeneous (Martino et al., 2008; Moritz et al., 2002), but conceptual understandings of the functionality primary remain linear or dichotomous.

To illustrate the major theoretical implication of study four, that multifinality of neurocognition in psychopathology is superior to linear or dichotomous approaches, I will contrast the approaches taken in study three and study four. In study three we applied a linear conceptualisation and showed that, across the population, deficits in speed of processing were associated with greater levels of high-level psychopathology (Haywood et al., 2022). While this finding is potentially useful to inform risk models or understandings of cognitive and psychological experiences across the population, it may also be taken to suggest that, theoretically, the other neurocognitive domains measured were not important to consider. In contrast, in study four we used a dynamic multidimensional conceptualisation and developed artificial neural network models that learned non-linear association patterns between the neurocognitive components. We found that the artificial neural network models were superior to the linear models used in study three. Therefore, it is likely that working memory, shifting, inhibition and speed of processing were all important to understanding psychopathology, while study three pointed towards speed of processing as being the only neurocognitive domain of interest. These alternate approaches have the potential of supporting and directing differing theoretical and aetiological explanations, and therefore it is paramount for future research to apply dynamic multidimensional conceptualisations. The findings from studies one and two were also integral to interpreting the findings of study four. The higher-order factors of psychopathology, in particular the *p*-factor, through our work in studies one and two, as well as other literature (Eid, 2020; Fried et al., 2021; Heinrich et al., 2020) are now understood to be flexible domains without universal substantive interpretation. Therefore, we suggest that to better understand the functionality between neurocognition and psychopathology, theoretical explanations should be contrived using a dynamic multidimensional approach at the lower-level (non-factorised), rather than the higher-level, of dimensional psychopathology. To conclude, the primary theoretical implication of the findings relating to aim four was that dynamic

multidimensional conceptualisations of neurocognition in psychopathology are integral to understanding their functional association.

6.4.2 Practical Implications

There are a number of pertinent practical implications resulting from our findings assessing aim four. The key implications relate to case conceptualisations, and treatment and rehabilitation approaches. Neurocognition in conditions such as intellectual disability and autism are often assessed at the individual level, and this informs case conceptualisations and responses (Danielsson et al., 2012; Kleinhans et al., 2005). This level of detailed assessment is not commonly completed within broader psychopathology (Egger et al., 2007), and may provide integral information to facilitate optimal care, including treatment, rehabilitation, and practical aids.

Our finding related to aim four suggesting that considering the dynamic multidimensionality of neurocognition is important to understand the cognitive and psychological experiences of an individual, points towards to the need for nuanced assessment of the dynamics of neurocognition in clinical settings. Related to assessment approaches, our findings suggest that in the clinical setting, ideally, tasks designed to assess individual domains of neurocognition (i.e., working memory, shifting, inhibition, and speed of processing), should be preferred to tasks that measure their performance simultaneously as individual contributions and interactions cannot be teased apart (Haywood & Baughman, 2021). Further, our findings suggest going beyond just strength and weakness assessments common when using tools such as the WAIS (Zimmerman et al., 1973). More advanced tools are needed that consider the functional dynamics between different domains of neurocognition and their associations with the outcome of interest. The dynamic multidimensional conceptualisation may also inform the development of more advanced cognitive remediation approaches that consider the multidimensionality of neurocognition in psychopathology. Ultimately, the findings related to aim four provide a host of practical implications for current service provision, as well as for the development of future assessment and treatment approaches.

6.5 Limitations and Future Directions

The primary limitations of the research and directions for future research are presented in each chapter and are related to the data simulation approach taken in studies one and two, and the collection of data for studies three and four using an online crowdsourcing platform. For studies one and two we developed data that simulated the data used by Caspi et al. (2014) in their pivotal work developing structural models of psychopathology. Instead of using Monte-Carlo simulation methods, that can be used to sample subsets of data from a range of population parameters (Mooney, 1997), we developed a single dataset (with a large sample size) for each study to parallel empirical examinations using human data. While, our approach had the strength of extending current knowledge developed by specifically paralleling the data and approach used by pivotal work (Caspi et al., 2014) in this domain of inquiry, a Monte-Carlo approach may allow for the exploration of boundaries of explanation using the structural models and improved generalisation across different samples (e.g., Greene et al., 2019). Future work should look to further use Monte-Carlo methods to explore the boundaries of utility of structural models of psychopathology.

Limitation in studies three and four relate primarily to the data collection methods. We used the crowd-sourcing platform Prolific to collect data on psychopathology, substance use and neurocognitive performance. By using this approach to data collection, we did not have control of the specific environments the participants completed the neurocognitive tasks, or the hardware used. This may have introduced noise in the data, specifically for the neurocognitive tasks. However, evidence points towards the quality of behavioural task data collected using crowd-sourcing platforms being comparable to lab-based studies (Crump et al., 2013; Johnson et al., 2021; Merz et al., 2020; Sauter et al., 2020). Further, we used the Prolific crowd-sourcing platform which has been shown to be the most valid platform to collect data of this type (Uittenhove et al., 2022). Future research should look to replicate our findings using tightly controlled lab settings.

A limitation of the each of the studies in this thesis generally is that the analyses were all cross-sectional. Even though the simulated data was developed from Caspi et al.'s (2014) models

that used longitudinal data, our explorations were primarily not across different time points. For example, in study two we used intelligence measures administered at timepoints 5 years of age, 7-11 years of age, and adult in the same S-1 bifactor model to represent lifetime IQ. Deficits in neurocognition that that are commonly seen in psychopathology have been shown to often pre-date the onset of psychopathological symptoms (Romer & Pizzagalli, 2021), and many aetiological explanations of psychopathology incorporate neurocognitive deficits as predisposing, precipitating, and perpetuating factors (e.g., Beck & Rector, 2005). However, our cross-sectional approach to exploring neurocognition did not facilitate insight into the developmental trajectory or association profiles between neurocognition and dimensional psychopathology. Future research should use longitudinal data and multilevel modelling analysis approaches to further elucidate developmental insights into neurocognition and structural psychopathology over time and may use S-1 bifactor modelling approaches to facilitate interpretation. Further, the longitudinal approaches, given the encouraging results relating to the dynamic multidimensional conceptualisation of neurocognition in psychopathology, should seek to apply a multidimensional approach to neurocognition in psychopathology. This future research may then facilitate the development of a neurocognitive assessment package, analysed by artificial neural networks, that is predictive of psychopathology over time, and is superior to the predictive accuracy of current linear-based approaches.

6.6 Conclusion to the Thesis

The objective of this thesis was to provide an improved understanding of the associations between neurocognition and psychopathology, taking into consideration the rise of dimensional and structural psychopathology. We examined the (a) applicability, utility and potential for a substantive meaning of the *p*-factor and other higher-order factors of psychopathology, (b) the utility of an alternative structural modelling approach to psychopathology to mitigate the substantive fluidity of the *p*-factor (c) the neurocognitive correlates of the factors of psychopathology and if the profile of neurocognitive correlates differed substantially between the factors, and (d) if a non-linear multidimensional conceptualisation of neurocognition in dimensional psychopathology was superior to traditional linear conceptualisations. Overall, the

findings suggested that structural models of psychopathology may not be particularly applicable to many subgroups of a population, and it will be difficult to develop a universal substantive meaning of the factors, particularly the *p*-factor. S-1 bifactor modelling approaches, using neurocognition as the general factor, offer a unique way to mitigate the issue of a general factor without consistent substantive interpretation and facilitate the interpretation of the associations between neurocognition and psychopathology. Although, while S-1 bifactor models have many unique strengths, they are unable to answer research questions involving the assessment of individual neurocognitive components, and therefore additional approaches are needed. We found that the factors of psychopathology did not have discrete patterns of neurocognitive correlates, and instead, a general negative association was found between tasks that measured speed of processing and each factor of psychopathology. We proposed that, although this result might on face-value point toward working memory, shifting and inhibition not being important to understanding the samples psychopathology, the traditional linear conceptualisation used cannot account for potential non-linear functional dynamics between the neurocognitive components. Supporting this proposition, we found that artificial neural network models, that can learn complex non-linear relations between predictors and outcomes, had significantly greater accuracy at predicting both lower-level and higher-level psychopathology, when compared to a traditional linear approach. We therefore suggested that using dynamic multidimensional conceptualisations regarding the associations between neurocognition and psychopathology is integral to advancing aetiological theories as well as assessment and treatment approaches. Future research should look to use Monte-Carlo simulation techniques to further understand the boundaries of utility of structural models of psychopathology. Future research should also look to replicate our findings using in-lab data collection and extend upon our findings by exploring longitudinal data. We suggest that future research, due to the fluidity of the *p*-factor and the higher-order factors, use the multidimensional neurocognitive approach primarily on the lower-levels of dimensional psychopathology. Practically, our findings, along with future research, may be used to develop multidimensional prediction, assessment, treatment, and rehabilitation approaches for psychopathology.

DARREN HAYWOOD SCHOOL OF POPULATION HEALTH

References

Akhavan Abiri, F., & Shairi, M. R. (2020). Validity and Reliability of Symptom Checklist-90-Revised (SCL-90-R) and Brief Symptom Inventory-53 (BSI-53). *Clinical Psychology and Personality, 17*(2), 169-195.

Albinet, C. T., Boucard, G., Bouquet, C. A., & Audiffren, M. (2012). Processing speed and executive functions in cognitive aging: How to disentangle their mutual relationship? *Brain and Cognition, 79*(1), 1-11.

Aldao, A., Gee, D. G., De Los Reyes, A., & Seager, I. (2016). Emotion regulation as a transdiagnostic factor in the development of internalizing and externalizing psychopathology: Current and future directions. *Development and Psychopathology, 28*(4), 927-946.

Anderson, M., Reid, C., & Nelson, J. (2001). Developmental changes in inspection time: What a difference a year makes. *Intelligence, 29*(6), 475-486.

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33*(4), 245-266.

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology, 15*(1), 31-36.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

Barceló, F., & Knight, R. T. (2002). Both random and perseverative errors underlie WCST deficits in prefrontal patients. *Neuropsychologia, 40*(3), 349-356.

Baughman, F. D., & Thomas, M. S. C. (2008). Specific impairments in cognitive development: A dynamical systems approach. In: 30th Annual Meeting of the Cognitive Science Society, 23-26 July 2008, Washington, U.S.

Bechtel, W., & Mundale, J. (1999). Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science, 66*(2), 175-207.

Beck, A. T., & Rector, N. A. (2005). Cognitive approaches to schizophrenia: theory and therapy. *Annu. Rev. Clin. Psychol., 1*, 577-606.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238.

Bloemen, A. J. P., Oldehinkel, A. J., Laceulle, O. M., Ormel, J., Rommelse, N. N. J., & Hartman, C. A. (2018). The association between executive functioning and psychopathology: general or specific. *Psychological Medicine, 48*(11), 1787-1794.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective* (Vol. 467). John Wiley & Sons. Hoboken, NJ: Wiley

Bornovalova, M. A., Choate, A. M., Fatimah, H., Petersen, K. J., & Wiernik, B. M. (2020). Appropriate use of bifactor analysis in psychopathology research: appreciating benefits and limitations. *Biological Psychiatry*.

Bosia, M., Bechi, M., Bosinelli, F., Politi, E., Buonocore, M., Spangaro, M., Bianchi, L., Cocchi, F., Guglielmino, C., & Cavallaro, R. (2019). From cognitive and clinical substrates to functional profiles: Disentangling heterogeneity in schizophrenia. *Psychiatry Research, 271*, 446-453.

Brandes, C. M., Herzhoff, K., Smack, A. J., & Tackett, J. L. (2019). The p factor and the n factor: Associations between the general factors of psychopathology and neuroticism in children. *Clinical Psychological Science, 7*(6), 1266-1284.

Bregant, J. (2006). John Bickle, Philosophy and Neuroscience: A Ruthlessly Reductive Account. *Croatian Journal of Philosophy, 6*(16), 133-140.

Bringmann, L. F., & Eronen, M. I. (2018). Don't blame the model: Reconsidering the network approach to psychopathology. *Psychological Review, 125*(4), 606.

Brooks, S. J., Funk, S. G., Young, S. Y., & Schiöth, H. B. (2017). The role of working memory for cognitive control in anorexia nervosa versus substance use disorder. *Frontiers in Psychology, 8*, 1651.

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research, 21*(2), 230-258. https://doi.org/10.1177/0049124192021002005

Burdick, K. E., Goldberg, J. F., Harrow, M., Faull, R. N., & Malhotra, A. K. (2006). Neurocognition as a stable endophenotype in bipolar disorder and schizophrenia. *The Journal of Nervous and Mental Disease, 194*(4), 255-260.

Burke, J. D., & Johnston, O. G. (2020). The bifactor S-1 model: A psychometrically sounder alternative to test the structure of ADHD and ODD? *Journal of Abnormal Child Psychology, 48*(7), 911-915.

Burns, G. L., Geiser, C., Servera, M., Becker, S. P., & Beauchaine, T. P. (2019). Application of the bifactor S–1 model to multisource ratings of ADHD/ODD symptoms: An appropriate bifactor model for symptom ratings. *Journal of Abnormal Child Psychology*, 1-14.

Cannon, T. D. (2015). How schizophrenia develops: cognitive and brain mechanisms underlying onset of psychosis. *Trends in Cognitive Sciences, 19*(12), 744-756.

Carruthers, S. P., Gurvich, C. T., Meyer, D., Bousman, C., Everall, I. P., Neill, E., Pantelis, C., Sumner, P. J., Tan, E. J., & Thomas, E. H. X. (2019). Exploring heterogeneity on the Wisconsin card sorting test in schizophrenia spectrum disorders: a cluster analytical investigation. *Journal of the International Neuropsychological Society, 25*(7), 750-760.

Carruthers, S. P., Van Rheenen, T. E., Gurvich, C., Sumner, P. J., & Rossell, S. L. (2019). Characterising the structure of cognitive heterogeneity in schizophrenia spectrum disorders. A systematic review and narrative synthesis. *Neuroscience & Biobehavioral Reviews, 107*, 252-278.

Carver, C. S., Johnson, S. L., & Timpano, K. R. (2017). Toward a functional view of the p factor in psychopathology. *Clinical Psychological Science, 5*(5), 880-889.

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., & Poulton, R. (2014). The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science, 2*(2), 119-137.

Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *American Journal of Psychiatry, 175*(9), 831-844.

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology, 46*(1), 121-153.

Clarkson, T., Kang, E., Capriola-Hall, N., Lerner, M. D., Jarcho, J., & Prinstein, M. J. (2019). Meta-Analysis of the RDoC Social Processing Domain across Units of Analysis in Children and Adolescents. *Journal of Clinical Child & Adolescent Psychology*, 1-25.

Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition, 34*(8), 1754-1768.

Cowen, P. J. (2016). Neuroendocrine and neurochemical processes in depression. *Psychopathology Review, 3*(1), 3-15.

Craddock, N., Owen, M. J., & O'Donovan, M. C. (2006). The catechol-O-methyl transferase (COMT) gene as a candidate for psychiatric phenotypes: evidence and lessons. *Molecular Psychiatry, 11*(5), 446-458.

Crow, A. J. D. (2019). Associations between neuroticism and executive function outcomes: response inhibition and sustained attention on a Continuous Performance Test. *Perceptual and Motor skills, 126*(4), 623-638.

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PloS One, 8*(3), e57410.

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: A multilevel framework for attitudes and evaluation. *Social Cognition, 25*(5), 736-760.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological methods, 1*(1), 16.

Cuthbert, B. N. (2020). The role of RDoC in future classification of mental disorders. *Dialogues in Clinical Neuroscience, 22*(1), 81.

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Medicine, 11*(1), 126.

Daban, C., Mathieu, F., Raust, A., Cochet, B., Scott, J., Etain, B., Leboyer, M., & Bellivier, F. (2012). Is processing speed a valid cognitive endophenotype for bipolar disorder? *Journal of Affective Disorders, 139*(1), 98-101.

Danielsson, H., Henry, L., Messer, D., & Rönnberg, J. (2012). Strengths and weaknesses in executive functioning in children with intellectual disability. *Research in Developmental Disabilities, 33*(2), 600-607.

De Lissnyder, E., Koster, E. H. W., Derakshan, N., & De Raedt, R. (2010). The association between depressive symptoms and executive control impairments in response to emotional and non-emotional information. *Cognition and Emotion, 24*(2), 264-280.

Derogatis, L. R., & Melisaratos, N. (1983). The brief symptom inventory: an introductory report. *Psychological Medicine, 13*(3), 595-605.

Dolce, P., Marocco, D., Maldonato, M. N., & Sperandeo, R. (2020). Toward a machine learning predictive-oriented approach to complement explanatory modeling. an application for evaluating psychopathological traits based on affective neurosciences and phenomenology. *Frontiers in Psychology, 11*, 446.

Durazzo, T. C., Gazdzinski, S., Yeh, P.-H., & Meyerhoff, D. J. (2008). Combined Neuroimaging, Neurocognitive and Psychiatric Factors to Predict Alcohol Consumption Following Treatment for Alcohol Dependence. *Alcohol and Alcoholism, 43*(6), 683-691. https://doi.org/10.1093/alcalc/agn078

Eadeh, H.-M., Markon, K. E., Nigg, J. T., & Nikolas, M. A. (2021). Evaluating the Viability of Neurocognition as a Transdiagnostic Construct Using Both Latent Variable Models and Network Analysis. *Research on Child and Adolescent Psychopathology, 49*(6), 697-710.

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., Goldman, D., & Weinberger, D. R. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences, 98*(12), 6917-6922.

Egger, J. I. M., De Mey, H. R. A., & Janssen, G. T. L. (2007). Assessment of executive functioning in psychiatric disorders: Functional diagnosis as the ouverture of treatment.

Eid, M. (2020). Multi-Faceted Constructs in Abnormal Psychology: Implications of the Bifactor S-1 Model for Individual Clinical Assessment. *Journal of Abnormal Child Psychology*, 1-6.

Endres, M. J., Donkin, C., & Finn, P. R. (2014). An information processing/associative learning account of behavioral disinhibition in externalizing psychopathology. *Experimental and Clinical Psychopharmacology, 22*(2), 122.

Endres, M. J., Rickert, M. E., Bogg, T., Lucas, J., & Finn, P. R. (2011). Externalizing psychopathology and behavioral disinhibition: Working memory mediates signal discriminability and reinforcement moderates response bias in approach–avoidance learning. *Journal of Abnormal Psychology, 120*(2), 336.

Fried, E. I. (2021). Studying mental disorders as systems, not syndromes. *PsyArXiv*.

Fried, E. I., Coomans, F., & Lorenzo-Luaces, L. (2020). The 341 737 ways of qualifying for the melancholic specifier. *The Lancet Psychiatry, 7*(6), 479-480.

Fried, E. I., Greene, A. L., & Eaton, N. R. (2021). The p factor is the sum of its parts, for now. *World Psychiatry, 20*(1), 69.

Galandra, C., Crespi, C., Basso, G., Manera, M. R., Giorgi, I., Poggi, P., & Canessa, N. (2021). Decreased information processing speed and decision-making performance in alcohol use disorder: combined neurostructural evidence from VBM and TBSS. *Brain Imaging and Behavior*, *15*(1), 205-215.

Galletly, C. A., MacFarlane, A. C., & Clark, C. R. (2007). Impaired updating of working memory in schizophrenia. *International Journal of Psychophysiology, 63*(3), 265-274.

Gamboz, N., Borella, E., & Brandimonte, M. A. (2009). The role of switching, inhibition and working memory in older adults' performance in the Wisconsin Card Sorting Test. *Aging, Neuropsychology, and Cognition*, *16*(3), 260-284.

Geurts, H., Sinzig, J., Booth, R., & HappÉ, F. (2014). Neuropsychological heterogeneity in executive functioning in autism spectrum disorders. *International Journal of Developmental Disabilities*, *60*(3), 155-162.

Gilleen, J., David, A., & Greenwood, K. (2016). Self-reflection and set-shifting mediate awareness in cognitively preserved schizophrenia patients. *Cognitive Neuropsychiatry*, *21*(3), 185-196.

Green, R., Till, C., Al-Hakeem, H., Cribbie, R., Téllez-Rojo, M. M., Osorio, E., Hu, H., & Schnaas, L. (2019). Assessment of neuropsychological performance in Mexico city youth using the cambridge neuropsychological test automated battery (CANTAB). *Journal of Clinical and Experimental neuropsychology*, *41*(3), 246-256.

Greene, A. L., Eaton, N. R., Li, K., Forbes, M. K., Krueger, R. F., Markon, K. E., Waldman, I. D., Cicero, D. C., Conway, C. C., & Docherty, A. R. (2019). Are fit indices used to test psychopathology structure biased? A simulation study. *Journal of Abnormal Psychology*, *128*(7), 740.

Greene, A. L., Watts, A. L., Forbes, M. K., Kotov, R., Krueger, R., & Eaton, N. (2021). Misbegotten Methodologies and Forgotten Lessons from Tom Swift's Electric Factor Analysis Machine: A Demonstration with Competing Structural Models of Psychopathology. *Psychological Methods.*

Hartman, M., Bolton, E., & Fehnel, S. E. (2001). Accounting for age differences on the Wisconsin Card Sorting Test: Decreased working memory, not inflexibility. *Psychology and Aging*, *16*(3), 385.

Hauke, J., & Kossowski, T. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae, 30*(2), 87.

Haywood, D., Baughman, F., Mullan, B., & Heslop, K. R. (2021a). Going “Up” to Move Forward: S-1 Bifactor Models and the Study of Neurocognitive Abilities in Psychopathology. *International Journal of Environmental Research and Public Health, 18*(14), 7413.

Haywood, D., & Baughman, F. D. (2021). Multidimensionality in Executive Function Profiles in Schizophrenia: A Computational Approach Using the Wisconsin Card Sorting Task. *Computational Brain & Behavior*, 1-14.

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2021b). One p-Factor for All? Exploring the Applicability of Structural Models of Psychopathology within Subgroups of a Population. *International Journal of Environmental Research and Public Health, 18*(13), 7108.

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2021c). Psychopathology and Neurocognition in the Era of the p-Factor: The Current Landscape and the Road Forward. *Psychiatry International, 2*(3), 233-249.

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2022, Mar 22). What Accounts for the Factors of Psychopathology? An Investigation of the Neurocognitive Correlates of Internalising, Externalising, and the p-Factor. *Brain Sciences, 12*(4), 22.

Haywood, D., Baughman, F. D., Mullan, B. A., & Heslop, K. R. (2022). What Accounts for the Factors of Psychopathology? an Investigation of the Neurocognitive Correlates of Internalising, Externalising, and the P-factor. *Brain Sciences, 12*(4), 421.

Haywood, D., Lawrence, B. J., Baughman, F. D., & Mullan, B. A. (2021). A Conceptual Model of Long-Term Weight Loss Maintenance: The Importance of Cognitive, Empirical and

Computational Approaches. *International Journal of Environmental Research and Public Health, 18*(2), 635.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1981). Wisconsin Card Sorting Test Manual Psychological Assessment Resources. *Incl Odessa, 521*.

Heeringa, S. G., Wagner, J., Torres, M., Duan, N., Adams, T., & Berglund, P. (2004). Sample designs and sampling methods for the Collaborative Psychiatric Epidemiology Studies (CPES). *International Journal of Methods in Psychiatric Research, 13*(4), 221-240.

Heinrich, M., Geiser, C., Zagorscak, P., Burns, G. L., Bohn, J., Becker, S. P., Eid, M., Beauchaine, T. P., & Knaevelsrud, C. (2020). On the meaning of the general factor of psychopathology ("P-Factor") in symmetrical bifactor models. *Psyarxiv*

Hovenkamp-Hermelink, J. H. M., Riese, H., Batelaan, N. M., Penninx, B. W. J. H., & Schoevers, R. A. (2016). Low stability of diagnostic classifications of anxiety disorders over time: A six-year follow-up of the NESDA study. *Journal of affective disorders, 190*, 310-315.

Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Sage Publications, Inc

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.

International Schizophrenia, C. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature, 460*(7256), 748.

Ip, K. I., Jester, J. M., Sameroff, A., & Olson, S. L. (2019). Linking Research Domain Criteria (RDoC) constructs to developmental psychopathology: The role of self-regulation and

emotion knowledge in the development of internalizing and externalizing growth trajectories from ages 3 to 10. *Development and Psychopathology*, *31*(4), 1557-1574.

Jacobson, L. A., Ryan, M., Martin, R. B., Ewen, J., Mostofsky, S. H., Denckla, M. B., & Mahone, E. M. (2011). Working memory influences processing speed and reading fluency in ADHD. *Child Neuropsychology*, *17*(3), 209-224.

Johnson, B. P., Dayan, E., Censor, N., & Cohen, L. G. (2021). Crowdsourcing in cognitive and systems neuroscience. *The Neuroscientist*, 10738584211017018.

Johnson, D. R. (2009). Emotional attention set-shifting and its relationship to anxiety and emotion regulation. *Emotion*, *9*(5), 681.

Johnson, V. A., Kemp, A. H., Heard, R., Lennings, C. J., & Hickie, I. B. (2015). Childhood-versus adolescent-onset antisocial youth with conduct disorder: Psychiatric illness, neuropsychological and psychosocial function. *PloS One*, *10*(4), e0121627.

Joormann, J., & Gotlib, I. H. (2008). Updating the contents of working memory in depression: interference from irrelevant negative material. *Journal of Abnormal Psychology, 117*(1), 182.

Joormann, J., Yoon, K. L., & Zetsche, U. (2007). Cognitive inhibition in depression. *Applied and Preventive Psychology*, *12*(3), 128-139.

Junghänel, M., Rodenacker, K., Dose, C., & Döpfner, M. (2020). Applying the bifactor S-1 model to ratings of ADHD/ODD symptoms: A commentary on Burns et al.(2019) and a re-analysis. *Journal of Abnormal Child Psychology*, *48*(7), 905-910.

Kail, R. (1991). Development of processing speed in childhood and adolescence. *Advances in Child Development and Behavior*, *23*, 151-185.

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin, 144*(11), 1147.

Kaufman, A. S., & Van Hagen, J. (1977). Investigation of the WISC-R for use with retarded children: Correlation with the 1972 Stanford-Binet and comparison of WISC and WISC-R profiles. *Psychology in the Schools, 14*(1), 10-14.

Kéri, S., & Janka, Z. (2004). Critical evaluation of cognitive dysfunctions as endophenotypes of schizophrenia. *Acta Psychiatrica Scandinavica, 110*(2), 83-91.

Kessler, R. C., & Merikangas, K. R. (2004). The national comorbidity survey replication (NCS-R): background and aims. *International Journal of Methods in Psychiatric Research, 13*(2), 60-68.

Kiehl, K. A., Smith, A. M., Hare, R. D., & Liddle, P. F. (2000). An event-related potential investigation of response inhibition in schizophrenia and psychopathy. *Biological Psychiatry, 48*(3), 210-221.

Kimberg, D. Y., Aguirre, G. K., & D'Esposito, M. (2000). Modulation of task-related neural activity in task-switching: an fMRI study. *Cognitive Brain Research, 10*(1-2), 189-196.

Kleinhans, N., Akshoomoff, N., & Delis, D. C. (2005). Executive functions in autism and Asperger's disorder: flexibility, fluency, and inhibition. *Developmental Neuropsychology, 27*(3), 379-401.

Kofler, M. J., Irwin, L. N., Soto, E. F., Groves, N. B., Harmon, S. L., & Sarver, D. E. (2019). Executive functioning heterogeneity in pediatric ADHD. *Journal of Abnormal Child Psychology, 47*(2), 273-286.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., & Clark, L. A. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*(4), 454.

Kotov, R., Krueger, R. F., Watson, D., Cicero, D. C., Conway, C. C., DeYoung, C. G., Eaton, N. R., Forbes, M. K., Hallquist, M. N., & Latzman, R. D. (2021). The Hierarchical Taxonomy of Psychopathology (HiTOP): A Quantitative Nosology Based on Consensus of Evidence. *Annual Review of Clinical Psychology*, *17*.

Krueger, R. F., & Eaton, N. R. (2015). Transdiagnostic factors of mental disorders. *World Psychiatry*, *14*(1), 27-29.

Kurtz, M. M., & Gerraty, R. T. (2009). A meta-analytic investigation of neurocognitive deficits in bipolar illness: Profile and effects of clinical state. *Neuropsychology, 23*(5), 551.

Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., & Rathouz, P. J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *Journal of Abnormal Psychology*, *121*(4), 971.

Lahey, B. B., Moore, T. M., Kaczkurkin, A. N., & Zald, D. H. (2021a). Hierarchical models of psychopathology: empirical support, implications, and remaining issues. *World psychiatry: Official journal of the World Psychiatric Association (WPA)*, *20*(1), 57-63.

Lahey, B. B., Moore, T. M., Kaczkurkin, A. N., & Zald, D. H. (2021b). Hierarchical models of psychopathology: empirical support, implications, and remaining issues. *World Psychiatry*, *20*(1), 57-63.

Levin-Aspenson, H. F., Watson, D., Clark, L. A., & Zimmerman, M. (2020). What is the general factor of psychopathology? Consistency of the p factor across samples. *Assessment*.

Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological assessment*. Oxford University Press, USA.

Lichtenstein, P., Yip, B. H., Björk, C., Pawitan, Y., Cannon, T. D., Sullivan, P. F., & Hultman, C. M. (2009). Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *The Lancet*, *373*(9659), 234-239.

Lynch, S. J., Sunderland, M., Newton, N. C., & Chapman, C. (2021). A systematic review of transdiagnostic risk and protective factors for general and specific psychopathology in young people. *Clinical Psychology Review*, 102036.

Mackintosh, N. J. (1965). Selective attention in animal discrimination learning. *Psychological Bulletin*, *64*(2), 124.

MacLeod, C. M. (1992). The Stroop task: The" gold standard" of attentional measures. *Journal of Experimental Psychology: General*, *121*(1), 12.

Mahmood, O. M., Goldenberg, D., Thayer, R., Migliorini, R., Simmons, A. N., & Tapert, S. F. (2013). Adolescents' fMRI activation to a response inhibition task predicts future substance use. *Addictive Behaviors*, *38*(1), 1435-1441.

Malcolm, A., Pikoos, T., Castle, D. J., & Rossell, S. L. (2021). Cross-diagnostic cognitive heterogeneity in body dysmorphic disorder and obsessive-compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *73*, 101674.

Manoach, D. S., Lindgren, K. A., Cherkasova, M. V., Goff, D. C., Halpern, E. F., Intriligator, J., & Barton, J. J. S. (2002). Schizophrenic subjects show deficient inhibition but intact task switching on saccadic tasks. *Biological Psychiatry*, *51*(10), 816-826.

Mansell, W., Carey, T. A., & Tai, S. (2012). *A transdiagnostic approach to CBT using method of levels therapy: Distinctive features*. Routledge.

Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., Manfro, G. G., Miguel, E. C., Paus, T., & Bressan, R. A. (2017). A general psychopathology factor (P factor) in children: Structural model analysis and external validation through familial risk and child global executive function. *Journal of Abnormal Psychology, 126*(1), 137.

Martino, D. J., Strejilevich, S. A., Scápola, M., Igoa, A., Marengo, E., Ais, E. D., & Perinot, L. (2008). Heterogeneity in cognitive functioning among patients with bipolar disorder. *Journal of Affective Disorders, 109*(1-2), 149-156.

McManus, F., Shafran, R., & Cooper, Z. (2010). What does a transdiagnostic approach have to offer the treatment of anxiety disorders? *British Journal of Clinical Psychology, 49*(4), 491-505.

McTeague, L. M., Goodkind, M. S., & Etkin, A. (2016). Transdiagnostic impairment of cognitive control in mental illness. *Journal of Psychiatric Research, 83*, 37-46.

Merz, Z. C., Lace, J. W., & Eisenstein, A. M. (2020). Examining broad intellectual abilities obtained within an mTurk internet sample. *Current Psychology*, 1-9.

Miyake, A., Emerson, M. J., & Friedman, N. P. (2000, 2000). Assessment of executive functions in clinical settings: Problems and recommendations. In *Seminars in speech and language* (Vol. 21, No. 02, pp. 0169-0183).

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100.

Mooney, C. Z. (1997). *Monte carlo simulation*. Sage. London: UK.

Moritz, S., Birkner, C., Kloss, M., Jahn, H., Hand, I., Haasen, C., & Krausz, M. (2002). Executive functioning in obsessive–compulsive disorder, unipolar depression, and schizophrenia. *Archives of Clinical Neuropsychology, 17*(5), 477-483.

Murray, A. L., Eisner, M., & Ribeaud, D. (2016). The development of the general factor of psychopathology 'p factor'through childhood and adolescence. *Journal of Abnormal Child Psychology*, *44*(8), 1573-1586.

Newman, D. L., Moffitt, T. E., Caspi, A., & Silva, P. A. (1998). Comorbid mental disorders: implications for treatment and sample selection. *Journal of Abnormal Psychology, 107*(2), 305.

Noël, X., Brevers, D., & Bechara, A. (2013). A neurocognitive approach to understanding the neurobiology of addiction. *Current Opinion in Neurobiology*, *23*(4), 632-638.

Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1-18). Springer, UK.

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19*(6), 625-666.

Nuño, L., Gómez-Benito, J., Carmona, V. R., & Pino, O. (2021). A Systematic Review of Executive Function and Information Processing Speed in Major Depression Disorder. *Brain Sciences, 11*(2), 147.

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance, 17*, 22-27.

Paolillo, E. W., Inkelis, S. M., Heaton, A., Saloner, R., Moore, R. C., & Moore, D. J. (2019). Age of last alcohol use disorder relates to processing speed among older adults living with HIV. *Alcohol and Alcoholism, 54*(2), 139-147.

Park, S.-C., Kim, J.-M., Jun, T.-Y., Lee, M.-S., Kim, J.-B., Yim, H.-W., & Park, Y. C. (2017). How many different symptom combinations fulfil the diagnostic criteria for major depressive disorder? Results from the CRESCEND study. *Nordic Journal of Psychiatry, 71*(3), 217-222.

Patrick, C. J., Venables, N. C., Yancey, J. R., Hicks, B. M., Nelson, L. D., & Kramer, M. D. (2013). A construct-network approach to bridging diagnostic and physiological domains: Application to assessment of externalizing psychopathology. *Journal of Abnormal Psychology, 122*(3), 902.

Pavão, R., Tort, A. B. L., & Amaral, O. B. (2015). Multifactoriality in psychiatric disorders: A computational study of schizophrenia. *Schizophrenia Bulletin, 41*(4), 980-988.

Pernu, T. K. (2019). Elimination, not reduction: Lessons from the Research Domain Criteria (RDoC) and multiple realisation. *Behavioral and Brain Sciences, 42*, 22.

Pettersson, E., Larsson, H., & Lichtenstein, P. (2021). Psychometrics, interpretation and clinical implications of hierarchical models of psychopathology. *World Psychiatry, 20*(1), 68.

Polger, T. W., & Shapiro, L. A. (2016). *The multiple realization book*. Oxford University Press, US.

Putnam, H. (1988). *Representation and reality*. MIT press, US.

Raffard, S., & Bayard, S. (2012). Understanding the executive functioning heterogeneity in schizophrenia. *Brain and Cognition, 79*(1), 60-69.

Reser, M. P., Slikboer, R., & Rossell, S. L. (2019). A systematic review of factors that influence the efficacy of cognitive remediation therapy in schizophrenia. *Australian & New Zealand Journal of Psychiatry, 53*(7), 624-641.

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K.-H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., & Huang, H. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature, 511*(7510), 421-427.

Rodríguez-Sánchez, J. M., Crespo-Facorro, B., González-Blanch, C., Perez-Iglesias, R., & Vázquez-Barquero, J. L. (2007). Cognitive dysfunction in first-episode psychosis: the processing speed hypothesis. *The British Journal of Psychiatry, 191*(S51), s107-s110.

Romer, A. L., & Pizzagalli, D. A. (2021, 2021/10/01/). Is executive dysfunction a risk marker or consequence of psychopathology? A test of executive function as a prospective predictor and outcome of general psychopathology in the adolescent brain cognitive development study. *Developmental Cognitive Neuroscience, 51*, 100994.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of statistical software, 48*(2), 1-36.

Ruggero, C. J., Kotov, R., Hopwood, C. J., First, M., Clark, L. A., Skodol, A. E., Mullins-Sweatt, S. N., Patrick, C. J., Bach, B., & Cicero, D. C. (2019). Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into clinical practice. *Journal of Consulting and Clinical Psychology, 87*(12), 1069.

Rybakowski, J. K. (2019). 120th anniversary of the Kraepelinian dichotomy of psychiatric disorders. *Current Psychiatry Reports, 21*(8), 1-8.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403.

Sauter, M., Draschkow, D., & Mack, W. (2020). Building, Hosting and Recruiting: A Brief Introduction to Running Behavioral Experiments Online. *Brain Sciences, 10*(4), 251.

Schwannauer, M., & Chetwynd, P. (2007). The Brief Symptom Inventory: A validity study in two independent Scottish samples. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice, 14*(3), 221-228.

Shanahan, M. A., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G., Olson, R. K., & DeFries, J. C. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology, 34*(5), 584.

Shevlin, M., McElroy, E., Bentall, R. P., Reininghaus, U., & Murphy, J. (2016). The psychosis continuum: testing a bifactor model of psychosis in a general population sample. *Schizophrenia Bulletin, 43*(1), 133-141.

Shi, D., Maydeu-Olivares, A., & DiStefano, C. (2018). The relationship between the standardized root mean square residual and model misspecification in factor analysis models. *Multivariate Behavioral Research, 53*(5), 676-694.

Simeoli, R., Milano, N., Rega, A., & Marocco, D. (2021). Using Technology to Identify Children With Autism Through Motor Abnormalities. *Frontiers in Psychology*, 1710.

Smith, G. T., Atkinson, E. A., Davis, H. A., Riley, E. N., & Oltmanns, J. R. (2020). The general factor of psychopathology. *Annual Review of Clinical Psychology, 16,* 75-98.

Smith, P. J., Need, A. C., Cirulli, E. T., Chiba-Falek, O., & Attix, D. K. (2013). A comparison of the Cambridge Automated Neuropsychological Test Battery (CANTAB) with "traditional" neuropsychological testing instruments. *Journal of Clinical and Experimental Neuropsychology, 35*(3), 319-328.

Smucny, J., Lesh, T. A., Newton, K., Niendam, T. A., Ragland, J. D., & Carter, C. S. (2018). Levels of cognitive control: a functional magnetic resonance imaging-based test of an RDoC domain across bipolar disorder and schizophrenia. *Neuropsychopharmacology, 43*(3), 598-606.

Snyder, H. R., & Hankin, B. L. (2017). All models are wrong, but the p factor model is useful: Reply to Widiger and Oltmanns (2017) and Bonifay, Lane, and Reise (2017). *Clinical Psychological Science, 5*(1), 187-189.

Snyder, H. R., Miyake, A., & Hankin, B. L. (2015). Advancing understanding of executive function impairments and psychopathology: bridging the gap between clinical and cognitive approaches. *Frontiers in Psychology, 6*, 328.

Soper, D. (2022). *Significance of the Difference between Two Correlations Calculator [Software]*. https://www.danielsoper.com/statcalc

Spector, H. (2013). Long-term IQ stability using the WISC-IV and WAIS-IV among a sample of special education students. https://www.proquest.com/openview/47e0a467cd43a5c885a6b04f24f9d172/1?pq-origsite=gscholar&cbl=18750

Stefanopoulou, E., Manoharan, A., Landau, S., Geddes, J. R., Goodwin, G. U. Y., & Frangou, S. (2009). Cognitive functioning in patients with affective disorders and schizophrenia: a meta-analysis. *International Review of Psychiatry, 21*(4), 336-356.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*(2), 173-180.

Szöke, A., Schürhoff, F., Mathieu, F., Meary, A., Ionescu, S., & Leboyer, M. (2005). Tests of executive functions in first-degree relatives of schizophrenic patients: a meta-analysis. *Psychological Medicine, 35*(6), 771-782.

Tan, E. J., Rossell, S. L., Subotnik, K. L., Ventura, J., & Nuechterlein, K. H. (2021). Cognitive heterogeneity in first-episode psychosis and its relationship with premorbid developmental adjustment. *Psychological Medicine*, 1-10.

Thomas, M. S. C., McClelland, J. L., Richardson, F. M., Schapiro, A. C., & Baughman, F. (2009). Dynamical and connectionist approaches to development: Toward a future of mutually beneficial co-evolution. *Toward a unified theory of development: Connectionism and dynamical systems theory re-considered*, 337-353.

Thöne, A. K., Junghänel, M., Görtz-Dorten, A., Breuer, D., del Giudice, T., Hanisch, C., ... & Döpfner, M. (2022). Empirically based dimensions of externalizing symptoms in children and adolescents: a multitrait-multisource approach. *Journal of Psychopathology and Behavioral Assessment*, *44*(3), 844-861.

Trivedi, J. K. (2006). Cognitive deficits in psychiatric disorders: Current status. *Indian Journal of Psychiatry*, *48*(1), 10.

Tryon, W. W., Hoffman, J., & McKay, D. (2017). Neural network models as explanatory frameworks of psychopathology and its treatment. *Treatments for Psychological Problems and Syndromes*, 452.

Tsourtos, G., Thompson, J. C., & Stough, C. (2002). Evidence of an early information processing speed deficit in unipolar major depression. *Psychological Medicine*, *32*(2), 259-265.

Uittenhove, K., Jeanneret, S., & Vergauwe, E. (2022). From lab-based to web-based behavioural research: Who you test is more important than how you test. *Psyarxiv*.

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior research methods*, *49*(2), 653-673.

Verghese, J., Mahoney, J. R., Ayers, E., Ambrose, A., Wang, C., & Holtzer, R. (2021). Computerised cognitive remediation to enhance mobility in older adults: a single-blind, single-centre, randomised trial. *The Lancet Healthy Longevity*, *2*(9), e571-e579.

Vetter, T. R., & Schober, P. (2018). Regression: the apple does not fall far from the tree. *Anesthesia & Analgesia, 127*(1), 277-283.

Warner, B., & Misra, M. (1996). Understanding neural networks as statistical tools. *The American Statistician, 50*(4), 284-293.

Watts, A. L., Lane, S. P., Bonifay, W., Steinley, D., & Meyer, F. (2020a). Building theories on top of, and not independent of, statistical models: The case of the p-factor. *Psyarxiv*

Watts, A. L., Lane, S. P., Bonifay, W., Steinley, D., & Meyer, F. A. C. (2020b). Building theories on top of, and not independent of, statistical models: The case of the p-factor. *Psychological Inquiry, 31*(4), 310-320.

Watts, A. L., Poore, H. E., & Waldman, I. D. (2019). Riskier tests of the validity of the bifactor model of psychopathology. *Clinical Psychological Science, 7*(6), 1285-1303.

Wei, Y., Womer, F. Y., Sun, K., Zhu, Y., Sun, D., Duan, J., Zhang, R., Wei, S., Jiang, X., & Zhang, Y. (2022). Applying dimensional psychopathology: transdiagnostic prediction of executive cognition using brain connectivity and inflammatory biomarkers. *Psychological Medicine*, 1-11.

WHO. (2002). The alcohol, smoking and substance involvement screening test (ASSIST): development, reliability and feasibility. *Addiction, 97*(9), 1183-1194.

Wilson, K. M., Finkbeiner, K. M., De Joux, N. R., Russell, P. N., & Helton, W. S. (2016). Go-stimuli proportion influences response strategy in a sustained attention to response task. *Experimental Brain Research, 234*(10), 2989-2998.

Wilson, R. C., & Niv, Y. (2012). Inferring relevance in a changing world. *Frontiers in Human Neuroscience, 5*, 189.

Wolinsky, F. D., Vander Weg, M. W., Martin, R., Unverzagt, F. W., Ball, K. K., Jones, R. N., & Tennstedt, S. L. (2009). The effect of speed-of-processing training on depressive symptoms in ACTIVE. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 64*(4), 468-472.

Xie, J., Bi, Q., Shang, W., Yan, M., Yang, Y., Miao, D., & Zhang, H. (2012). Positive and negative relationship between anxiety and depression of patients in pain: a bifactor model analysis. *PloS One, 7*(10), e47577.

Zelazo, P. D. (2015). Executive function: Reflection, iterative reprocessing, complexity, and the developing brain. *Developmental Review, 38*, 55-68.

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1*(2), 198-226.

Zimmerman, I. L., Woo-Sam, J. M., & Glasser, A. J. (1973). *Clinical Interpretation of the Wechsler Adult Intelligence Scale*. Grune & Stratton, US.

Zimmerman, M. (2016). A review of 20 years of research on overdiagnosis and underdiagnosis in the Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) Project. *The Canadian Journal of Psychiatry, 61*(2), 71-79.

Appendices

Appendix A - Ethics Approval Letter for Data Used in Chapter 4 and 5

Research Office at Curtin

GPO Box U1987
Perth Western Australia 6845

Telephone +61 8 9266 7863
Facsimile +61 8 9266 3793
Web research.curtin.edu.au

12-Apr-2021

Name: Karen Heslop
Department/School: Curtin School of Nursing
Email: K.Heslop@curtin.edu.au

Dear Karen Heslop

RE: Amendment approval
Approval number: HRE2021-0105

Thank you for submitting an amendment request to the Human Research Ethics Office for the project **Cognitive Abilities and Mental Health Symptoms**.

Your amendment request has been reviewed and the review outcome is: **Approved**

The amendment approval number is HRE2021-0105-02 approved on 12-Apr-2021.

The following amendments were approved:
Addition of an initial phase (a phase 1) to the current study approval. This is to facilitate current regulatory and practical limitations related to COVID-19 at Cockburn Integrated Health. The additional, phase 1, will occur before data collection at Cockburn Integrated Health and facilitate comparison and inform data interpretation for phase 2 (data collection at Cockburn Integrated Health).

Special Condition of Approval:
It is the responsibility of the Chief Investigator to ensure that any activity undertaken under this project adheres to the latest available advice from the Government or the University regarding COVID-19.

Any special conditions noted in the original approval letter still apply.

Standard conditions of approval

1. Research must be conducted according to the approved proposal
2. Report in a timely manner anything that might warrant review of ethical approval of the project including:
   - proposed changes to the approved proposal or conduct of the study
   - unanticipated problems that might affect continued ethical acceptability of the project
   - major deviations from the approved proposal and/or regulatory guidelines
   - serious adverse events
3. Amendments to the proposal must be approved by the Human Research Ethics Office before they are implemented (except where an amendment is undertaken to eliminate an immediate risk to participants)
4. An annual progress report must be submitted to the Human Research Ethics Office on or before the anniversary of approval and a completion report submitted on completion of the project
5. Personnel working on this project must be adequately qualified by education, training and experience for their role, or supervised
6. Personnel must disclose any actual or potential conflicts of interest, including any financial or other interest or affiliation, that bears on this project
7. Changes to personnel working on this project must be reported to the Human Research Ethics Office
8. Data and primary materials must be retained and stored in accordance with the Western Australian University Sector Disposal Authority (WAUSDA) and the Curtin University Research Data and Primary Materials policy

9. Where practicable, results of the research should be made available to the research participants in a timely and clear manner
10. Unless prohibited by contractual obligations, results of the research should be disseminated in a manner that will allow public scrutiny; the Human Research Ethics Office must be informed of any constraints on publication
11. Ethics approval is dependent upon ongoing compliance of the research with the Australian Code for the Responsible Conduct of Research, the National Statement on Ethical Conduct in Human Research, applicable legal requirements, and with Curtin University policies, procedures and governance requirements
12. The Human Research Ethics Office may conduct audits on a portion of approved projects.

Should you have any queries regarding consideration of your project, please contact the Ethics Support Officer for your faculty or the Ethics Office at hrec@curtin.edu.au or on 9266 2784.

Yours sincerely

Associate Professor Sharyn Burns
Chair, Human Research Ethics Committee

Appendix B - Brief Symptom Inventory (Derogatis & Melisaratos, 1983) Used in Chapters 4 and 5

Brief Symptom Inventory

Below are a list of problems people sometimes have. Please indicate HOW MUCH THAT PROBLEM HAS DISTRESSED OR BOTHERED YOU DURING THE PAST 7 DAYS INCLUDING TODAY.

0 = Not at all

1 = A little bit

2 = Moderately

3 = Quite a bit

4 = Extremely

R = Refused

DURING THE PAST 7 DAYS, how much were you distressed by:

1. Nervousness or shakiness inside	0 1 2 3 4 R
2. Faintness or dizziness	0 1 2 3 4 R
3. The idea that someone else can control your thoughts	0 1 2 3 4 R
4. Feeling others are to blame for most of your troubles	0 1 2 3 4 R
5. Trouble remembering things	0 1 2 3 4 R
6. Feeling easily annoyed or irritated	0 1 2 3 4 R
7. Pains in the heart or chest	0 1 2 3 4 R
8. Feeling afraid in open spaces	0 1 2 3 4 R
9. Thoughts of ending your life	0 1 2 3 4 R

DURING THE PAST 7 DAYS, how much were you distressed by:

10. Feeling that most people cannot be trusted	0 1 2 3 4 R
11. Poor appetite	0 1 2 3 4 R

12. Suddenly scared for no reason 0 1 2 3 4 R

13. Temper outbursts that you could not control 0 1 2 3 4 R

14. Feeling lonely even when you are with people 0 1 2 3 4 R

15. Feeling blocked in getting things done 0 1 2 3 4 R

16. Feeling lonely 0 1 2 3 4 R

17. Feeling blue 0 1 2 3 4 R

18. Feeling no interest in things 0 1 2 3 4 R

DURING THE PAST 7 DAYS, how much were you distressed by:

19. Feeling fearful 0 1 2 3 4 R

20. Your feelings being easily hurt 0 1 2 3 4 R

21. Feeling that people are unfriendly or dislike you 0 1 2 3 4 R

22. Feeling inferior to others 0 1 2 3 4 R

23. Nausea or upset stomach 0 1 2 3 4 R

24. Feeling that you are watched or talked about by others 0 1 2 3 4 R

25. Trouble falling asleep 0 1 2 3 4 R

26. Having to check and double check what you do 0 1 2 3 4 R

27. Difficulty making decisions 0 1 2 3 4 R

DURING THE PAST 7 DAYS, how much were you distressed by:

28. Feeling afraid to travel on buses, subways, or trains 0 1 2 3 4 R

29. Trouble getting your breath 0 1 2 3 4 R

30. Hot or cold spells 0 1 2 3 4 R

31. Having to avoid certain things, places, or activities because they frighten you 0 1 2 3 4 R

32. Your mind going blank 0 1 2 3 4 R

33. Numbness or tingling in parts of your body 0 1 2 3 4 R

34. The idea that you should be punished for your sins 0 1 2 3 4 R

35. Feeling hopeless about the future 0 1 2 3 4 R

36. Trouble concentrating 0 1 2 3 4 R

DURING THE PAST 7 DAYS, how much were you distressed by:

37. Feeling weak in parts of your body 0 1 2 3 4 R
38. Feeling tense or keyed up 0 1 2 3 4 R
39. Thoughts of death or dying 0 1 2 3 4 R
40. Having urges to beat, injure, or harm someone 0 1 2 3 4 R
41. Having urges to break or smash things 0 1 2 3 4 R
42. Feeling very self-conscious with others 0 1 2 3 4 R
43. Feeling uneasy in crowds 0 1 2 3 4 R
44. Never feeling close to another person 0 1 2 3 4 R
45. Spells of terror or panic 0 1 2 3 4 R

DURING THE PAST 7 DAYS, how much were you distressed by:

46. Getting into frequent arguments 0 1 2 3 4 R
47. Feeling nervous when you are left alone 0 1 2 3 4 R
48. Others not giving you proper credit for your achievements 0 1 2 3 4 R
49. Feeling so restless you couldn't sit still 0 1 2 3 4 R
50. Feelings of worthlessness 0 1 2 3 4 R
51. Feeling that people will take advantage of you if you let them 0 1 2 3 4 R
52. Feeling of guilt 0 1 2 3 4 R
53. The idea that something is wrong with your mind 0 1 2 3 4 R

Appendix C - The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) V3.1 (WHO, 2002)

Question 1. In your life, which of the following substances have you ever used (non-medical use only)?					
a. Tobacco products (cigarettes, chewing tobacco, cigars, etc.)	No	Yes	-	-	-
b. Alcoholic beverages (beer, wine, spirits, etc.) No Yes	No	Yes	-	-	-
c. Cannabis (marijuana, pot, grass, hash, etc.)	No	Yes	-	-	-
d. Cocaine (coke, crack, etc.) No Yes	No	Yes	-	-	-
e. Amphetamine-type stimulants (speed, meth, ecstasy, etc.)	No	Yes	-	-	-
f. Inhalants (nitrous, glue, petrol, paint thinner, etc.)	No	Yes	-	-	-
g. Sedatives or sleeping pills (diazepam, alprazolam, flunitrazepam, midazolam, etc.)	No	Yes	-	-	-
h. Hallucinogens (LSD, acid, mushrooms, trips, ketamine, etc.)	No	Yes	-	-	-
i. Opioids (heroin, morphine, methadone, buprenorphine, codeine, etc.)	No	Yes	-	-	-

j. Other – specify_____	No	Yes	-	-	-
Question 2. In the *past three months*, how often have you used the substances you mentioned (first drug, second drug, etc)?					
a. Tobacco products (cigarettes, chewing tobacco, cigars, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
b. Alcoholic beverages (beer, wine, spirits, etc.) No Yes	Never	Once or twice	Monthly	Weekly	Daily or almost daily
c. Cannabis (marijuana, pot, grass, hash, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
d. Cocaine (coke, crack, etc.) No Yes	Never	Once or twice	Monthly	Weekly	Daily or almost daily
e. Amphetamine-type stimulants (speed, meth, ecstasy, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
f. Inhalants (nitrous, glue, petrol, paint thinner, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
g. Sedatives or sleeping pills (diazepam, alprazolam, flunitrazepam, midazolam, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
h. Hallucinogens (LSD, acid, mushrooms, trips, ketamine, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
i. Opioids (heroin, morphine, methadone, buprenorphine, codeine, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily

j. Other – specify_____	Never	Once or twice	Monthly	Weekly	Daily or almost daily
Question 3. During the *past three months*, how often have you had a strong desire or urge to use (first drug, second drug, etc)?					
a. Tobacco products (cigarettes, chewing tobacco, cigars, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
b. Alcoholic beverages (beer, wine, spirits, etc.) No Yes	Never	Once or twice	Monthly	Weekly	Daily or almost daily
c. Cannabis (marijuana, pot, grass, hash, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
d. Cocaine (coke, crack, etc.) No Yes	Never	Once or twice	Monthly	Weekly	Daily or almost daily
e. Amphetamine-type stimulants (speed, meth, ecstasy, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
f. Inhalants (nitrous, glue, petrol, paint thinner, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
g. Sedatives or sleeping pills (diazepam, alprazolam, flunitrazepam, midazolam, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
h. Hallucinogens (LSD, acid, mushrooms, trips, ketamine, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
i. Opioids (heroin, morphine, methadone, buprenorphine, codeine, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily

j. Other – specify_____	Never	Once or twice	Monthly	Weekly	Daily or almost daily
Question 4. During the *past three months*, how often has your use of (first drug, second drug, etc) led to health, social, legal or financial problems?					
a. Tobacco products (cigarettes, chewing tobacco, cigars, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
b. Alcoholic beverages (beer, wine, spirits, etc.) No Yes	Never	Once or twice	Monthly	Weekly	Daily or almost daily
c. Cannabis (marijuana, pot, grass, hash, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
d. Cocaine (coke, crack, etc.) No Yes	Never	Once or twice	Monthly	Weekly	Daily or almost daily
e. Amphetamine-type stimulants (speed, meth, ecstasy, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
f. Inhalants (nitrous, glue, petrol, paint thinner, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
g. Sedatives or sleeping pills (diazepam, alprazolam, flunitrazepam, midazolam, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
h. Hallucinogens (LSD, acid, mushrooms, trips, ketamine, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
i. Opioids (heroin, morphine, methadone, buprenorphine, codeine, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily

j. Other – specify_____	Never	Once or twice	Monthly	Weekly	Daily or almost daily
Question 5. During the *past three months*, how often have you failed to do what was normally expected of you because of your use of (first drug, second drug, etc)?					
a. Tobacco products (cigarettes, chewing tobacco, cigars, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
b. Alcoholic beverages (beer, wine, spirits, etc.) No Yes	Never	Once or twice	Monthly	Weekly	Daily or almost daily
c. Cannabis (marijuana, pot, grass, hash, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
d. Cocaine (coke, crack, etc.) No Yes	Never	Once or twice	Monthly	Weekly	Daily or almost daily
e. Amphetamine-type stimulants (speed, meth, ecstasy, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
f. Inhalants (nitrous, glue, petrol, paint thinner, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
g. Sedatives or sleeping pills (diazepam, alprazolam, flunitrazepam, midazolam, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
h. Hallucinogens (LSD, acid, mushrooms, trips, ketamine, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily
i. Opioids (heroin, morphine, methadone, buprenorphine, codeine, etc.)	Never	Once or twice	Monthly	Weekly	Daily or almost daily

j. Other – specify_____	Never	Once or twice	Monthly	Weekly	Daily or almost daily
Question 6. Has a friend or relative or anyone else *ever* expressed concern about your use of (first drug, second drug, etc)?					
a. Tobacco products (cigarettes, chewing tobacco, cigars, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
b. Alcoholic beverages (beer, wine, spirits, etc.) No Yes	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
c. Cannabis (marijuana, pot, grass, hash, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
d. Cocaine (coke, crack, etc.) No Yes	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
e. Amphetamine-type stimulants (speed, meth, ecstasy, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
f. Inhalants (nitrous, glue, petrol, paint thinner, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
g. Sedatives or sleeping pills (diazepam, alprazolam, flunitrazepam, midazolam, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
h. Hallucinogens (LSD, acid, mushrooms, trips, ketamine, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
i. Opioids (heroin, morphine, methadone, buprenorphine, codeine, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-

j. Other – specify_____	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
Question 7. Have you *ever* tried to cut down on using (first drug, second drug, etc) but failed?					
a. Tobacco products (cigarettes, chewing tobacco, cigars, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
b. Alcoholic beverages (beer, wine, spirits, etc.) No Yes	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
c. Cannabis (marijuana, pot, grass, hash, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
d. Cocaine (coke, crack, etc.) No Yes	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
e. Amphetamine-type stimulants (speed, meth, ecstasy, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
f. Inhalants (nitrous, glue, petrol, paint thinner, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
g. Sedatives or sleeping pills (diazepam, alprazolam, flunitrazepam, midazolam, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
h. Hallucinogens (LSD, acid, mushrooms, trips, ketamine, etc.)	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
i. Opioids (heroin, morphine, methadone,	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-

buprenorphine, codeine, etc.)					
j. Other – specify_____	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-
Question 8. Have you *ever* used any drug by injection (non-medical use only)?					
	No, never	Yes, in the past 3 months	Yes, but not in the past 3 months	-	-

Appendix D - Copyright Information For Published Articles

“For all articles published in MDPI journals, copyright is retained by the authors. Articles are licensed under an open access Creative Commons CC BY 4.0 license, meaning that anyone may download and read the paper for free. In addition, the article may be reused and quoted provided that the original published version is cited. These conditions allow for maximum use and exposure of the work, while ensuring that the authors receive proper credit.

In exceptional circumstances articles may be licensed differently. If you have specific condition (such as one linked to funding) that does not allow this license, please mention this to the editorial office of the journal at submission. Exceptions will be granted at the discretion of the publisher.”

https://www.mdpi.com/authors/rights