UNIVERSITY OF DERBY

The Dark Side of Humanity Scale: A Reconstruction of the Dark Tetrad Constructs

Louise F. Katz

Doctor of Philosophy March 2021

Table of Contents

List of Tables

List of Figures

Preface

Declaration

This work presented within this thesis, encompassing all written work, research and data analyses, are entirely my own and produced as the result of my own original research. Advice and guidance was provided by my supervisory team. All research studies received ethical approval before commencement, from the University of Derby Ethics Committee.

Acknowledgements

This may be the hardest part of the thesis to write as it marks the end of a very long journey – one that I never dreamt I would make. From starting my academic career at the age of 44 with the Open University, to completing my MSc with the University of Derby Online Learning and finally nearing the end of my PhD with the University of Derby, it has been a somewhat lonely road. Yet, what I have learnt about myself as well as the vast range of skills along the way, is inexpressible through words alone. What this journey has shown me, is that no matter what anyone says to you, if you really want to achieve something and put your heart and soul into it, the impossible becomes the possible.

I have been amazingly lucky with my three wonderful supervisors, Caroline, Ian and Chris. I don't think they truly know how much their support and guidance has pushed me through these past years. How their suggestions, feedback and questioning of my decisions made me re-think directions I was planning to go in (which would have been the wrong way). Their support through the rocky times, of which, luckily there were few and their encouragement to keep pushing through. Thank you will never be enough. One thing I would say however, is that they were always right.

To my husband, Stephen, his patience has had no limits. His life with me has changed considerably over these years. He has put up with my stress induced moods, my preoccupation with my studies and constant staring at the computer screen. Yet he has always encouraged me and pushed me forward. Indeed, it was he who taught me the definition of insanity, when things weren't going quite to plan, where I was doing the same thing over and over hoping for a different outcome. That is one expression I will never forget. He has allowed me to chase this dream, he has stood by me through the highs and lows, words of thanks can never do that justice.

To my daughters, Ilana and Tania, who are very grateful they no longer live at home and have to listen to my ramblings. I think they have watched in bewilderment as to why I would put myself through this. Yet at each graduation they have looked proudly on. I need to thank them for their patience for all the times I wasn't as present as they would have liked me to be.

Finally, to my dogs, Coco who is sadly no longer with me, Odgie, Saffie, Mylo and Millie. They have made me go out for long walks to clear my head and give my eyes a rest. They have listened to my endless discussions with them about statistics, theories, ethics applications and whether a paragraph reads well, without passing judgement. They have known when to stay quiet so I can concentrate and when I've been stressed, to sit with me for a cuddle. Each and every one of them has been integral to this journey and are inextricably entwined with it.

Louise Katz
March 2021

Abstract

A myriad of criticisms have been directed at the widely available scales which measure the Dark Tetrad (DT), constructs of Machiavellianism, psychopathy, narcissism and everyday sadism. These have included construct incongruence, unstable factor structures, sex and age variance. Sex and age are salient variables in research studies and as such, their reported variance was considered to be a weakness of the measures, which contributed to their instability. To ascertain whether this was indeed the case, an investigation of the widely available DT measures was conducted in study one (*n*=605). Through the use of Classical Test Theories, Exploratory Graph Analysis and Item Response Theory, the criticisms of past research were supported, including variance across age and sex. Mediated by the evidenced issues and in order to address them, the Dark Side of Humanity Scale (DSHS), was developed, which was the main aim of this thesis. The focus of the measure was to assesses the DT personality constructs as they manifest in the general population, whilst also being sex and age invariant. Using a diverse range of statistical methods and guided by theory, expert ratings and past research, where available, the development of the DSHS began in the second study (*n*=667). During analyses, a divergence from the widely available DT measures emerged, whereby primary psychopathy and Machiavellianism were unified. However, this corroborated past research which has discussed the two constructs as being parallel. It further afforded the DSHS with a shift away from the traditional DT conceptualisation. The resulting scale encompasses four factors. The first represents the successful psychopath, factor two addresses the grandiose form of entitlement, factor three taps into everyday sadism and includes sadistic fantasies, direct and vicarious psychological malevolence, whilst factor four pertains to narcissistic entitlement rage. Each factor is specific to the traits and behaviours of the construct they address. Study three reports on the convergent and discriminant validity of the DSHS (*n*=712), and test-retest validity (*n* = 413), for which temporal reliability as well as nomological validity was achieved. The unique contributions of this thesis are discussed in the final chapter. The DSHS provides an alternative approach to investigating the dark side of human nature in general population samples, whilst also being sex and age invariant.

Chapter One

Introduction

This thesis began with a spark, ignited through discrepancies in the literature, which fuelled the flames to conduct further exploration. The origins of the current investigations were prompted by research pertaining to psychopathy found within institutionalised female samples (e.g., Lehmann & Ittel, 2012; Warren, South, Burnette et al., 2005). The tool used to measure psychopathy in these samples was the Psychopathy Checklist-Revised (PCL-R; Hare, Harpur et al., 1990), which was initially developed for use with incarcerated males. Research showed that 25% of male inmates scored above the PCL-R cut-off of 30 (Strand & Belfrage, 2005), which is the criteria for clinical psychopathy (Hare, Hart & Harpur, 1991). Conversely, only 6% to 17% of females achieved or exceeded this cut-off point (Vitale et al., 2002; Warren & South 2006). To address this, a lower cut-off score of 25 was used for females (Coid & Ulrich, 2010; Lehmann & Ittel, 2012; Weizmann-Henelius et al., 2004). Yet, across studies, this lower cut-off score varied, thereby yielding conflicting results (de Vogel et al., 2016; Jackson, Rogers et al., 2002; Logan & Blackburn, 2009). It was therefore questioned as to whether the traits and behaviours of males could meaningfully be analogised to females (Vitale et al., 2002). It was further deliberated whether consideration of the similarities and differences between the two sexes (as defined by the biological differences between males and females, (APA, 2015)), had been given to measures which assess dark personality constructs in general population samples (Anestis et al., 2011).

This led to the investigation of the four malevolent or dark personality traits which are found under the umbrella of the Dark Tetrad (DT; Paulhus, 2014), Machiavellianism, psychopathy, narcissism and everyday sadism. These constructs will be discussed in further depth in the ensuing chapters but in brief; Machiavellianism is characterised by a cynical world view whereby individuals embodying high levels of the construct are conniving, manipulative, long-term strategists, with an inclination to use others through nefarious means to achieve their goals (Christie & Geis,1970; Collison, South et al., 2020; Jones & Paulhus, 2009). Those with primary psychopathic traits are callous and egocentric, who through manipulation and deception, enacted without remorse, cause significant anguish within interpersonal relationships (Miller, Sleep et al., 2020; Sellbom & Drislane, 2020). Conversely, individuals who have secondary psychopathy propensities,

enact impulsive antisocial behaviours. These may be influenced by negative affectivity and external comorbidities, including drug and alcohol addiction (Sellbom & Drislane, 2020). In relation to narcissism, the literature consistently finds evidence for two facets, grandiose and vulnerable narcissism (Brailovskaia, Rohmann et al., 2020; Dickinson & Pincus, 2003; Gore & Widiger, 2016; Jauk et al., 2017). Grandiose narcissism is characterised by entitled exploitation, inflated self-esteem and a disproportionate need for power (Corry et al., 2008; Green et al., 2020; Krizan & Herlache, 2018; Wright & Edershile, 2018), whilst vulnerable narcissism addresses hypersensitivity, low self-esteem, entitled resentment and rage (Kealy, Laverdière & Pincus, 2020). Everyday sadism relates to individuals who have dispositional tendencies to fantasise about, watch or directly inflict psychological and/or physical pain and humiliation on others, whilst taking pleasure from their suffering (Kowalksi, Di Pierro et al., 2019; Lui et al., 2020). Any feelings of guilt are discarded through rationalisation of their callous behaviour (Buckels, Jones & Paulhus, 2013; Buckels et al., 2019; O'Meara et al., 2011; Trémolière & Djeriouat, 2016). An overlap amongst these constructs has been attributed to a common antagonistic core of disagreeableness (Paulhus & Williams, 2002), callousness and manipulation (Jones & Figuerdo, 2013), which facilitate exploitative interpersonal relationships (Kajonius et al., 2016).

A salient aspect which underlies the DT constructs is their theoretical origins. Psychopathy, narcissism and everyday sadism are historically grounded in the traits and behaviours of mainly institutionalised males (e.g., Davies & Hand, 2003; Hare, Harpur et al., 1990; Raskin & Hall, 1979), which will be further discussed in Chapter Two. Machiavellianism, although emerging from a social psychological aetiology, is primarily based on the writings of Machiavelli, a male political advisor from the sixteenth century (Christie & Geis, 1970). These male oriented traits and behaviours have been incorporated within the 'gold standard' measures and alludes to the consideration that they have been imposed on women through the scales used in general population research, without due consideration of sex specific manifestations, which will be discussed further in Chapter Four.

Yet, it is not just females who have been marginalised in the development of DT measures but also individuals who fall outside of the student age range of 18-25 (Barlett & Barlett, 2015). Many of the DT scales used young adult undergraduate student samples during the development process, whom, for many researchers, are

convenient (Carter, Campbell et al., 2015; Trzesniewski & Donnellan, 2010). However, some studies incorporate participants across a broad range of ages and evidence has shown that the factor structures of measures and item endorsements have differed between more mature adult samples from those of students (Carter, Campbell et al., 2015; James et al., 2014; Moss, 2005). It is therefore contended that the widely available scales which measure the DT constructs may have neglected the fundamental requirement of sex and age equivalence, whereby at the minimum, factor structures should be comparable (Brinkley et al., 2008; Forouzan & Cooke, 2005).

Consequently, the initial focus is to provide evidence to ascertain whether the widely available DT scales are variant between and amongst the groups of interest: sex and age, whereby age pertains to all adults over the age of 18. The samples will be separated into two groups. The first being those aged 18-25, who are reflective of students and determined as emerging adults (Barlett & Barlett, 2015), the second will be adults aged 26 and over. The rationale for the groups is discussed in Chapter Four. A lack of invariance would suggest that intergroup mean differences cannot be meaningfully compared in research studies (Shchebetenko et al., 2020).

The main aim of this thesis is to develop a new psychometric measure, The Dark Side of Humanity Scale, using a diverse range of statistical methods. These will elucidate how the DT constructs manifest invariantly over age and sex, within the general population.

The structure of this thesis is as follows:

Chapter Two provides a systematic literature review in relation to the DT traits of psychopathy, Machiavellianism, narcissism and everyday sadism. The focus of the review is the factor structures, scale indicators sex and/or age invariance of the widely available scales used in DT research. It further seeks to ascertain whether the measures adequately tap into the constructs they seek to measure.

Chapter Three, the methodology chapter, addresses the methods used in the ensuing studies and provides a discussion of Classical Test Theories, Exploratory Graph Analysis and Item Response Theories.

Chapter Four details the first study which seeks to investigate the age and sex invariance of the widely available scales which are commonly used in DT research. Psychometric scrutiny will determine whether the measures provide robust intergroup generalisability.

Chapter Five incorporates the second study and begins with a discussion in relation to the core traits and behaviours of each of the DT constructs. From this, definitions for each construct are established based on theory, expert ratings and previous research, which provides the framework for a pool of items to begin development on the Dark Side of Humanity Scale.

Chapter Six focuses on the third study, which is the validation of the Dark Side of Humanity Scale, investigating convergent and discriminant validity along with temporal reliability through test-retest.

This then leads into Chapter Seven, which draws upon the findings from the studies and discusses the theoretical implications, psychometric considerations and unique contributions that this thesis provides to the field of dark personality research. Suggestions are offered for future research along with consideration of limitations.

Chapter Two

An Interrogation of the Structures, Sex and Age Invariance of the Dark Tetrad Scales: A Systematic Review

The goal of any psychometric scale is to measure a latent construct as precisely as possible (Clark & Watson,1995), as well as homogenously across groups. However, scales which measure the Dark Tetrad (DT) personality constructs of sub-clinical psychopathy, sub-clinical narcissism, Machiavellianism and everyday sadism (e.g., Mach IV; Christie & Geis, 1970), have been criticised by many researchers for their unstable factor structures and construct incongruence. Structural and indicator variance has also been found for males and females across ages (Corry et al., 2008; Dotterer et al., 2017; Gummelt et al., 2012; Kawamoto et al., 2020; McHoskey et al., 1998; Persson, 2019; Persson, Kajonius & Garcia, 2017a, 2017b). The focus of this review will be to consider the evidence provided for these criticisms, to ascertain the extent the scales align with the constructs they seek to measure, the variation in factor structures across studies and whether the instability of factor structures may be mediated by a lack of sex and/or age invariance. The implications from the latter point would suggest that the scales may not measure the constructs equivalently across the specified groups.

Results have consistently found through the DT measures, that males score higher means than females across the four constructs (Jonason & Davis, 2018; Paulhus, Buckels et al., 2020). This has been explained through an evolutionary viewpoint in terms of life history strategies. It is discussed that males with dark personality traits adopt a “fast life” history strategy, whereby they embody characteristics which are analogous with an antisocial world view, including impulsivity, a short-term mating strategy and egoism (Jonason, Koenig & Tost, 2010). Conversely, females are suggested to follow a “slower life” history strategy, due to their dependence upon social structures and secure attachments (Olderbak & Figuerdo, 2009), and a preference for long-term planning and long-term mating (Gladden et al., 2009). They are considered reluctant to take risks, as doing so may come at a significant social cost (Buss & Duntley, 2008; Jonason, Foster et al., 2018; Jonason & Webster 2010). The social cost for females can be exemplified by those who seek leadership positions and are perceived as agentic. These women may be devalued comparatively to males, as they are seen to be encroaching on traditional

male dominions by assuming male-stereotypic leadership approaches or usurping male-dominated leadership roles (Eagly, Makhijani & Klonsky, 1992). A failure by women to conform to their sex role stereotypes is seen as a violation of what females should be and as a consequence, they face social backlash. Thus, agency is put forth as advantageous for males and disadvantageous for females (Rudman, Moss-Racusin et al., 2010). Accordingly, females who embody dark personality traits may be in direct conflict with the stereotypical assumptions imposed on them.

Stereotypical associations may be entwined within the male oriented aetiologies of the DT constructs, from which the general population measures emerged. Psychopathic traits were originally determined by male clinical psychiatrists based on the behaviours of predominantly male patients (Cleckley, 1941; Karpman 1948). Cleckley (1941), conceptualised psychopathy through the presence of emotional shallowness, grandiosity, callousness and a lack of remorse. He argued that many individuals with psychopathic traits had no history of antisocial behaviours and can function successfully in society. Yet, an alternative behavioural view put forth, stated that psychopathy should be assessed through a history of antisocial misconduct (Cloninger, 1978; Robins, 1966). When developing the Psychopathy Checklist and Psychopathy Checklist Revised (PCL; PCL-R; Hare, 1980; Hare, Harpur et al., 1990), for the primary purpose of measuring the psychopathic traits of incarcerated males (Harpur, Hakstian & Hare, 1988), the authors incorporated both the personality and behavioural facets (Salekin, Rogers et al., 1998), with the latter based on characteristics used in clinical assessments (Hare, 1980). However, It was found in relation to females, that certain items from the PCL-R were not directly relevant to females (Rutherford et al., 1996; Strachan et al., 1990).

Despite the lack of intersex homogeneity, when the construct metamorphosed from the clinical realm into the social psychological domain, the foundations of the male oriented PCL-R, including the antisocial and criminality facets, were preserved in psychometric measures (Williams, Nathanson & Paulhus, 2003). There are two considerations resulting. Firstly, that the traits and behaviours chosen to measure institutionalised males has been analogised to females, without further consideration and separate analysis of whether the female manifestation differs. Secondly, that the entwinement of two competing theories (Cleckley, 1941, Robins, 1966), mediated a construct drift, which has since permeated social psychological research. Indeed, initial research of sub-clinical psychopathy was minimal, which may have led to a

distorted understanding of the core of psychopathy and consequently a spurious definition of the construct (Boduszek, Dhingra, et al., 2016; Boduszek & Debowska, 2016; Cooke & Logan, 2015; Skeem & Cooke, 2010). The literature has contended that sub-clinical psychopathy, as based on clinical and forensic samples, yielded a disproportionate inclusion of criminal and antisocial behaviours, incorporated within measures of the secondary psychopathy facet (Boduszek, Dhingra, et al., 2016; Boduszek & Debowska, 2016).

Turning to narcissism, historically, the most prevalent measure has been the Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979), in which, the traits and behaviours of males are interwoven. The scale was developed based on the DSM-III's (APA, 1980), definition of Narcissistic Personality Disorder (NPD), which was mainly derived from male patients (Morf et al., 2000). Research has shown that the manifestation of NPD differs in males and females, whereby males respond with higher levels of impulsivity to negative affect, whilst females may react with a higher degree of self-focus. A lack of empathy was also seen at lower levels of the disorder for males more so than females. Importantly, seven out of the nine criteria for NPD were found to be sex invariant, including, interpersonal exploitation, entitlement, grandiosity and fantasies of power. Yet, these characteristics have been found to differ in general population samples (Hoertel et al., 2018; Grijalva et al., 2015). This has been attributed to assessments of narcissism which are based on self-report. These measures may mediate sex differences in how a lack of empathy is encapsulated, which may lead some researchers to assume sex role stereotypes (Baez et al., 2017).

This latter point is highlighted by narratives in research, which discuss that the traits and behaviours measured by the NPI may be more relevant for males rather than females. The NPI measures grandiose narcissism (Brown, Stanton & Watson, 2020; Krizan & Herlache, 2018), and incorporates agentic and adaptive characteristics of exhibitionism, entitlement, leadership and authority which are mediated by superiority and interpersonal antagonism (Rosenthal & Hooley, 2010). Facets of grandiose narcissism have been put forth as incompatible with female stereotypical associations because they violate culturally held Western expectations of female behaviours (Atlas & Them, 2008; Corry et al., 2008; Tschanz et al., 1998; Weiser, 2015). As a narcissism measure, the NPI minimises the vulnerable facet, which encompasses characteristics of introversion, hypersensitivity, defensiveness, rage and

anxiety (Pincus, Ansell et al., 2009; Wink, 1991). Studies have shown the NPI to have an unsatisfactory model fit and divergent factor structures with mixed sex samples. This may be attributed to the impact of the female data, or the scale may not tap into how the construct manifests across sexes (Corry et al, 2008 ; Kubarych et al., 2004).

A further dark personality construct which has been transmuted from the clinical domain into the social psychological realm and also criticised for holding male associations is sadism. Sadistic Personality Disorder (SPD), was suggested for inclusion in the DSM-III-R (APA, 1987), by psychiatrists and consequently categorised as 'Proposed Diagnostic Categories Needing Further Study'. Following the modelling of a prototypic male and female design to diagnose SPD, it emerged that the disorder was strongly aligned to stereotypical male behaviours, and as such the diagnostic criteria was determined as sex biased (Fiester & Gay, 1991). To corroborate this, research had shown a significant sex-bias, with most individuals with SPD being male (Myers et al., 2006).

Yet, the DSM-III-R SDP diagnostic characteristics were incorporated into social psychological general population measures, thereby perpetuating the identified male bias, beginning with the Sadistic Attitudes and Behaviour Scale (SABS; Davies & Hand, 2003). Ensuing measures have been developed which address sadistic fantasies, vicarious, psychological and direct sadism (O'Meara et al., 2011; Paulhus & Jones, 2015; Plouffe, Saklofske & Smith, 2017). Findings from studies using these measures have shown that both males and females embody the construct, although females invariably score lower means than males (Gonzalez & Greitemeyer, 2018; Kowalski, Di Pierro et al., 2019). To date, there is minimal research which has investigated sex invariance of the measures (Plouffe, Smith & Saklofske, 2018), to ascertain whether the scales have shifted away from the original male stereotypical characteristics.

Machiavellianism is the only DT construct which originates from a social psychological aetiology. Its theoretical underpinnings stem from focus group discussions between Christie and his colleagues (Christie & Geis, 1970), motivated by the writings of Machiavelli, a male political advisor in the 1500's, primarily his book 'The Prince'. The deliberations formulated the ethos of the construct, which was the investigation of the politically oriented individual, who embodies misanthropy, cynicism, deceit and manipulation. Consequently, the Mach IV was developed and comprises third person statements drawn from 'The Prince' (Christie & Geis, 1970).

Intersex differences were present from the outset, which mediated qualitative inequalities, for example, ‘Most men are brave’. Further sex inequalities were found, whereby female manipulative strategies were not being assessed by the Mach IV. This led to the suggestion that the scale may not be compatible with the female manifestation of Machiavellianism (Brown & Guy, 1983).

Machiavellianism as a construct has also been subject to criticism. Not only has it been suggested as invalid (e.g., Ray, 1983), but it has also been shown as empirically and conceptually similar enough to psychopathy, whereby the two constructs are indistinct (Glenn & Sellbom, 2015; McHoskey et al., 1998; Miller, Hyatt, et al., 2017; Persson, 2019). Researchers have contended that Machiavellianism is a global measure of psychopathy (McHoskey et al., 1998), whereby existing studies may be better reinterpreted as pertaining to psychopathy (Miller, Hyatt et al., 2017). Using the bass-ackwards method, Persson (2019), elucidated that Machiavellianism, at a hierarchical level converged with psychopathy and narcissism models more so than with that of Machiavellianism. The construct has been found to negatively correlate with variables such as cautiousness and positively with impulsivity (DeShong et al., 2017), suggesting that the measure may be incongruent with its theoretical description (Christie & Geis, 1970).

Further widely available brief measures used in dark personality research are the Dirty Dozen (DD: Jonason & Webster, 2010) and Short Dark Triad (SD3: Jones & Paulhus, 2014). These scales have minimised the constructs of psychopathy, Machiavellianism and narcissism, to measure the Dark Triad. However, their diminished content has brought extensive criticism, whereby the depths and nuances of the constructs have been negated (e.g., Kajonius et al., 2016). Development of the DD was partially guided by the consistent evidence that men score higher on all three constructs (Jonason & Webster, 2010). This mediated a prediction and validation criteria that men *should* score higher on all three constructs (Jonason & Webster, 2010). Related to the SD3, the constructs as defined by the researchers diverged from that of the DD. Psychopathy was determined by the secondary facet, comprising impulsivity and antisocial behaviour, thereby separating it from Machiavellianism, narcissism was depicted through the grandiose facet (Jones & Paulhus, 2014). Impulsivity and grandiose narcissism have both been suggested as being more germane to males (Corry et al., 2008; Massar et al., 2017).

The examination of age in relation to the DT constructs may be under investigated. It has been put forth that the age of participants may be a confounding variable in DT research (Paulhus, Buckels et al., 2020). It has been contended that by not controlling for age effects partially collapsed the DT factor structure in the Short Dark Tetrad. Yet, constraining the sample age to participants of 25 and under significantly lowered the median inter-factor correlations (Paulhus, Buckels et al., 2020). However, qualitative differences in the way indicators are interpreted and/or how the construct is conceptualised in the measures may mediate age affects. For example, pertaining to psychopathy, albeit with findings from an institutionalised male sample, the primary characteristics as put forth by Cleckley (1941), were not affected by age, however, the antisocial behaviour facet did decline (Harpur & Hare, 1994). Conversely, results from the Self-Report Psychopathy Scale-III (Williams, Nathanson & Paulhus, 2003), which is based on the PCL-R, showed that all psychopathy facets were negatively associated with increasing age (Gill & Crino, 2012). This finding was reflected with the Mach IV, with participants aged 38 and over showing significantly lower scores than the younger sample (Mudrack, 1989). Embodiment of Machiavellianism has been reported to be more marked during the transition period from late childhood to adolescence, reaching the lowest levels by the age of 65. However, this was found using the four Machiavellian items from the DD (Götz, Bleidorn & Rentfrow, 2020), which do not tap into the fullest extent of the construct.

Mean score differences have provided the evidence from the NPI to suggest younger participants had higher scores on the facets of authority, superiority, entitlement, exhibitionism and vanity, whilst older age groups displayed higher means on self-sufficiency (Kushari et al., 2017). A longitudinal study elucidated that participants aged between 18 and 41, showed mean level decreases in total score narcissism and the facets of leadership, vanity and entitlement (Wetzel et al., 2019). However, further research has suggested that men’s levels of narcissism decreases far slower than women’s by between ten and fifteen years (Wilson & Sibley, 2011).

In light of this discussion, through a synthesis of available evidence, the aim of this review is to investigate the stability, construct congruence, sex and age invariance of the prevalent scales used in DT research in order to establish whether research findings can be confidently generalised.

2.1 Method

The review was conducted following a systematic approach informed by guidance offered by the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA; Moher, Liberati, Tetzlaff & Altman, 2009). The systematic review process is illustrated in Figure 1.

Literature search

A systematic search was conducted to identify studies on the scale validity, factor structures, indicators, construct validity, sex and age invariance of dark personality scales. The inclusion criteria sought studies which encompassed adult male and female general population samples over the age of 18. Studies were restricted to the English language and those that had been peer reviewed. Unpublished theses were excluded as these may not have been subjected to the rigorous criteria imposed on published studies. The original scale development papers were also excluded in order to ascertain external post-development evidence.

The search was conducted in PsycArticles, PsychINFO, Web of Science and Google Scholar, with papers accessed through the University of Derby and British Psychological Society libraries. Search terms and Boolean operators were (Dark Triad, Dark Tetrad or narcissism, narcissistic or narcissist or Machiavellian or psychopathy or psychopath or psychopathic, everyday sadism or sadism) AND (sex or gender or men or women) OR (age or age group or age difference) AND (structure; IRT or item response theory; Dimension; Psychometric validity or reliability) NOT (PCL-R) NOT (criminals or convict or offenders or prisoners) NOT (children or adolescents or youth or child). The date ranges searched were in line with the year the first scale for each construct was developed for general population samples: Psychopathy, 1985-present; Narcissism, 1979 – present; Machiavellianism, 1970-present; Everyday sadism, 2003 - present; Dark Triad, 2013-present.

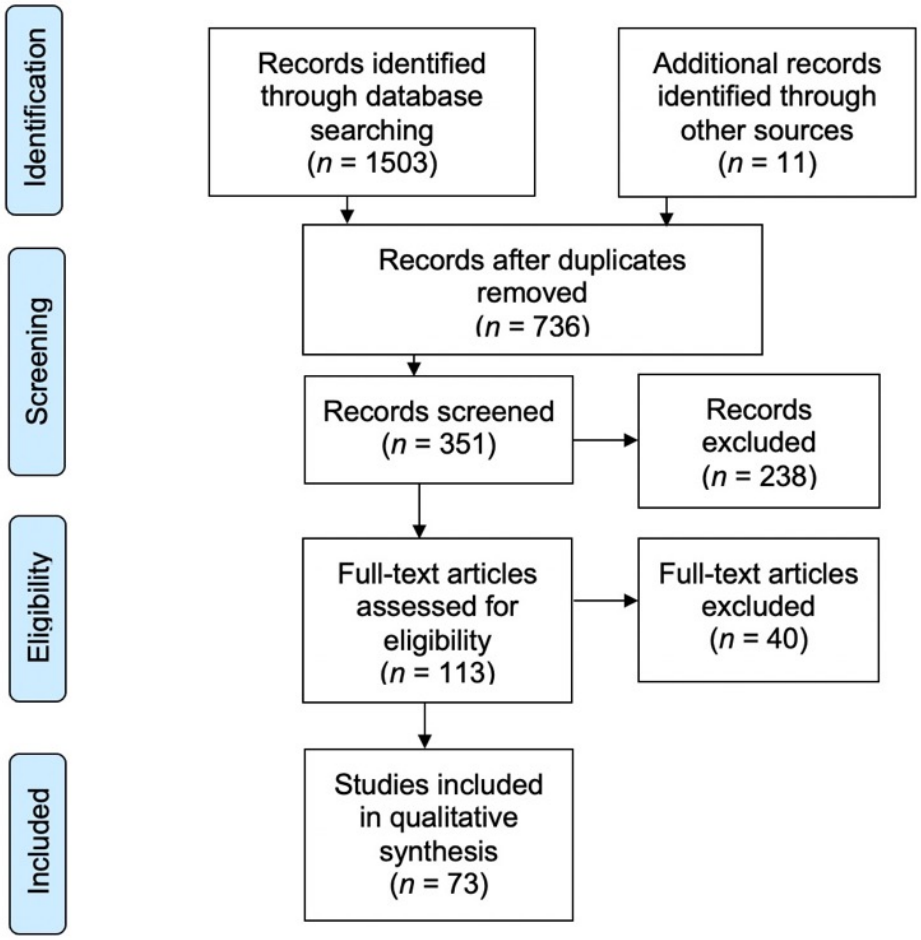
Data synthesis

Due to the heterogeneity in the studies, a formal meta-analysis was not possible because of the differences between studies in their contexts and methods, as such a qualitative synthesis was applied (Higgins & Green, 2011). Initially, 1503 publications were identified and imported into Endnote bibliographic database, where an electronic review highlighted 778 duplicates, which were deleted. In line with the inclusion/exclusion criteria, the remaining titles were assessed, resulting in the removal of 385 extraneous papers. The abstracts and methods sections, when further

clarification was required, of the remaining 351 studies were reviewed, reducing the list to 113 papers. The full texts were then evaluated and studies which did not meet the minimum requirements were excluded. Authors for studies identified in reference sections but which were inaccessible within the used libraries were contacted (*n* = 11), resulting in four further papers eligible for inclusion. Ultimately, 73 studies met the criteria to be included within the review (Figure 1).

Figure 1

PRISMA Flow Diagram



The studies included were critically evaluated using a modified version of the Critical Appraisal Skills Programme (CASP, 2018), cohort study appraisal form, prior to conducting the synthesis. This enabled the required data to be collated and extracted using a CASP form (Appendix 1.1) and a data extraction form, which can be

provided on request. While the potential for bias within the reviewed studies cannot be excluded entirely, none of the identified studies were deemed so at risk of bias that this warranted their exclusion.

2.2 Results

2.2.1 Study sample

Of the 73 studies included in the review, the breakdown for each construct is as follows: Psychopathy (*n*= 21); narcissism (*n*= 14); Machiavellianism (*n*=8); Everyday Sadism (*n* = 4); Dark Triad/Tetrad (*n*= 23), Machiavellianism and psychopathy (*n* =3).

2.2.2 Psychopathy

Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995)

The original two-factor structure of the LSRP was supported by five studies (Douglas et al., 2012; Hauck-Filho & Telxeira, 2014; Lynam, Whiteside & Jones, 1999; McHoskey et al., 1998; Salekin, Chen et al., 2014). Results showed that allowing the factors to correlate and estimating the model without including modifications, resulted in a reasonable fit (Douglas et al., 2012), or a poor fit, which significantly improved after allowing 17 errors of measurement to correlate (Lynam, Whiteside & Jones, 1999). This respecified model was criticised because these errors traversed the two-factors, were conceptually varied and therefore enacted without theoretical justification. However, incorporation of these modifications did not significantly improve the model fit (Douglas et al., 2011). There was a problem with an item from the secondary psychopathy facet, 'Love is overrated', shown as more relevant to the primary subscale (Hauck-Filho & Telxeira, 2014), or that it loaded significantly on both factors (Lynam, Whiteside & Jones,1999). The two-factor model was defended with the argument that it upholds the primary and secondary distinctions, allowing for stronger convergent and discriminant validity (Hauck-Filho & Telxeira, 2014; McHoskey et al., 1998).

There was support for the three-factor solution (Christian & Sellbom, 2016; Garofalo et al., 2018; Salekin et al., 2014; Shou et al., 2017; Somma et al., 2014), even though it was argued that the three-factor model mediates a construct drift (Christian & Sellbom, 2016; Garofalo et al., 2018). Evidence provided by Garofalo et al., (2018) and Shou et al., (2017), showed an adequate fit to the data only after the model had been respecified and errors allowed to covary. Scale items did not load

onto their intended factors (Somma et al., 2014), and the item, ‘When frustrated I often let off steam by blowing my top’ had weak associations with the latent construct (Garofalo et al., 2018). Two pairs of items had significant construct overlap and an acceptable fit was only found once the residuals of these items were permitted to correlate (Shou et al., 2017). The item ‘I have been in a lot of shouting matches with other people’, displayed a small factor loading with correlated error terms, therefore, it was suggested that this item may need discarding (Garofalo et al., 2018),

Sex invariance was assessed by six studies (Christian & Sellbom, 2016; Gummelt et al., 2012. Hauck-Filho & Telxeira, 2014; Lynam, Whiteside & Jones,1999; Marion & Sellbom, 2011; Salekin, Chen et al., 2014). Although factorial invariance was evidenced (Lynam, Whiteside & Jones, 1999; Salekin, Chen et al., 2014), Christian and Sellbom (2011) found that in order to achieve metric invariance in the three-factor model, the item ‘‘For me what’s right is whatever I can get away with’, needed to be freely estimated across both sexes. Lynam, Whiteside and Jones (1999), discussed that in the two-factor model, the item ‘I often admire a really clever scam’, loaded more strongly on the primary factor for men and in another study the item ‘Love is overrated’, from the secondary factor had a much higher loading on the primary factor across sexes, (Hauck-Filho & Telxeira, 2014). Ultimately, it has been argued that the same score for both sexes may capture different facets of psychopathic traits, with different meanings for males and females (Marion & Sellbom, 2011). This qualitative difference was corroborated in a further study, where males were found to endorse items tapping into boredom, impulsivity and a lack of long-term goals, whilst women endorsed indicators which addressed manipulation and egocentricity (Gummelt et al., 2012).

Self-Report Psychopathy Scale (SRP-III; Williams, Nathanson & Paulhus, 2003)

The original four-factor structure was corroborated across seven studies (Dotterer et al., 2017; Gordts et al., 2017; Lester et al., 2013; Mahmut et al., 2011; Neal & Sellbom, 2012; Neumann, Schmitt et al., 2012; Williams, Nathanson & Paulhus, 2003). However, Dotterer et al., (2017), suggested that a bifactor model with four sub-factors, whereby there was a general factor of psychopathy and four sub-factors of interpersonal, affective, lifestyle and antisocial, was a better fit. The authors also found the scale indicators to be problematic, with two items ‘get a kick out of scamming’ and ‘getting in trouble for some things’, not loading highly. On the affective factor, four of the seven items had negative loadings, including items which addressed

callousness, yet, removing these items would alter the core foundations of the construct and were retained.

Sex invariance was investigated by three studies (Dotterer et al., 2017; Neal & Sellbom, 2012, Neumann et al., 2012). Dotterer et al., (2017), found significant differences across sex, with endorsement levels for each factor varying for both threshold and mean scores. One item 'enjoy watching fighting' was excluded as it did not fit the model for females, and many of the interpersonal items had either low or negative loadings and/or non-significant loadings on the factor, a converse finding to males. The affective facet displayed several items with divergent loadings in both degree and direction across sex. Although configural and metric invariance was evidenced, neither the four-factor correlated, nor bi-factor model demonstrated scaler invariance. However, Neumann et al., (2012), showed strong invariance when the factor loadings were constrained and the threshold parameters were held to be equal, with minimal loss of fit, confirming the construct is equivalent across sexes. This was partially confirmed by Neal & Sellbom (2012), who found that when factor loadings and item intercepts were constrained and held to be equal across sexes, invariance was not evidenced however, allowing the intercepts but not the factor loadings to be feely estimated indicated strong factorial invariance across sexes.

Psychopathic Personality Inventory-Revised (PPI-R; Lilienfeld & Andrews, 1996)

The factor structure of the PPI-R was investigated by three studies (Anestis et al., 2011; Benning et al., 2003; Uzieblo et al., 2010). Benning et al., (2003), evidenced a two-factor structure, which diverged from the original three-factor model. However, both Anestis et al., (2011) and Uzieblo et al., (2010), did not find a good fit for the two-factor model, with items cross loading onto more than one factor. In relation to the Cold-Heartedness factor, Benning et al., (2003), found it to be problematic, whereby it emerged as an independent scale with the indicators considered unrelated to psychopathy. These problems were attributed to all items being reverse scored as well as the item content suggested to tap into sentimentality, imagination and emotional reactions, to a greater extent than the callous unemotionality they were developed to measure. However, Uzieblo et al., (2010), disagreed. They argued that the indicators tapped into the affective and interpersonal characteristics of psychopathy, which emphasised its usefulness as a factor. Anestis et al., (2011), expanded on this, whereby the cold-heartedness facet addresses psychopathic characteristics of a lack of remorse, empathy and guilt, which are core to the construct. The authors

contended that removal of this factor would fundamentally alter the definition of psychopathy, even though inclusion of the cold-heartedness facet resulted in a statistically poor model fit.

Scale indicators were investigated with a graded response model (Eichenbaum et al., 2019). From the 131 items, 39 were found to provide a minimal amount of information for the respective factors, and 80 items were evidenced as functioning differently between males and females. Specifically, males and females who embodied similar levels of the latent construct responded differently, especially at the high and low levels. There was a substantial difference in responses for indicators from the Machiavellian egocentricity factor relating to manipulation and conning people as well as the carefree non-planfulness facet which addresses strategic planning and career aims. Eichenbaum et al., (2019), contested that an intersex lack of invariance was likely either due to psychopathy manifesting differently in males and females and/or the semantics of the indicators mediating inadvertent sex-specific responses.

Only one study examined the sex invariance of the PPI-R (Anestis et al., 2011). Although one, two and three factor models were investigated, full invariance was not found. However, when constraints were freed, partial invariance was supported. The evidence suggested that when a mixed sex sample is used, none of the models provided adequate fit for the data and applying either of these to a mixed sex sample was considered unsuitable. It was stressed that the PPI-R was developed solely with male samples and therefore the poor model fit when using a mixed sex sample was unsurprising. This emphasises the ineffectiveness of the scale for females, suggesting a male bias in both design and application and highlights the importance of considering both males and females when developing a psychometric measure.

In line with the parameters of this review, no studies were found which investigated the psychopathy measures and age invariance.

2.2.3 Narcissism

Narcissistic Personality Inventory (NPI: Raskin & Hall, 1979)

The NPI has undergone numerous transformations in relation to its factor structure, with a unidimensional model (Ames et al., 2005; Braun et al., 2016); models of one or two factors (Barelds & Dijkstra, 2010); two factors (Corry et al., 2008); two or three factors (Kubarych et al., 2004); three factors (Ackerman, Witt et al., 2011); four factors (Emmons, 1984) and seven factors (Raskin & Terry, 1988). Model fits for the

one, two and three factor models have been evidenced as poor (Kubarych et al., 2004), with the unidimensional scale argued to be constructed on weak criteria and the four and seven factor structures unable to be replicated (Barelds & Dijkstra, 2010). The leadership/authority factor was stated as being the most accurate, with strong validity (Ackerman, Witt et al., 2011; Clarke et al., 2015; Corry et al., 2008; Raskin & Terry, 1988), however, it was argued that the maladaptive facet of narcissism, was inadequately represented within the entitlement and exploitativeness factor, with some indicators suggested as incompatible with the construct (Ackerman, Witt et al., 2011; Barelds & Dijkstra, 2010; Clarke et al., 2015). Although Braun et al., (2016), proposed that the scale is a valid measure of grandiose narcissism, both Ackerman, Witt et al., (2011) and Barelds and Dijkstra (2010), argued that NPI does not conform with the definition of the construct, with narcissism being determined by leadership and attention seeking.

Network analysis of the NPI (Briganti & Linkowski, 2020), showed there were weak items which should be made redundant. However, there were strong connections between indicators belonging to the same facets of authority, exhibitionism and superiority. These items displayed the highest strength centrality estimates. Entitlement was also central and was seen as the bridge between grandiose and vulnerable narcissism.

Sex invariance for the NPI has not been investigated to date. However, sex differences found within the NPI factors showed that males endorsed items on the leadership/authority and self-absorption/self-admiration factors, whilst women were more likely to endorse items pertaining to exploitation (Corry et al., 2008; Jackson et al., 1992; Tschanz et al., 1998), which was attributed to sex role socialisation (Jackson et al.,1992; Tschanz et al., 1998; Wilson & Sibley, 2011). Yet, this was disputed, with evidence suggesting that both sexes obtained similar scores on the exploitation/entitlement factor, as well as self-sufficiency, leadership/authority and superiority/arrogance, denoted to be male oriented (Brown, Akers & Giacomino, 2013; Corry et al., 2008; Jackson et al., 1992).

In line with the parameters of this review, no studies were found which investigated age invariance.

2.2.4 Machiavellianism

Mach IV (Christie & Geis, 1970)

The original three-factor structure of the Mach IV has been demonstrated by seven studies as not being a good fit for data (Ahmed & Stuart, 1981; Corral & Calvete, 2000; Hunter et al., 1982; McHoskey et al., 1998; Monaghan et al., 2016; Miller, Hyatt et al., 2017; Rauthmann, 2013; Ray, 1983; Williams, Hazleton & Renshaw, 1975). Varying structures have been suggested including two factors (Monaghan et al., 2016); four factors (Corral & Calvete, 2000; Williams et al.,1975) and five factors (Ahmed & Stewart, 1981). Three studies demonstrated that the Mach IV had low reliability (McHoskey et al., 1998; Rauthmann, 2013; Ray, 1983). Problems with scale items were found (Ahmed & Stewart, 1981; Hunter et al., 1982; Monaghan et al., 2016; Rauthmann, 2013), with three studies removing the item ‘‘The biggest difference between most criminals and other people is that criminals are stupid enough to get caught’ (Ahmed & Stewart, 1981; Hunter et al., 1982; Monaghan et al., 2016), as its inclusion rendered the factor as psychologically uninterpretable (Ahmed & Stuart, 1981), and the inter-item correlation was below acceptable limits (Hunter et al., 1982; Monaghan et al., 2016). Monaghan et al., (2016), found that by reducing the indicators from twenty to ten, alleviated some of the scale’s weaknesses, whilst Rauthmann (2013), reduced the scale to five items, due to twelve items loading over more than one factor, and many of the indictors evidencing low probabilities of endorsement, thereby providing minimal information.

In relation to the conflation of Machiavellianism and psychopathy, five studies suggested the two constructs are analogous (Czibor et al., 2017; McHoskey et al., 1998; Miller, Hyatt et al., 2017; Rogoza & Cieciuch, 2018; Vize, Lynam et al., 2018). McHoskey et al., (1998), argued that Machiavellianism was associated with primary and secondary psychopathy, whilst the findings of Czibor et al., (2017), showed a stronger correlation with the primary facet. Psychopathy and Machiavellianism were demonstrated as being highly correlated (r = .89), in a study investigating the three dark triad traits (Miller, Hyatt et al., 2017), which was corroborated by Vize, Lynam et al., (2018), in a meta-analysis, which considered both the nomological networks and the associations psychopathy and Machiavellianism held to criterion variables, the results evidenced Machiavellianism and psychopathy to be (r_{ICC} =. 86).

In line with the parameters of this review, no studies were identified which investigated the Mach IV in relation to sex or age invariance.

2.2.5 Everyday Sadism

In relation to the factor structures of the Short Sadistic Impulse Scale (SSIS), Assessment of Sadistic Personality (ASP) and Varieties of Sadistic Tendencies (VAST), both the SSIS and ASP demonstrated good fit indices for the one factor model, whilst for the VAST, a two-factor model was a significantly better fit than the one factor model (Dinić, Bulut Allred et al., 2020).

Sex invariance was investigated for the ASP by Plouffe, Smith and Saklofske (2018) and metric invariance was supported, however this was not corroborated by Min et al., (2019), who evidenced that the ASP only achieved configural invariance. The researchers also examined the sex invariance of the SSIS, where metric invariance was found. It was suggested that males and females with comparable levels of everyday sadism, responded differently to indicators and as such, for studies seeking to investigate sex differences, the SSIS was recommenced for use over the ASP. However, as scalar invariance was not achieved, the comparison of mean differences across the groups cannot be meaningfully compared.

Further investigating sex differences at an item level, Plouffe, Kowalski et al., (2021), using samples from Canada, Italy and Poland, found through a graded response model and differential item functioning that the indicators were mainly unbiased across sex. However, they did show that item 9, 'I would not purposely hurt anybody, even if I didn't like them', a reverse scored item, was ineffective at assessing everyday sadism across the samples and should be removed from the scale.

In line with the parameters of this review, no studies were found which investigated the everyday sadism measures with age invariance.

2.2.6 Dark Triad

Dirty Dozen (DD: Jonason & Webster, 2010)

Investigating the factor structure of the DD, four studies determined that the original three-factor model was the best fit for the data (Dinić, Petrović & Jonason, 2018; Klimstra et al., 2020; Rogoza, Zemoitel-Piotrowska, et al., 2020; Webster & Jonason, 2013), however, Watts et al., (2017), found that items did not load in line with Machiavellianism, psychopathy and narcissism but were heterogenous aggregates distributed across the constructs. Support for a dark dyad, whereby Machiavellianism and psychopathy were subsumed under one construct, with narcissism as a separate factor, opposed to the dark triad of three separate constructs, was demonstrated by three studies (Carter, Campbell et al., 2015; Dinić,

Petrović & Jonason, 2018; Maneiro et al., 2018). However, Kajonius et al., (2016), suggested that the DD measured narcissism and a separate facet which taps into antisociality.

In relation to scale items, it was found that those from the psychopathy facet were the most difficult to endorse, with participants needing higher thresholds of the construct (Carter, Campbell et al., 2015; Dinić, Petrović & Jonason, 2018; Garcia et al., 2018). Narcissism indicators were evidenced as the easiest to endorse across the constructs (Carter, Campbell et al., 2016, Kajonius et al., 2016). These items held the least discriminatory power, with participants requiring a lower threshold of the latent trait. Indeed, Garcia et al., (2018), suggested the facet provided more information about a lack of narcissism.

Sex invariance was considered by five studies (Carter, Campbell et al., 2015; Chiorri et al., 2019; Dinič, Petrović & Jonason, 2018; Maneiro et al., 2018; Webster & Jonason, 2013). Carter, Campbell et al., (2015), demonstrated that the DD is not invariant over both sex and age. However, the lack of sex invariance was disputed with further evidence suggesting both sex invariance and goodness-of-fit for the three-factor structure (Chiorri et al., 2019; Dinič, Petrović & Jonason, 2018; Maneiro et al., 2018; Rogoza, Żemojtel-Piotrowska, et al., 2020). Relative to scale items, Kajonius et al., (2016), showed through IRT analysis that the number of females endorsing the lowest scale category was disproportionate to males. The authors contended that in studies with small sample sizes, this may robustly affect statistical and external validity.

The DD, and specifically the Dark Core, which is a single factor that underlies psychopathy, Machiavellianism and narcissism, has been discussed as being taxonic (a natural category), for men and dimensional for women (Tran et al., 2018). This means, that for males, there are qualitatively different categories within the latent structure, whereby cut-off scores can be justified. For females, the DD constructs are best described as a continuum (Beller & Bosse, 2017). Related to both age and sex, Carter, Campbell et al., (2015), when comparing student aged samples and older participants across sex, found that the DD is better suited for older participants, where both a unitary scale as well as the three subscales functioned well across both sexes. The scale showed greater consistency for student females than males, with all three facets loading separately, whilst for the student males, Machiavellianism and

psychopathy were subsumed under one construct and narcissism emerged as a separate factor.

A further study which investigated age and sex differences (Klimstra et al., 2020), with a sample ranging from 11-77, found metric invariance across sex in all age groups. However scalar invariance was only found in the 14-16 and 17-18 age groups, whilst partial scalar invariance across sex was shown for the 11-13, 14-16 and 55-77 age groups. In young adults, neither partial nor scalar invariance was evidenced. Using an adult sample, Kawamoto et al., (2020), found scalar invariance across three age groups, ranging from 20-69.

Short Dark Triad (SD3: Jones & Paulhus, 2014)

Ten studies investigated the structure of the SD3 (Arseneault & Catano, 2019; Atari & Chegeni, 2016; Dinić, Petrović & Jonason, 2018; Gamache et al., 2018; Grigoras et al., 2020; Kawamoto et al., 2020; Persson, Kajonius & Garcia, 2017a; Persson, Kajonius & Garcia., 2017b; Rogoza & Cieciuch, 2018; Siddiqi et al., 2020; Vaughan et al., 2019). Although Arseneault and Catano, (2019), showed that the SD3 models failed to achieve an adequate fit, Vaughan et al., (2019), evidenced a good model fit for the three-factor structure. However, they found that a bifactor model of a general factor and three specific factors provided a much better fit to the data. Rogoza and Cieciuch (2018), determined that the original three-factor model did not converge, however a two-factor structure of Machiavellianism/psychopathy and narcissism fitted the data. This was further corroborated across multiple studies (Gamache et al., 2018; Persson, Kajonius & Garcia, 2017a, 2017b; Vize, Lynam et al., 2018). Siddiqi et al., (2020), also found a dark dyad with narcissism as a separate factor fit the data but only after removing fourteen items from the scale. Further research showed that amalgamated scores for all three constructs contained relatively little specific variance (Persson, Kajonius & Garcia, 2017b). Measurement invariance showed that the SD3 had configural invariance but not metric invariance across the groups investigated. However, these groups were job-applicants and non-job applicants rather than age or sex (Grigoras et al., 2020).

Turning to the indicators, more items were found to have stronger loadings on the dark core than their respective subscales. This was particularly pertinent for psychopathy and Machiavellianism (Gamache et al., 2018). To increase internal consistency of the SD3, Atari and Chegeni, (2016), discarded seven items, two from Machiavellianism which loaded onto psychopathy, three from the narcissism factor

and one item from psychopathy. However, internal consistency for narcissism was low, with remarkably high levels of endorsement and the corrected item scale correlations were particularly low (Gamache et al., 2018). Further item analysis showed that more than 50% of indicators were affected by differential item functioning (DIF; Grigoras et al., 2020), indicating that those who completed the scale with the same levels of the latent constructs responded differently to more than half of the SD3 items. Machiavellianism was the most affected with seven of the items showing DIF. Persson, Kajonius and Garcia (2017a), found through a graded response model, that Machiavellianism indicators were more easily endorsed than those of psychopathy, which may pertain to the content being comparatively less socially toxic.

Sex invariance was investigated by Vaughan et al., (2019), who found that the model differed across males and females. As each model with increasing restraints was assessed, there was a decrease in model fit from the baseline configural model. It was also observed that the factor structure for males showed fewer instances of misspecification comparatively to females, However, both sexes evidenced discrepancies with the conceptualisation of the constructs in the SD3, although the general dark core remained stable across groups.

Age invariance was investigated with the sample split into three groups; 20-29; 30-49 and 50-69 (Kawamoto et al., 2020). The fit indices for invariance were below the recommended guidelines, thus invariance was not supported. Age was significantly and negatively correlated with the DT traits and the composite score. Only narcissism showed a weak but significant and positive association with age in males. Psychopathy and Machiavellianism decreased with age, with Machiavellianism decreasing at a faster rate for women.

2.2.7 Network Analysis of the Dark Constructs

Three studies investigated the dark triad with network analysis (Dinić, Wertag & Tomašević, 2020; Trahair et al., (2020); Truhan et al., 2020). Truhan et al., (2020), used the NPI, Mach IV, SD3, DD, SRP-4 and the Five-Factor Narcissism Inventory-Short From. Exploratory Graph Analysis (EGA), showed that the NPI and SD3 narcissism were only connected to facets of grandiose narcissism, with the SD3 more strongly associated with superiority, authority and exhibitionism. Leadership and authority were situated on the outer edges of the network, suggesting they may tap into a separate domain. DD narcissism was associated with both vulnerable and grandiose dimensions. For psychopathy, the SD3 factor mainly aligned with antisocial

and interpersonal characteristics, whist DD psychopathy was clustered with affective and antisocial facets. The analysis showed that the DD addresses specific aspects of Machiavellianism and psychopathy rather than definitive constructs. It further elucidated that psychopathy and Machiavellianism indicators clustered on one dimension, providing further support for the conceptual overlap.

A network analysis which incorporated the NPI, Narcissistic Admiration and Rivalry Questionnaire, Mach IV and LSRP (Trahair et al., 2020), showed the network aligning with a dark triad model, with the agentic facet of narcissism clustered furthest away from Machiavellian and psychopathy. Primary psychopathy was found closer to the Machiavellian measures, whilst secondary psychopathy was clustered close to the SD3 psychopathy facet. Yet, Dinić, Wertag, et al., (2020), who included a measure of everyday sadism with their network, the SSIS, along with the SD3, Mach IV, LSRP, NPI and Psychopathic Personality Trait Scale, contended that secondary psychopathy should be redundant from the dark core. Primary psychopathy was found to be the most central aspect of the DT network and at the facet level primary psychopathy was the central feature of the LSRP and the core of the dark traits.

2.3 Discussion

The popularity of scales has been discussed as a reason not to discard them (Ackerman, Witt et al., 2011). However, the evidence has shown that across the measures, factor structures are unstable, item content has facilitated construct drifts, indicators are not homogenous across sex and scales are mainly neither sex nor age invariant. There appears to be an inconsistency amongst construct definitions (e.g., Jones & Paulhus, 2014; Williams, Paulhus & Hare, 2007), and a lack of consideration of the female manifestation of the constructs during the scale development process (e.g., Anestis et al., 2011). This does not imply that past studies have provided untenable findings, as researchers have used the tools that were available to them, yet it is indisputable that there are limitations with these measures (Ackerman, Witt et al., 2011).

Although it can be argued that all scales will be subjected to criticism and also have limitations, the nature of the limitations may create issues in the representativeness of the findings, whereby the measures work for some groups better than others (e.g., Carter, Campbell et al., 2015). However, some of these limitations may have been mediated by psychometric restrictions within the time

period several of the measures were developed. Although aspects of measurement invariance have been discussed in the literature over the decades; configural invariance (e.g., Buss & Royce, 1975; Irvine, 1969), metric invariance (e.g., Horn & McArdle, 1992; Meredith, 1993) and scalar invariance (e.g., Meredith, 1993; Steenkamp & Baumgartner, 1998), the methods were discussed in a technical mathematical language which was not widely comprehensible (Wu et al., 2007). With advances in accessible statistical software, Meredith's (1993) concept of measurement invariance, set under the confirmatory factor analysis framework, was explained in a more widely accessible format (Wu et al., 2007). As, for example, the LSRP, Mach IV and NPI, were developed between 1970 and 1995, the findings of this review should be considered within this context. A more in-depth discussion of measurement invariance can be found in Chapter Three.

Turning to the findings, the unstable factor structure of the NPI has been widely debated (e.g., Kubarych et al., 2004; Pincus & Lukowitsky, 2010; Rosenthal & Hooley, 2010), which may align with the conflation of non-narcissistic and narcissistic traits (Ackerman, Correti & Carson, 2018), and partially account for the variance across samples. This has led some to argue that the NPI is an ineffective measure of narcissism (Cain et al., 2008; Rosenthal & Hooley, 2010). Although not suggested as a core component (Rosenthal & Hooley, 2010), the Leadership/Authority subscale has been intimated as the most accurate factor (Ackerman, Witt et al., 2011), Yet the content of the indicators do not necessarily differentiate between an individual with high self-esteem and positive psychological traits or one who embodies narcissism (Peterson et al., 2009; Rosenthal & Hooley, 2010). Although it was deemed a failure that leadership was not included in the DSM-III definition (Miller, Price & Campbell, 2012), the DSM-IV-TR (APA, 2000), also omitted references to leadership. Conversely to Freud's (1931/1950), consideration that narcissism is a salubrious and fundamental process in normal development, the DSM criteria focused on pathological manifestations (Rosenthal, 2006).

Narcissism as measured by the brief scales of the DD and SD3, also diminish the vulnerable facet. Consequently, these scales may provide evidence of a lack of narcissism in participants (Garcia et al., 2017; Jonason & Krause, 2013). Relationships with external correlates have shown the SD3 narcissism factor to produce findings of a lack of impulsivity and an ability to plan ahead (Szabó & Jones, 2019), which are congruent with Machiavellian characteristics (Christie & Geis, 1970).

Narcissism as measured by the DD was found through item response analysis, to have the least difficulty and discriminating power (Kajonius, Persson et al, 2016). Consequently, there remains concern over whether the NPI, SD3 narcissism and DD narcissism are congruent with the construct they seek to measure and whether from a semantics perspective, the indicators in the DD and SD3 were selected to try and ensure a separation from psychopathy and Machiavellianism (Persson, Kajonius & Garcia, 2017b).

The SD3 indicators which tap into psychopathy mainly correspond with the lifestyle and antisocial facets of secondary psychopathy (Williams, Paulhus & Hare, 2007). The Machiavellian indicators mostly align with the primary psychopathy features of interpersonal and affective traits (Williams, Paulhus & Hare, 2007). Thus, it is unsurprising that researchers found a dark dyad structure of Machiavellianism/psychopathy, with narcissism as a separate factor (Gamache et al., 2018; Persson, Kajonius & Garcia, 2017b; Vize, Lynam et al., 2018). These findings were mirrored in the DD, with evidence supporting the dark dyad structure (Maneiro et al., 2018), mediated by the minimal semantic differentiation between the psychopathy and Machiavellian indicators. This was further suggested in exploratory graph analysis, where it was contended that all the DD indicators mainly tap into manipulation. As such, this prevents the determination of definitive facets (Truhan et al., 2020). Consequently, the oversimplification of the constructs in the two brief measures, has diminished the ability to disentwine primarily Machiavellianism and psychopathy (Miller, Price & Campbell, 2012).

The scales which measure psychopathy were also found to have unstable factor structures. Although some authors defended the two-factor model of the LSRP (Hauk-Filho & Telxeira, 2014), a good model fit could not be obtained (Douglas et al., 2012). This may imply that an alternative perspective on how psychopathy manifests in general populations is considered. Indeed, on the affective facet of the SRP-III, many items had negative loadings (Dotterer et al., 2017), whilst the cold-heartedness factor on the PPI-R, which is comprised of reverse scored items, did not fit the model (Benning et al., 2003), and 80 items from 131, were found to function differently across males and females (Eichenbaum et al., 2019). Thus, across psychopathy measures, the evidence points to a problem with both the indicators and the scale structures (Christian & Sellbom, 2011; Marion & Sellbom, 2011: Gummelt et al., 2012; Dotterer et al., 2017). As suggested by Dinić, Wertag et al., (2020), it may be that

primary psychopathy is the most central aspect of the DT network and secondary psychopathy is a redundant dimension within general population samples.

There was limited evidence provided for everyday sadism measures to draw any firm conclusions, thus, further interrogation of the scales will be addressed in Chapter Four. Turning to sex invariance, although there is a minimal amount of research which addresses this issue, what has been evidenced is that during the development of the scales, the fundamental requirement of sex equivalence may not have been considered (Forouzan & Cooke, 2005), or for some of the measures, the psychometric tools were unavailable. When invariance was investigated, in the LSRP, SRP-III, ASP and SSIS, mainly metric invariance was achieved, which means that the factor structure is equal across sexes and thus, an equivalent construct is being measured (Christian & Sellbom, 2011; Dinić & Vujić, 2019; Dotterer et al., 2017; Min et al., 2019; Plouffe, Smith & Saklofske, 2018). For the DD, in the 19-30-year age range sex invariance was not found (Klimstra et al., 2020). Without a consensus of studies evidencing scalar invariance of the measures, it is difficult for intergroup mean differences to be robustly compared, however diverse samples may provide alternative findings (Neumann, Schmitt et al., 2012).

A salient possibility may be that the impact of female data is incongruent with the scales' contents and structures. The authors of the LSRP stated that sex differences found between men and women in primary psychopathy was due to a callous interpersonal attitude being accentuated in the social norms of men but not women, and additionally, that women may have learned to deny these attitudes even if they embody them (Levenson et al., 1995). The authors of the Dirty Dozen specifically stated the validation criteria that males should score higher than females (Jonason & Webster, 2010), whilst Lilienfeld and Andrews (1996), initially used male samples during development of the PPI, because they tend to have higher levels of aggression, boldness and other psychopathic traits comparatively to females. Although individuals from either sex would not savour being labelled as embodying dark personality traits, the evidence has elucidated that the measures do not semantically correspond at an intersex level.

An interesting point made by Raskin and Hall (1981), in relation to the NPI, is that both age and sex of the participants did not significantly correlate with narcissism. The mixed sex participant sample were undergraduates with a mean age of 23. The authors stated that finding no significant relationship between sex and narcissism was

unsurprising, as no sex differences had emerged in their previous studies. They further contended that the lack of relationship between age and narcissism was due to the restricted age of participants. Yet, despite this, age invariance of the prevalent measures was rarely investigated and the only found studies related to the DD and SD3 (Kawamoto et al., 2020; Klimstra et al., 2020). The concerning finding was the lack of invariance found in a Western sample (Klimstra et al., 2020), for the DD, conversely to the scalar invariance evidenced in a Japanese sample (Kawamoto et al., 2020), which indicates cultural variances.

It is also unclear whether the DT measures perpetuate the male sex-role stereotype in Western cultures. As was mentioned in the introduction in relation to narcissism, measures of self-report may not encapsulate how specific traits are manifested across sex, which may lead to assumptions of sex role stereotypes (Baez et al., 2017). Thus, the measures used may inadvertently continue to preserve sex stereotypes and the voices of females, who embody dark personality traits may be silenced, especially if they are removed from data as an outlier (Aguinis et al., 2013; Neumann, Kosson & Salekin, 2017). What has been elucidated is the lack of clarity as to whether findings from research, based on the prevalent measures, truly reveal how both sexes manifest the constructs.

The evidence from this review suggests that the lack of invariance across age and sex may have an adverse effect on the robustness of the measures. This can be seen through the unstable factor structures and the qualitative differences in how males and females across ages endorse scale indicators (Carter, Campbell et al., 2015; Lynam, Whiteside & Jones, 1999), yet these differences are seldom considered (Marion & Sellbom, 2011). Consequently, studies which assume sex and/or age invariance without analysis either through classical test theories or alternate methods, such as item response theories (e.g., Gummelt et al., 2012), are negating the nuances of the similarities and differences between and across sexes and ages. Females within the field of dark personality research may be considered socioculturally unproblematic, yet this culminates in them being psychologically ‘othered’ (Eagly, 1987; Tavris, 1993), and thus, side-lined when considering the constructs of Machiavellianism, psychopathy, narcissism and everyday sadism.

2.3.1 Limitations

This review is not without its limitations. Grey literature was excluded, which may increase the risk of publication bias. However, it was decided to prioritise peer reviewed papers and the limitations of this decision are acknowledged. A further limitation is that only the widely used dark personality scales were included. Although other scales have been developed to measure the constructs, e.g., The Machiavellian Personality Scale (Dahling et al., 2009), these do not appear to have been put to widespread use. It may be argued that this is not just a limitation of this review but also of dark personality studies. A further limitation of research, that was reflected in its absence in this review, is the paucity of studies which have investigated whether the dark personality measures are both sex and age invariant. However, this may be a function of researchers not seeing it as being important.

2.3.2 Conclusion

The scales used in research that measure the DT constructs appear to have a range of issues. Construct drifts have occurred in psychopathy, the validity of Machiavellianism as a construct has been questioned, narcissism is only being partly assessed and the measures across constructs have lacked sex invariance, whilst those with supporting evidence have mainly showed age variance. The review findings have highlighted that the scales are exposed to a cycle of criticism, yet those such as the LSRP, NPI and Mach IV, should be situated within their historical timeframe when powerful statistical software was not readily accessible. The measures were developed with the tools the researchers had available. Even so, despite all the criticisms, these scales have emerged as the ‘gold standard’, and over time, there is a possibility that they have become conflated with the constructs, such that the construct becomes defined by the measure (Schimmack, 2010). This may lead to a disconnect from the aetiologies of the traits and behaviours.

The field of dark personality research has given minimal consideration to the manifestations across sex and age. Although it is seemingly apparent that males and females across ages demonstrate idiosyncratic behaviours, this goes against the grain of homogeneity in psychological research. Yet, if these differences are avoided because they have wider ranging repercussions, then the time may come when psychology ignores salient societal questions and the field declines into an evaporating relevance (Eagly, 1987).

Chapter Three

Methodology

This chapter details the diverse statistical methods that will be used in the subsequent studies. These encompass Classical Test Theories (CTT), exploratory graph analysis (EGA) and Item Response Theories (IRT). Initially, the merits of using both CTT and IRT will be considered and this will be followed by more in-depth explanations of the latent variable models. The discussion will further address the EGA and IRT methods. As mentioned in Chapter Two, technological advances have provided researchers with accessible and freely available statistical software, which facilitates the robust analysis of data from both the top down and bottom up. The benefits of this were highlighted in Chapter Two, by the evidence which elucidated the psychometric problems with the investigated measures. Psychometric scales are used as the foundation for ensuing research (Crede & Harms, 2019), and their reliability and the confidence that study findings can be generalised should be assured. Technological advances have also provided researchers with a range of options for sourcing participants. Relevant to the studies in this thesis, the benefits of crowd sourcing platforms such as Prolific, will be discussed, when both large sample sizes and equal groups are required.

3.1 Classical Test Theories and Item Response Theories

The main consideration between CTT and IRT is that CTT uses the measure as the unit of analysis, whereas IRT uses the item as the element of analysis. IRT methods have unique qualities, which complement CTT and as such offer significant advantages when investigating and/or developing Likert-type scales. Specifically, the item-level focus of IRT contrasts with the test-level focus of CTT (Kean & Reilly, 2014), whereby, a scale is considered at both micro and macro levels.

Researchers have predominantly used CTT methods such as Confirmatory Factor Analysis (CFA) and Exploratory Factor Analysis (EFA; e.g., Jonason & Webster, 2010; Jones & Paulhus, 2014), when seeking to establish the foundations of psychometric scales (Crede & Harms, 2019). These methods are encompassed within the overarching framework of Structural Equation Modelling (SEM; Kline, 2015), where the underlying ethos seeks to explain the observed scores. These scores are produced by semantic indicators which a participant responds to and consequently, the indicators are assigned to a set of unobserved factors or latent variables. One

main assumption of factor analysis models is that the observed scores are assumed to be caused by the personality trait that the latent variable addresses (Hoyle & Duvali, 2004). For example, data could be collected to investigate a psychological construct and based on previous theory, it may be hypothesised that there are four factors, thus four different causes motivating the participants responses. To investigate whether four factors are sufficient to explain the observed data, either EFA or CFA has historically been used.

A further assumption of latent models is that the true score and error score are uncorrelated. The true score is the participants score on a scale if there is no measurement error. Yet, CTT assumes that although elements of the true score are present for the measured construct, this is rarely observable. The observed score, which may differ across participants, dependent on the levels of the underlying construct they embody, is therefore comprised of the true score and a degree of measurement error (Salkind, 2010). Consequently, the difference between the true score and the observed score is determined by measurement error (Adedoyin & Mokobi, 2013). That is, the expected value of the error score for any participant is zero, implying that the scale is unbiased and the error score is caused by random error. This suggests that the error scores on scales should not be correlated (Hambleton & Jones, 1993; Ogunsakin & Shogbesan, 2018; Tarka, 2011). These methods therefore allude to ones which measure the same underlying construct whereby the same true score is found across participants and the errors across tests are equal (Ogunsakin & Shogbesan, 2018).

This highlights a limitation of CTT, whereby there is no indication as to how the scale items or the participant characteristics differ across samples. This is an important consideration as the essence of the items may change, depending on the sample used. For example, if a sample embodies high levels of the latent trait being measured, then all scale indicators may seem easy to endorse. Conversely, if a sample embodies low levels of the construct, the same set of scale items would be more difficult to endorse (An & Yung, 2014). Thus, participants responses to specific scale items are not considered with CTT, consequently, attained scores are solely dependent on the measure as well as being sample specific (Ogunsakin & Shogbesan, 2018). However, these weaknesses in CTT can be transcended through IRT, which produces models of item level statistics, whereby each scale indictor is uniquely validated and remains valid even when a subset of the items are used from

the validated facet. Distinctive parameters are formulated from both the characteristics of each indicator and the participants levels of the latent construct. Once the indicators have been determined for a population, the participants scores from that sample can be compared directly, even if different subsets of the items are used (An & Yung, 2014; Edelen & Reeve, 2007). IRT models can therefore provide both item and latent trait parameter invariance for scale indicators and participants. An important contribution from these models, not found in CTT, is that the same items can be used with different samples, whilst keeping their statistical properties (Zanon et al., 2016).

Reliability is also an important aspect of a psychometric scale. In CTT models, reliability is based on the entire measure, yet as the properties of the scale vary across populations, this can be a problem when small and/or heterogenous samples are used. IRT models encompass a range of robust indicators at the item and model level, which ensure the data fits to the model, which is more inclusive than those found within CTT methods. Reliability varies across the continuum, with greater accuracy found around the mean of the sample (Kean & Reilly, 2014). Finally, a further benefit of IRT is the reduction of scale items, ensuring that only discriminative indicators are included in the scale. These items are elucidated by the discrimination parameter, which will be discussed in further detail below. A scale which is determined by IRT can provide accuracy and significantly improved reliability, whilst fewer indicators can potentially reduce time and effort for participants (An & Yung, 2014).

3.1.1 Exploratory Graph Analysis

Despite exploratory graph analysis (EGA), being a somewhat new method, the ethos of psychometric network models is that traits emerge not because of a latent common cause but rather from a bidirectional relationship between the indicators (Cramer et al., 2012). The variables directly and jointly reinforce each other which results in a causally linked network (Borsboom, 2008). The emerging network does not depend upon theory or a priori assumptions but is determined solely by the data. This facilitates the appraisal of the theoretical structure (Christensen, Gross et al., 2020). Conversely, exploratory CTT methods such as exploratory factor analysis (EFA), seek latent variables which reflect characteristics that are thought to occur concurrently due to an underlying common cause (Schmittmann et al., 2013), which thereby explains the covariance between sets of observed variables.

In Monte Carlo simulations as well as real-world datasets, EGA has shown equivalent or increased accuracy in identifying facets over other reduction techniques (Christensen & Golino, 2020; Golino & Demetriou, 2017; Golino & Epskamp, 2017; Goliino, Shi, et al., 2020). For example, parallel analysis and the minimum average partial procedure are adequate methods when interfactor correlations are not high, there is a minimum sample size of 500 and when the item factor loadings range from medium to high (Keith et al., 2016; Crawford, Green et al., 2010; Green et al., 2016). Yet, these methods tend to underestimate factor quantity when the interfactor correlations are high, there is a small sample size and few indicators per factor (Keith et al., 2016; Crawford, Green et al., 2010; Green et al., 2016). Evidence has shown that the Kaiser-Guttman rule overestimates the number of factors, particularly when indicator numbers and sample sizes are large (Ruscio & Roche, 2012).

EGA has also shown advantages over EFA. For example, different rotation methods used in EFA can have significant effects on the data in relation to the estimation of factor loadings (Sass & Schmitt, 2010). EGA does not use rotation but produces orthogonal dimensions and further removes the subjective decisions sometimes found in EFA. The EGA algorithm places indicators into a dimension, with no direction from the researcher. This is converse to the factor loading matrix produced by an EFA, which initially needs interpretation. This highlights a further limitation of the EFA matrix, whereby the complexity of indicators are equivalently represented and can load onto more than one factor. Thus, EGA elucidates whether indicators are loading onto their intended facets and further provides graphical depictions of the interactions between indicators (Bringmann & Eronen, 2018; Christensen, Golino & Silvia, 2020). EGA has been used across different domains of psychology, including psychopathology (Borsboom & Cramer, 2013) and dark personality research (Dinić, Wertag et al., 2020; Trahair et al., 2020).

3.2 Classical Test Theory: Confirmatory Factor Analysis

CFA is theory driven, with the aim of investigating how well the hypothesised factor structure fits the data, through a specified model (McDonald, 2014). CFA models are designed with indicators assigned to specific factors, which align with theory. This enables multiple models to be explored and facilitates comparisons between these models to ascertain the best model fit for the data (Lancaster et al.,2009). CFA enables the examination of the scale structure through confirmation of the number of factors alongside the pattern of item-factor relationships (Schreiber et

al., 2006). It is an essential analysis in the latter phases of scale development or when investigating existing measures, when the underlying structure has been theoretically and empirically determined (Brown & Moore, 2012).

3.2.1 Sample Sizes for CFA

Disputes in the literature have resulted in diverse and contradictory advice in relation to the recommended sample size for CFA. Historically, it has been suggested that a minimum sample of 50 is adequate (Sapnas & Zeller, 2002), although others have argued that 100 participants is poor (Comrey and Lee, 1992). Further contentions have stated that a CFA model with between six and twelve items per factor could be specified with 50 participants, whilst 100 would be needed with a model where factors contains three to four items (Boomsma, 1985; Marsh & Hau, 1999). When conducting multiple group CFA's, a general rule of thumb has been put forth as 100 participants in each group (Kline, 2016; Wang & Wang, 2012). However, using robust estimators, such as WLSMV (discussed below), alleviates the necessity for large sample sizes (Raykov, 2012). Indeed, with these estimators in certain circumstances, as few as 60 participants can be used to determine a CFA model (Kline, 2015). However, when conducting factor analysis, where the number of factors in a model is unknown or disputed, the suggestions as to sample size have been stated as implausible on an a priori basis. It is therefore advocated that as large a sample as possible is obtained and the analysis conducted (MacCallum et al., 1999).

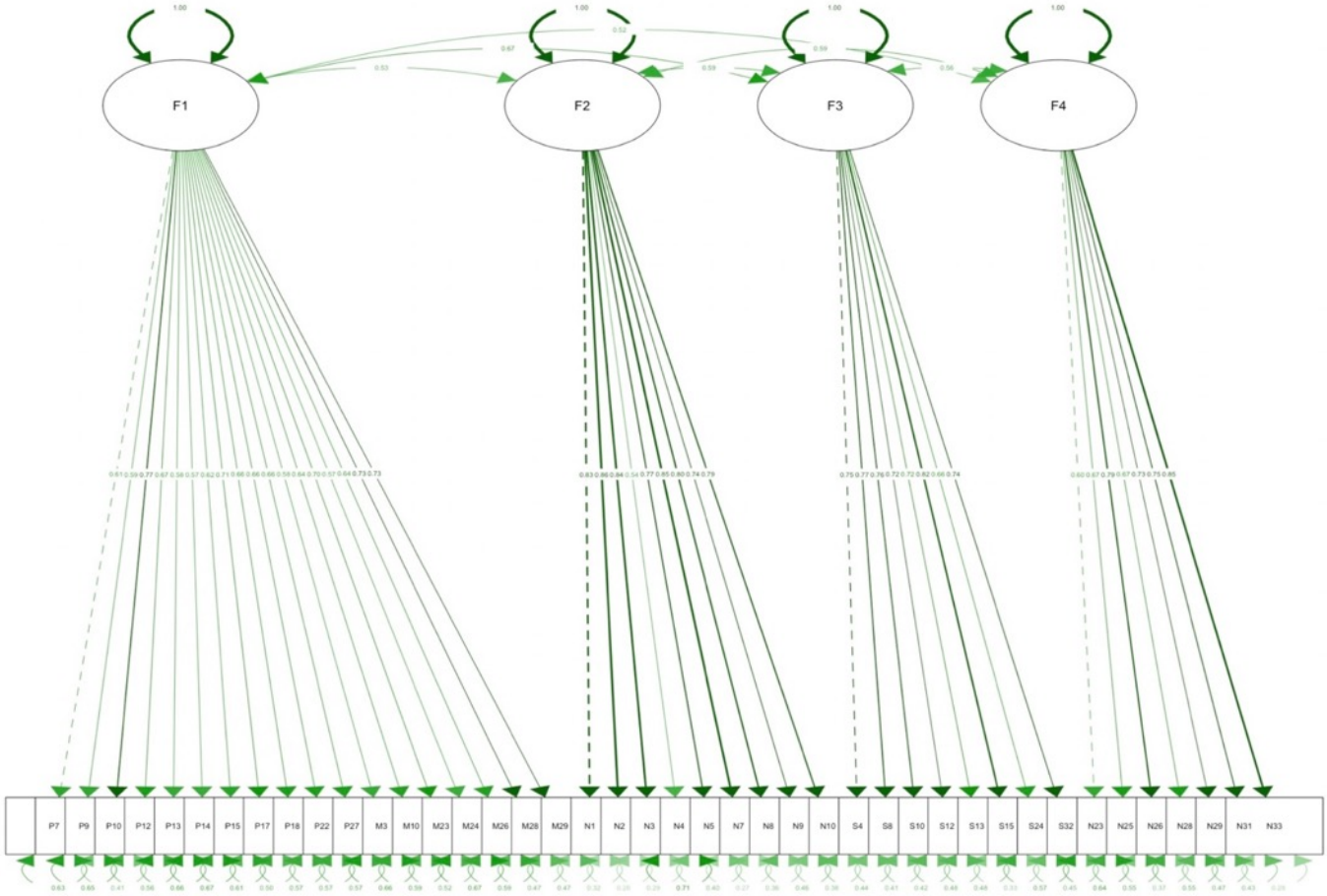
3.2.2 CFA Fit Indices

To determine the adequacy of the model fit, the following indices are considered. A chi-square (χ^2), difference test indicates the difference between the expected and observed covariance matrices. If the value is close to zero, then this denotes that there is minimal difference between the two matrices (Suhr, 2006). The Comparative Fit Index (CFI) shows the extent to which the tested model is a better fit than the alternative model found with the observed covariance matrix (Chen, 2007). The Tucker-Lewis Index (TLI) is a Non-Normed Fit Index, which measures the relative fit of the model through consideration of the degrees of freedom from the specified and baseline models (Schermelleh-Engel et al., 2003). A good model fit is determined when CFI and TLI values are $\geq .90$, in line with Kline's (2015), recommendations. The Root Mean Square Error of Approximation (RMSEA) assesses whether the model fits approximately well for the sample used, such that the null hypothesis of exact fit is

replaced by the null hypothesis of close fit. RMSEA values range from 0 to 1, and a good fit is found if the value is 0.06 or below and an adequate model fit accepted at 0.08 or below. Models are rejected if they display a value above 1.0 (Brown, 2015; Hu & Bentler, 1998; Marsh, Hau & Wen, 2004). The suggested estimator to use for non-normal data is weighted least means and variance (WLSMV; Salekin et al., 2014). This estimator is also advised to be used with binary and ordinal data, with the ordered command in lavaan (Rosseel, 2020). Graphical representations of the model can be produced in R, using the lavaan package (Figure 2). The figure below is for representation purposes only and shows a four-factor structure, with the correlated relationships indicated by the arrows across and between the factors.

Figure 2

CFA Four-Factor Model



3.3 Measurement Invariance

Measurement invariance (MI) is used to assess whether a scale measures the same construct across different groups and determines whether participants interpret the constructs in the same way. MI across and between sex and age is a central argument throughout this thesis, as it is considered an essential requirement for the robustness of psychometric scales. A lack of MI suggests that the scale is biased, whereby participants with *n* degree of the latent trait from one group, consistently present lower or higher response scores than participants with the same level of the latent trait from another group. This bias is engendered by the scale which does not elicit true differences. Consequently, the scores from scales which do not achieve scalar invariance are not able to be meaningfully compared across groups (Halamová et al., 2019; Xu & Tracey, 2017). MI models are based on the CFA framework, and are compared through the fit indices discussed above. If the difference in the fit indices of ΔCFI and ΔRMSEA (with Δ denoting 'difference'), between a model and the preceding less constrained model is not larger than 0.01 for ΔCFI and equal or less than 0.015 for ΔRMSEA, this suggests that the investigated level of measurement invariance has been attained (Chen, 2007; Cheung & Rensvold, 2002; Marsh, Nagengast & Morin, 2013; Orri et al., 2018; Rudnev et al., 2018).

MI is detected through a series of CFA models which impose increasing restrictions. The first stage is configural invariance, which serves as a baseline model and is determined by unconstrained factor loadings. It assesses whether the same set of indicators associated with the same factors are the same across groups (Orri et al., 2018). Configural invariance is achieved if this model evidences a good fit in line with the CFA guidelines discussed in section 3.2. Configural invariance shows that the factor structure is equivalent across groups (Xu & Tracey, 2017). If the baseline model provides a good fit, then metric invariance is investigated.

Metric invariance determines whether the relationship with the factors and the indicators hold across the groups. This is achieved by constraining the factor loadings to be equal across groups (Abrams et al., 2013). If this model is supported, it signifies that the indicators are psychologically meaningful within the same factors across groups (Vandenberg & Lance, 2000). This model is compared to the configural invariance results. Metric invariance is present if the fit is not substantially worse than the baseline model (Putnick & Bornstien, 2016). Metric invariance is necessary in

order to meaningfully compare the relationships between the latent variables, across different groups. If metric invariance is established, then scalar invariance can be considered.

Scalar invariance is the most important model needed to achieve MI. It indicates that the same mean level of the latent construct is supported across groups and also signifies that the measure does not hold response biases (Chen, 2007; Vandenberg & Lance, 2000). Scalar invariance is determined by constraining equal factor loadings and item intercepts, implying that not only the item loadings but also the item intercepts are equivalent (Putnick & Bornstein, 2016). Scalar invariance is established if the model fit is not markedly worse than that of the metric invariance model. A measure which holds scalar invariance denotes that researchers, when testing hypotheses which concern the means and interrelations of the investigated construct, can be confident that the scale produces robust results (Hirschfeld & von Brachel, 2014).

3.4 Exploratory Graph Analysis

There are no main assumptions for conducting EGA, however if data has a non-normal distribution the threshold function can be used to transform the data. This culminates in the reflection of a latent normally distributed score (Epskamp & Fried, 2018). Exploratory Graph Analysis (EGA), investigates the dimensional structure of the constructs at a facet level through the *EGAnet* package in R (Golino & Christensen, 2020). The visualisation of the complex relationships amongst variables is provided by a Gaussian Graphical Model (GGM: Constantini et al., 2015; Epskamp & Fried, 2018). GGM's are undirected, which means relationships between variables are not constrained to causal pathways. Nodes (circles) within the network form clusters, whereby many nodes are interconnected, alternatively they may form cliques, which is when all the nodes are connected to each other. These clusters or cliques can be viewed as latent dimensions within the network. The edges (lines), between the nodes indicate that wider and more saturated edges signify stronger partial correlations whilst the less saturated and thinner lines denote that the edge weight is close to 0, after controlling for all other nodes in the network (Epskamp & Fried, 2018).

Edges are regularised through the graphical least absolute shrinkage and selection operator (GLASSO; Friedman, Haste & Tibshirani, 2014), which shrinks small correlations to zero and produces a sparse inverse covariance matrix, thereby avoiding overfitting (Epskamp & Fried, 2018; Golino & Demetriou, 2017; Golino &

Epskamp, 2017). The sparsity is mediated by the lambda parameter (λ). Fewer edges are removed when λ is set at lower values, which increases the possibility of spurious correlations. Higher λ values reject more edges, which increases the possibility of important edges being removed. When λ=0, estimates are equal to the ordinary least squares solution for the partial correlation matrix. This denotes conditional independence and facilitates the interpretability of the network structure, requiring fewer connections to explain the covariance between variables (Epskamp & Fried, 2018). In the ensuing studies, the ratio of the maximum and minimum λ will be set to 0.1 (Golino & Christensen, 2020).

The regularisation parameter is defined with the extended Bayesian information criteria (EBIC; Chen & Chen, 2008; Epskamp, Borsboom & Fried, 2017), and employs a hyperparameter, gamma (γ). Greater γ values produce simpler models, whilst lower γ values mediate denser models. When γ = 0, the EBIC is equal to the Bayesian information criterion. In the ensuing studies, γ will be set to 0.25, which produces a conservative network, where specious edges between two variables, which have no actual relationship but are linked statistically, are discarded. This provides a more interpretable network where only statistically meaningful edges remain, it also controls for Type 1 errors, that may occur from sampling error (Epskamp & Fried, 2018).

Walktrap, a community detection algorithm (Pons & Latapy, 2006), which deterministically assigns variables to specific dimensions, finds the number of clusters identified which equals the number of latent factors in the dataset. A transition matrix is calculated, whereby each element signifies the probability of one node crossing to another. This is determined by the strength of the node or the sum of connections to each node. Random walks, using a number of steps (e.g., 4), combined with the transition matrix, search for credible destinations. Mediated by Ward's agglomerative clustering approach (Ward, 1963), each node begins in one cluster and then amalgamates with adjacent clusters. Clusters are determined based on the squared distances between each one, thereby minimising the sum of squared distances between other clusters.

It has been evidenced that clusters in the network are equivalent to latent variables, demonstrating both an empirical and mathematical feature of the method (e.g., Constantini et al., 2015; Golino & Epskamp, 2017). The dimensions revealed by EGA are deterministic, whereby no guidance is needed from the researcher. This

affords advantages over other exploratory methods as the items and quantity of facets are immediately interpretable, without any deliberation about how to interpret the factor loadings of individual indicators (Golino & Demetriou, 2017).

In relation to the required minimum sample size for EGA, psychological studies tend towards small datasets. Indeed, EGA has provided reliable results with a sample of 100 This is due to the LASSO estimator, which provides a penalised maximum likelihood estimation, which protects against overfitting (Golino & Epskamp, 2017). EGA has previously identified facets within personality networks as well as symptoms of psychopathology (Christensen, Cotter, Silvia, 2019; Christensen, Gross et al., 2019).

3.5 Network Item Redundancy

Establishing which indicators to include in a psychometric measure involves determining the specificity and coverage of the traits across the construct, whilst also reducing idiosyncrasies (DeVellis, 2017; McCrae, 2015). Reducing redundant items minimises indicator and facet confounding (Hallquist et al., 2019), for example, two items from the Narcissistic Personality Inventory (Raskin & Hall, 1979), 'I like to be the centre of attention' and 'I really like to be the centre of attention', can clearly be seen to have a common underlying focus. Thus, these items are not unique but comprise a single unique aspect of the factor it is part of.

A method of reducing indicators from a pool of items is found in the 'EGAnet' package (Golino & Christensen, 2020) in R. Network redundancy uses an adaptive alpha multiple comparison correction method (Perez & Pericchi, 2014), which has been shown to have the highest accuracy and few false positives and negatives (Christensen, Golino & Silvia, 2020). The method further uses a weighted topological overlap from the 'wTo' package in R (Gysi et al., 2018), which measures the similarity between the nodes and facilitates multiple comparison corrections to suggest which items may be redundant.

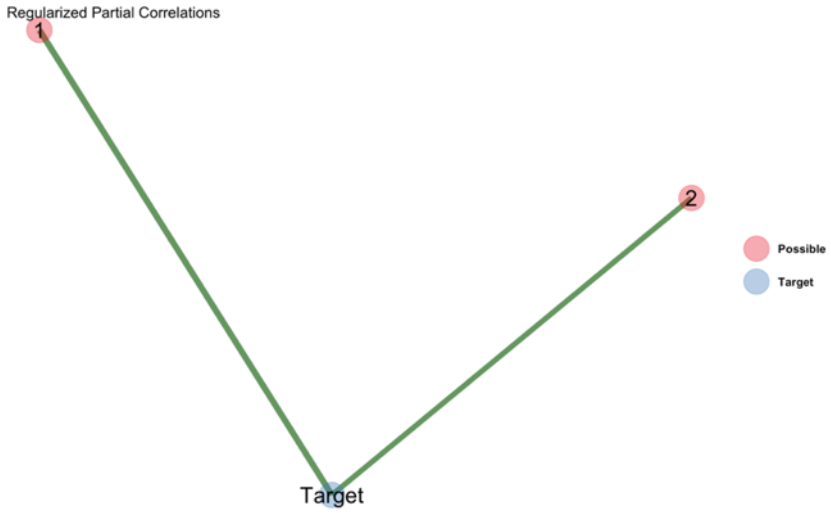
When deciding which items are redundant, there is an element of subjectivity involved (Christensen, Golino & Silvia, 2020). Indicators need to be considered from a qualitative perspective, with the subtleties of the semantics reflected upon. For example, returning to the Narcissistic Personality Inventory (Raskin & Hall, 1979), the items, 'I see myself as a good leader', 'I am a born leader' and 'I would prefer to be a leader', would be considered to have a high topological overlap. Yet semantically there are subtle differences and may be interpreted by participants in varying ways.

Thus, it is not always pertinent to select items which tap into the most general form of the construct or sub-factor.

The node redundancy command produces an output, with a choice of possible redundant nodes and an associated redundancy chain plot (Figure 5),

Figure 3

Redundancy Chain Plot



In these plots, the edges between the indicators denote statistically significant redundancies, whilst the thickness of the lines relate to the regularised partial correlations of those items in the network. The target item is the focus when considering which items may be redundant with it. This is done by perusing the item content against that of the target item. Once the process is completed, a new data matrix is returned and a further table documents the process, thereby providing transparency over the decisions taken.

3.6 Robust Correlations

In Chapter Five, correlations will be used to determine convergent and discriminant validity of the Dark Side of Humanity Scale, as well as the test-retest analysis. A correlation design provides information which shows the strength of a relationship between two variables as well as whether this relationship is significant (Weaver et al., 2017). Pearson product-moment correlation coefficients are not robust, as they are sensitive to outliers and any deviation from normality may yield ineffective results (Wilcox, 2017). If outliers are present, a robust correlation method, which guards against their effect will be used with the ‘WRS2’ package (Mair & Wilcox,

2019), in R. It uses an M-estimator, whereby the ‘M’, indicates maximum likelihood-type, using the Pearson’s correlation formula (Mair & Wilcox, 2019).

3.7 Item Response Theory

3.7.1 Mokken Scale Analysis

Mokken Scale Analysis (MSA), is a non-parametric form of IRT. It was developed to address the limitations of scales measuring ordinal latent variables, which may not be suitable for parametric IRT analysis. MSA can be used to investigate both multidimensional and unidimensional measures (Sijtsma & van der Ark, 2017). The main ethos of MSA is to determine whether hierarchically ordered Mokken scales are present within a set of indicators. Hierarchy is determined when assumptions of monotone homogeneity (MH) and invariant item ordering (IIO), which was previously referred to as double monotonicity, are met (Watson, Wang & Thompson, 2014).

Although an assumption of the MH model is unidimensionality, it is acknowledged that this is rarely met. Violations are not considered problematic if there is a clear, dominant first factor or Mokken scale (van der Linden, 2016). The monotone aspect refers to monotonicity, which means that participants who endorse more difficult items (requiring a higher level of the latent construct), is related to a higher probability of endorsing easier items. Therefore, the higher level of the construct the participant embodies, the more likely they are to achieve scores representative of their trait level (Sijtsma & van der Ark, 2017).

Local independence is also an assumption, whereby items should only be related to each other by the construct being measured. This suggests that if the common factor from any two items is partialled out, their residual covariance is zero. Again, this assumption is never precisely met and violations are not a cause for concern (van der Linden, 2016). However, a possible cause of local dependence is that the quality of the scale or scale indicators vary across samples and/or across groups. Two participants who embody the same level of the construct but are members of two different groups, may have different response probabilities for the same items (Sijtsma & van der Ark, 2017). Hence, the analysis will be conducted with the samples split separately across and between age and sex.

The ordering of the Mokken scales is facilitated by IIO. The ordering is determined by the mean scores of the indicators, whereby the item ordering is equivalent for participants with differing levels of the latent trait. The endorsement of

any two indicators by the same participant should only relate to their position on the latent trait continuum (Van Schuur, 2003). This process facilitates comparison between groups or samples (Sijtsma & van der Ark, 2017). The MH model provides a quantitative elucidation of the scale scores, whilst IIO offers a qualitative explanation. MSA uses scalability coefficients to assess the quality of the indicators with the automatic item selection procedure (AISP; Van Schuur, 2003; Sijtsma & Hemker, 1998; Sijtsma & Molenaar, 2002)

AISP is used at the start of the analysis to separate the items into Mokken scales (Van der Ark, 2012). Items which do not discriminate well between participants across the continuum of the construct do not contribute to a reliable IIO. AISP identifies these items as well as those which do not belong to any of the Mokken scales and determines these as unscalable. This may be due to the item being of a poor quality within the set of items or within the sample. However, these items may not necessarily be defunct as they may function well in different samples (Sijtsma & van der Ark, 2017).

Coefficient H^i values, accompanied by criterion or crit values, denote each items scalability. A crit value is calculated as a single value determined by the combined *H* coefficients of the Mokken scale items, where values of >80 are considered to violate MH (Watson, Deary & Shipley, 2008). If significant violations are identified, then the item is removed, and the analysis is repeated. If more than one item has violations with crit values of >80, they are excluded one by one (Van Schuur, 2003). IIO, which is indicated by H_{trans} (H^T), reverses the role of participants and items, conveying the extent participants agree on item ordering, finally the *H* coefficient is provided for the scale. The *H* and H^T coefficients indicate the strength of the scale. $H >$ 0.3 signifies a weak scale, $H >$ 0.4, a moderate scale and $H >$ 0.5, a strong scale, with the same ranges for H^T (Watson, Deary & Shipley., 2008). If the MH model fits the data and IIO is evidenced, then the ordering of the scale items can be determined as robust for use within populations (Sijtsma, Meijer & van der Ark, 2011).

A definitive minimum sample size has not been suggested for the AISP to obtain stable results. This is dependent on the positive difference between the item score *H* values and lower bound coefficient, which in the study found in Chapter Four, will be a minimum coefficient value of 0.40. Consequently, it is not possible to determine whether the sample size is sufficient until participants have taken part in the survey and the data collected and analysed (Straat, van der Ark & Sijtsma, 2014).

3.7.2 Generalised Partial Credit Model

The generalised partial credit model (GPCM; Muraki & Muraki. 2016), was the chosen IRT method as it does not make any assumptions about the rank ordering of the Likert scale categories (Hays et al., 2000) and is useful for both informing scale development and assessing models (Langer et al., 2008). The GPCM determines the degree in which the scale items for each construct captures the latent trait levels in participants and is specifically designed for polytomous data. The “generalised” part of the name means that the method includes a discrimination parameter for evaluating the measurement quality of an item (Muraki & Muraki, 2016).

The GPCM will be conducted through The Multidimensional Item Response Theory package for R (Chalmers, 2012), which was specifically developed for estimating multidimensional item response theory parameters. Technological advances have facilitated multidimensional IRT analysis as a practical statistical methodology, which is useful in psychological research as constructs are inevitably multidimensional (Chalmers, 2012).

GPCM reveals the increasing probability a participant has of endorsing items across the scale options as their latent trait level increases. For example, those with higher levels should be more likely to endorse category five on a Likert scale rather than four, if the lower category does not represent a meaningful option to them (Stewart et al., 2012). The measurement indices used signify the measurement properties of both the indicators and participants. The latent trait is represented by theta (θ), which has a mean of 0 and a standard deviation of 1, whilst the slope parameter (*a*), which is comparable to item discrimination, shows how well the scale indicator identifies participants at differing levels of the trait (Hussein, 2010). Based on the guidelines by Baker (2001), the slope values are determined by: very low (0.01-0.34), low (0.35-0.64), moderate (0.65-1.34), high (1.35-1.69) and very high (>1.70).

Participants receive partial credit for successfully completing a step, with the assumption that all steps below it have been completed first. The threshold, or step difficulty parameter (*b*), indicates how difficult it is for a participant to achieve a 50% probability of endorsing a scale option, given the participants level of the latent construct. The number of step difficulty parameters is equivalent to the number of scale categories minus one (Embretson & Reise, 2000). For example, if an item has six response options, then there will be five steps. Step one moves from the first

option to the second, step two from the second to the third option, step three from the third to the fourth option and step four, from the fourth to the fifth scale option (Langer et al., 2008). A high positive value suggests the item is located at the higher end of the construct or that few participants respond in the higher categories. A negative value indicates the item is easy to endorse or that few participants respond in the low categories (Embretson & Reise, 2000).

3.8 Participant Selection

Participants across studies will mainly be recruited from the crowd sourcing platform Prolific, in order to obtain large sample sizes with roughly equivalent numbers across groups. Although using social media groups such as Facebook offer a no cost option it cannot be guaranteed that a large enough sample of participants can be obtained or equal groups of males and females across ages (Kayrouz et al., 2016; Ramo et al., 2014). Using crowd sourcing websites such as Prolific allows researchers to take a shift away from the use of psychology students prevalently used in studies (e.g., Jonason & Webster, 2010), which have been shown to significantly skew results (Henrich, Heine & Norenzaya, 2010a, 2010b).

Yet crowd sourcing websites, despite their increasing popularity, have been criticised. These criticisms have included suggestions that participants may not fully focus on the survey whilst completing it, especially with no researcher present to reiterate the study instructions (Ramsey et al., 2016). However, this can be applied to any participant taking an online survey, whether they are members of a Facebook group, university students or an individual who chooses to be paid for participating. Indeed, a limitation of using Facebook groups or undergraduates is that they can restrict the scope of the sample, even though they are convenient and easily accessible to the researcher (Bornstein et al., 2013).

This highlights the benefits of crowd sourcing websites, such as Prolific, which can be valuable in providing participants who are outside of the Western, Educated, Industrialised, Rich and Democratic population (Gosling et al., 2010; Henrich, Heine & Norenzaya, 2010a, 2010b). In a comparison study of crowd sourcing platforms (Peer et al., 2017), Prolific was shown to offer a more diverse population sample opposed to Amazon Mturk, although the quality of data provided was equivalent in relation to the response rate, reliability and dishonesty of participants across the two platforms. Prolific was also found to be superior to MTurk in terms of usability, whilst also offering researchers an alternative option for robust research (Palan & Schitter, 2018).

3.9 Ethics

The studies which contribute to this thesis are not experimental and no manipulation will be conducted, ensuring there will be minimal impact on the participants. All studies will follow the guidelines of the British Psychological Society (BPS, 2018) and General Data Protection Regulations (Gov.UK, 2018). Studies will only be conducted once ethical approval is granted by the University ethics committee. All data across studies will be anonymous, however participants will have a unique identification code, with which they can contact the researcher and/or supervisor to withdraw their data from the study they participate in. If this situation arises, then the data becomes confidential, rather than anonymous.

3.10 Supplementary Material

A soft copy of all statistical output from the studies contained within this thesis is available on request. Copies of the request for ethical approval, the confirmation of approval and the measures used are provided in Appendix 3, 4 and 5. The R code used for the methods discussed can be found in Appendix 2. Further supporting evidence considered essential, will also be provided, including examples of the statistical output from the analyses.

3.11 Summary of statistical methods to be used in the studies

Study One

This study will examine the psychometric properties of the widely used measures in dark personality research, as well as investigate sex and age invariance.

Confirmatory Factor Analysis.

Measurement Invariance

Exploratory Graph Analysis.

Mokken Scale Analysis.

Study Two

The focus of this study is the development of the Dark Side of Humanity Scale.

Exploratory Graph Analysis.

Generalised Partial Credit Model.

Confirmatory Factor Analysis.

Measurement Invariance.

Study Three

The convergent and discriminant validity as well as the temporal reliability of the Dark Side of Humanity Scale is assessed in this study.

Confirmatory Factor Analysis.

Robust Correlations.

Chapter Four

The Dark Tetrad Measures: An Interrogation of the Factor Structures, Construct Validity and Invariance Across Sex and Age.

4.1 Introduction

The evidence from the literature review provides the rationale for a study to investigate the age and sex invariance of the widely available scales which measure psychopathy, Machiavellianism, narcissism and everyday sadism. Although past research suggested that the scales are mainly sex variant (Anestis et al., 2011; Marion & Sellbom, 2011), there was a scarcity of studies which investigated whether the scales measure the constructs homogenously across ages (Klimstra et al., 2020). Criticisms further contended that the measures may be incongruent with the constructs they seek to measure (e.g., Persson, Kajonius & Garcia, 2017b). Studies which have offered critiques of the scales have mainly used classical test theories (CTT), such as confirmatory and exploratory factor analysis (Chiorri et al., 2017; Christian & Sellbom, 2016), some authors, however, expanded their methods of investigation with the use of item response theories (IRT; Dinić, Bulut Allred et al., 2020; Garcia et al., 2017; Gummelt et al., 2012) and network analysis (Truhan et al., 2020). Further advances in psychometric research provides the opportunity to analyse the scales in alternative ways, which may offer new insights.

Hence, the current study will address the following considerations. Firstly, whether the widely available scales used in Dark Tetrad (DT), research are sex invariant, such that they sufficiently provide equivalent psychometric properties when used in mixed sex studies, whereby intergroup comparisons are meaningful (Putnick & Bornstein, 2016). Results from studies tend towards comparing intergroup mean differences (Gluck et al., 2019), therefore ascertaining whether these results are unaffected by group membership is relevant (Meade et al., 2008). For example, the literature review revealed that the Assessment of Sadistic Personality held inter-sex metric invariance in student aged groups (ASP; Plouffe, Smith & Saklofske, 2018). This implies that the factor structure is equivalent for both males and females, such that the same observed variables are equally important to both groups understanding of the everyday sadism construct. However, the authors added a caveat that generalisability had not been established beyond a student age group (Plouffe, Smith & Saklofske., 2018). Yet, the measure did not achieve intersex scalar invariance,

which would indicate that the intercept or mean of the latent variable varied between groups, when the intercepts of the variables were constrained to be equal for both groups. This may indicate possible measurement bias, suggesting that either age and/or sex influences the way participants respond to the indicators (Bialosiewicz et al., 2013).

This leads to the second consideration of age invariance for which the literature review emphasised the scarcity of research (Kawamoto et al., 2020). With the exception of the Short Dark Triad (SD3; Jones & Paulhus, 2014), which for the scale development study, used a general population sample (M_{age} = 30.72, *SD* = 11.09), other widely available DT measures (the Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995), Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979), Mach IV (Christie & Geis, 1970), ASP (Plouffe, Saklofske & Smith. 2017), Short Sadistic Impulse Scale (SSIS; O'Meara et al., 2011), Varieties of Sadistic Tendencies Scale (VAST; Paulhus & Jones, 2015) and the Dirty Dozen (DD; Jonason & Webster, 2010)), used student samples, with mean ages ranging between 18.74 (*SD* = 3.79) and 24 (with no *SD* provided), whom for many researchers are easily available and may provide larger sample sizes (Trzesniewski & Donnellan, 2010). Yet, the limitations of doing so removes the ability to assess whether the measures are homogenous across age ranges.

Age should be a fundamental factor when investigating personality and as such, this study will employ the theoretical framework as suggested by Arnett (2000). Individuals, who encompass the 18-25 age group are neither classified as adolescents nor young adults, but empirically and theoretically reside within a distinctive group of emerging adults. This temporal period, characterised by significant biopsychosocial changes, may include the instability of peer group relationships, a volatile sense of self-identity, an increase in risk-taking and changes in cognitive elasticity (Reifman et al., 2007; Spear, 2000). It can further mediate poor behavioural and impulse control (Arnett & Taber, 1994; Barlett & Barlett, 2015; Reifman et al., 2007).

The age of 25 may appear to be an arbitrary cut-off point and it could be questioned as to whether there is a significant difference in behaviour between the ages of 25 and 26. However, the generalised theory suggests that stability related to self-identity and societal roles is attained from aged 26 and over, following the transition from emerging adulthood into young adulthood (Arnett, 2000). Changes are found in the processing of socio-emotional information and emotional stimuli reactions

are more controlled. Decreases are also seen in negative emotionality and aggressiveness (Furlong, 2009). A further consideration is that the realisation of adults roles occur at an older age than in previous decades. Societal norms of the mid-twentieth century, which saw marriage and workplace entry occur shortly after finishing education, now happen later (Côté & Bynner, 2008). Many employment opportunities increasingly require at least a bachelor's degree, indicating people spend longer in education, which delays the onset of traditional adult roles. Yet from the age of 26 and over, individuals are considered to settle into long-term adult roles, including employment and relationships (Arnett, 2000; Reifman et al., 2007).

Relative to the current research and the DT constructs, research has shown that impulsivity, a characteristic associated with emerging adults, is related to Machiavellianism, narcissism and psychopathy (e.g., Hart et al., 2017; Szabó & Jones, 2019). Consideration of the traits and behaviours embodied by emerging adults are salient for dark personality research, especially when the development of DT measures have mainly used samples within the 18-25 age group (e.g., Levenson et al., 1995), as discussed above. More specifically, scales that were developed using psychology students, such as the DD, may be thought paradoxical.

Psychology students have been evidenced to hold the least malevolent personalities comparatively to law, business and economics students (Hanel & Vione, 2016), with female psychology students scoring the lowest mean scores conversely to females in the other domains. Moreover, female students of law and business scored a higher mean in Machiavellianism than male psychology students (Vedel & Thomsen, 2017). In respect of older age groups, a divergence has been found in the way they endorsed scale items in the Mach IV compared to student samples (Moss, 2005), whilst evidence has supported the decline of secondary psychopathy traits and behaviours of antisociality, impulsivity and criminality with age, in both institutionalised and general population samples (Gill & Crino, 2012; Harpur & Hare, 1994). Consequently, both age and sex invariance should be important factors in dark personality measures.

The final consideration is whether the measures are incongruent with the constructs, whereby the definitions of the constructs as determined by the authors, are incompatible across and between age groups as well as males and females. In relation to psychopathy, it is unclear whether the conceptualisation of primary and

secondary facets, which were divergent theories (Salekin, Rogers et al., 1998), synthesised by Hare (1980), for his research focus of institutionalised males, transcends to mixed sex and age general population samples. The LSRP has a two-factor structure of primary and secondary psychopathy, however, it does not specifically measure criminal behaviour but impulsivity and a self-defeating lifestyle (Brinkley et al., 2008). Yet, the SD3 has indicators which address criminality, for example, 'I have never gotten into trouble with the law'. There is a debate as to whether the secondary facet is pertinent to general population samples. Although some contend that excluding it from measures removes a salient dimension of the construct (Neuman, Vitacco et al., 2005), others put forth that it may be redundant (Dinić, Wertag et al., 2020). Indeed, associations of secondary psychopathy incorporate high levels of dysfunctional negative emotionality, external comorbidities of substance abuse as well as symptoms of borderline personality disorder (Sellbom & Drislane, 2020). This suggests the facet may be a valid dimension in forensic and clinical populations, but the antisocial and criminal behaviours incorporated into some measures may be disproportionate to the samples they seek to assess (Boduszek, Dhingra, et al., 2016; Boduszek & Debowska, 2016; Cooke & Logan, 2015; Skeem & Cooke, 2010).

Discordance with the narcissism construct is one of the criticisms directed at the NPI (Raskin & Hall, 1979), as it is discussed as mainly being a measure of grandiose narcissism, minimising the vulnerable facet (Atlas & Them, 2008; Corry et al., 2008). The measure is also considered to be incongruent with the female manifestation. This is attributed to characteristics of exhibitionism, entitlement, leadership and authority being antithetical to stereotypical female behaviours, whereby enactment of these behaviours violates culturally held Western expectations (Tschanz et al.,1998; Weiser, 2015). Yet, as discussed in the literature review, assessments of narcissism based on self-report may mediate sex differences, leading some researchers to assume sex role stereotypes (Baez et al., 2017). It should be considered that in relation to Narcissistic Personality Disorder, on which the NPI is based most facets were found to be sex invariant (Hoertel et al., 2018). However, the DSM criteria did not include leadership and authority, as these facets are not considered core facets of the construct (Rosenthal & Hooley, 2010), and this did not change through further DSM publications (APA, 2013). Thus, there may be a problem with the measure more so than the construct relative to intersex invariance.

The instability of the Machiavellianism measure, the Mach IV, was well documented. Researchers reported that the original structure is not a good fit for data, with uninterpretable factors emerging (Ahmed & Stuart, 1981; Monaghan et al., 2016), and some suggesting the Mach IV does not address the female manifestation (Brown & Guy, 1983; O'Connor & Simms, 1990). Across decades of research, there has been no agreement on a stable factor structure (e.g., Monaghan et al., 2016; Panitz, 1989; Williams et al.,1975), nor on which factor indicators should load (Miller, Nicols & Konopaske, 2019). A further consideration is the contention that Machiavellianism and primary psychopathy are parallel constructs (Czibor et al., 2017; McHoskey et al., 1998; Miller, Hyatt et al., 2017; Rogoza & Cieciuch, 2018; Vize, Lynam et al., 2018). The methods used in this current study will seek to further investigate whether Machiavellianism as it is conceptualised in the widely used measures, is a valid construct or if it is, as suggested, parallel to primary psychopathy.

Turning to everyday sadism, as highlighted in the literature review, there is sparse research available in relation to the psychometric properties of the measures (e.g., Dinić, Bulut Allred et al., 2020). Two studies to date have investigated the sex invariance of the ASP (Min et al., 2019; Plouffe, Smith & Saklofske, 2019), thus, it is not possible to ascertain whether the remaining scales which address this construct offer sex invariance and if so, to what degree. Moreover, the measures which are commonly used across the dark tetrad constructs, have, on the whole, not been analysed for age invariance and thus, have not been subjected to psychometric scrutiny to ensure their intergroup generalisability.

The above discussion provides the basis to investigate the factor structures, scale indicators, sex and age invariance of the Levenson Self-Report Psychopathy Scale, Narcissistic Personality Inventory, Mach IV, Dirty Dozen, Short Dark Triad, Assessment of Sadistic Personality, Short Sadistic Impulse Scale and Varieties of Sadistic Tendencies, using confirmatory factor analysis (CFA), exploratory graph analysis (EGA) and Mokken Scale Analysis (MSA). Not all prevalent DT scales, for example, The Psychopathic Personality Inventory Revised (PPI-R) and the Self-Report Psychopathy Scale (SRP-4), were incorporated within this study as they charge per use. Therefore, the decision to use freely available scales which are extensively used in research, was taken.

Groups comprising males and females 18-25 and males and females over the age of 26 (Arnett, 2000), will determine whether the scales are invariant across age

and sex. It is predicted that the structures of the scales will differ across age and sex. It is further predicted that a reduced number of indicators as well as varying structures will be evidenced by MSA across age and sex.

4.2 Method

4.2.1 Participants

An online study was designed to investigate whether the widely used DT scales were invariant across age and sex. Invitations to participate were posted on social media websites, the University of Derby participant scheme and a crowd-data provider, Prolific, to ensure both a cross-sectional sample, and roughly equal group sizes. Ethical approval was received from the University of Derby (Appendix 3.1). Students from the University of Derby were awarded 2 credits and the Prolific participants were paid £1.67, for completion of the survey. All participants were over the age of 18. Initially there were 656 recorded responses, 41 participants had missing data, which were removed. A final sample of 614 participants contributed data to this study, with an age range of 18-86 (M_{age} = 31.92, *SD* = 13.64). There were 299 females (48.7%) and 306 males (49.8%), 6 participants (1%) preferred to self-define and 3 preferred not to say (0.5%). As a core focus of the study was the similarities and differences between males and females, participants who chose to self-define or preferred not to say were removed from the analyses, leaving 605 people (M_{age} = 32.00, *SD* = 13.70), who contributed data to this study.

The sample was defined by the following groups: Females 18-25 (*n* =148; M_{age} = 21.78, *SD* = 2.08); Males 18-25 (*n* = 156; M_{age} = 21.42, *SD* = 2.08); Females 26 and over (*n* = 151; M_{age} = 43.62, *SD* = 13.02); Males 26 and over (*n*=150; M_{age} = 41.41, *SD* = 11.58).

4.2.2 Measures

The Mach IV (Christie & Geis, 1970).

The Mach IV (Appendix 3.2.1), comprises 20 items which are measured on a 7-point Likert scale from *strongly disagree* (1) to *strongly agree* (7). The scale can be used either as a unidimensional or three factor model. The three factors comprise; Tactics, ‘Anyone who completely trusts anyone else is asking for trouble’, Views, ‘Most people are brave’, and Morality, ‘All in all, it is better to be humble and honest than to be important and dishonest’. The total scale reliability has been shown to be adequate

α =.74 (Abell & Brewer, 2018), with subscale reliability of tactics α =.70, views α = .61 and morality α = .07 (Monaghan et al., 2016).

The Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995)

The LSRP (Appendix 3.2.2), is a 26-item scale, which incorporates two factors; Primary psychopathy measures callous and manipulative interpersonal characteristics and secondary psychopathy taps into impulsivity and antisocial behaviours. Measured on a four-point Likert scale from *disagree strongly* (1) to *agree strongly* (4), items for primary psychopathy include, 'I often admire a really clever scam', and for secondary psychopathy, 'Love is overrated'. Reliability has been found to be good; Total scale; α =.87; Primary; α =.86 and Secondary; α =.63 (Marion & Sellbom, 2011).

Narcissistic Personality Inventory (NPI: Raskin & Hall, 1979).

The NPI (Appendix 3.2.3), comprises 40 forced choice items, where participants choose between two statements, one which is a narcissistic response, 'I am an extraordinary person' and the other non-narcissistic, 'Compliments embarrass me'. The scale comprises seven factors: Authority, Self-sufficiency, Superiority, Exhibitionism, Exploitativeness, Vanity and Entitlement. A score of 1 was given for each narcissistic response and 0 for the non-narcissistic choice. Previous research has found that the NPI demonstrates good internal consistency (α =0.89; Carter & Douglass, 2018) The alpha reliabilities of the sub-factors from this study can be found in Table 1.

The Dark Triad Dirty Dozen (DD: Jonason & Webster, 2010)

This is a 12-item self-report scale (Appendix 3.2.4), which measures the Dark Triad traits, psychopathy, narcissism and Machiavellianism on three subscales comprising four items each. Participants were asked to indicate the extent to which they agreed or disagreed to each scale item on a five-point Likert scale from *strongly disagree* (1) to *strongly agree* (5). Scale items include 'I tend to manipulate others to get my way', (Machiavellianism), 'I tend to be cynical', (psychopathy), and 'I tend to want others to admire me' (narcissism). Scale items were averaged together to create the three subscales. Previous research has found that the DD demonstrates good reliability with Machiavellianism; α =.94, narcissism α =.84 and psychopathy α =.89 (Pan et al., 2018).

The Short Dark Triad (SD3: Jones & Paulhus, 2014)

This measure comprises 27 items (Appendix 3.2.5), which are rated on a 5-point Likert scale from *strongly disagree* (1) to *strongly agree* (5). Scale items consist of statements such as 'Make sure your plans benefit you, not others' (Machiavellianism), 'I insist on getting the respect that I deserve' (narcissism) and 'I like to pick on losers' (psychopathy). The items were averaged to create indices for Machiavellianism, narcissism and psychopathy. Previous research has found that the SD3 demonstrates good reliability with Machiavellianism α = 0.71, narcissism α = 0.74 and psychopathy α = 0.77 (Jones & Paulhus, 2014).

Varieties of Sadistic Tendencies Scale (VAST; Paulhus & Jones, 2015)

The VAST comprises 16 items (Appendix 3.2.6), with sub-scales of direct (9 items) and vicarious (7 items) sadism, which are rated on a 5-point Likert scale from *strongly disagree* (1) to *strongly agree* (5). Scale items for direct sadism consist of statements such as 'I enjoy physically hurting people', and vicarious sadism, 'In video games, I like the realistic blood spurts'. Items were averaged to calculate the subscales. The current study found that the VAST demonstrated adequate reliability α = 0.75.

Short Sadistic Impulse Scale (SSIS; O'Meara et al., 2011)

The SSIS is a 10-item scale (Appendix 3.2.7), where responses are recorded in a dichotomous form, using the categories, 'like me' and 'unlike me'. The maximum score is 10 with a minimum of zero. Scale items comprise statements such as 'I have hurt people because I could'. Previous research has found that the SSIS demonstrates good reliability α = 0.87 (Schumpe & Lafrenière, 2016).

The Assessment of Sadistic Personality (ASP; Plouffe, Saklofske & Smith, 2017)

The ASP comprises nine items (Appendix 3.2.8), which are rated on a 5-point Likert scale from *strongly disagree* (1) to *strongly agree* (5). Scale items consist of statements such as 'Watching people get into fights excites me'. The ASP has demonstrated good reliability α = 0.87 (Plouffe Smith & Saklfoske, 2018).,

4.2.3 Procedure

The invitation to participate provided a link to the online survey hosted by Qualtrics, where involvement was voluntary and anonymous. Exclusion criteria requested that participants had no clinically diagnosed mental health nor substance abuse issues. Participants were initially briefed about the study and its aims. Following

completion of the consent form, the participants were presented with demographic questions and the questionnaires. Contact information was given for the researchers, should participants wish to ask any questions about the nature of the study. They were also informed that they could withdraw their data within two weeks of taking part and their responses removed from the study, by providing the unique identification code they had entered before beginning the survey.

4.2.4 Analytic Strategy

The methods used in this study are discussed in depth in the Methodology Chapter. In brief, the design of this study incorporated confirmatory factor analysis (CFA), exploratory graph analysis (EGA) and Mokken scale analysis (MSA), to facilitate a more extensive understanding of the measures across groups. EGA and MSA provide advantages over CFA, when the scales being investigated consist of items which measure different facets and the extent of dimensionality is disputed (Emons et al., 2012). The total sample was split into four groups, of males and females 18-25 and males and females 26 and over.

The analyses used in this study were all conducted in R Studio, Version 1.2.1335 and Jamovi Version 1.0.5.0. Descriptive statistics were investigated for each sample. As multivariate normality of the data was not achieved (Appendix 3.3.1), the WLSMV estimator was employed for CFA analysis, which used diagonally weighted least squares to estimate the model parameters and the full weight matrix to produce robust values (Rosseel, 2020). As the Likert measures ranged from four to seven-point scales, thus, not conceptually meeting the assumption of measurement at an interval-scale level (Bertl et al., 2017), the WLSMV estimator is recommended with ordered categorical as well as binary indicators (Bertl et al., 2017; Rosseel, 2020; Salekin et al., 2014). To align with previous dark tetrad research (e.g., Persson, Kajonius & Garcia, 2017b), the CFA's conducted in both study one and two used polychoric correlations for Likert measures and tetrachoric correlations for binary scored measures.

The CFA model fit was assessed using chi square (χ^2), Confirmatory Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). A good model fit was determined when CFI and TLI values were $\geq .90$, in line with Kline's (2015), recommendations. RMSEA values evidenced a good fit if they were .06 or below and an adequate model fit was accepted at .08 or below. Models

were rejected if they displayed a value above 1.0 (Brown, 2015; Hu & Bentler, 1998; Marsh, Hau & Wen, 2004).

There is a caveat to the confirmatory factor analysis. Two scales used in the study, the Narcissistic Personality Inventory (NPI: Raskin & Hall, 1979) and the Short Sadistic Impulse scale (SSIS: O'Meara et al. 2011), both use a dichotomous response format. Previous research which has investigated the factor structure of the NPI have used a variety of methods including principal components analysis (PCA; Emmons, 1984; Kubarych et al., 2004; Raskin & Terry, 1988) and exploratory factor analysis (EFA; Ackerman, Witt et al., 2011; Corry et al., 2008; Emmons, 1987). In the development study of the SSIS, PCA was used to investigate the scales structure as well as parallel analysis. Research which has since investigated the SSIS have used a five-point Likert scale (Dinić, Bulut Allred et al., 2020), or a seven-point scale (Min et al., 2019).

As discussed by Ackerman, Witt et al., (2011), confirmatory factor analysis may not sufficiently address the statistical complexities associated with dichotomous data, and it is possible that ensuing results will be unreliable. It was further stated that the factor structure of the NPI may have not been successfully replicated, as the analytic methods are not ideal for determining the model due to the data being dichotomous (Boldero et al., 2015). Traditional factor analysis methods were not developed for use with binary items (Woods, 2002), yet, Ackerman, Witt et al., (2011), proceeded to investigate the NPI with CFA, using WLSMV as the estimator (Rosseel, 2020). As such, this study will analyse the SSIS and NPI with CFA, although the findings may be unreliable.

Measurement invariance through multigroup nested CFA models, was used to determine whether the scales held for configural, metric and scaler invariance. Analysis was conducted in R using lavaan (Rosseel, 2020). The CFI and RMSEA considerations were the same as the CFA criteria discussed above. Fit indices were assessed by the change in CFI and RMSEA, which are less sensitive to sample size than the chi-square and more sensitive to a lack of invariance than chi-square (Meade et al., 2008). If the difference in the fit indices of ΔCFI and ΔRMSEA (with Δ denoting 'difference'), between a model and the preceding less constrained model is not larger than 0.01 for ΔCFI and equal or less than 0.015 for ΔRMSEA, then it was considered

that the level of measurement invariance had been achieved (Chen, 2007; Cheung & Rensvold, 2002; Marsh, Nagengast & Morin, 2013; Rudnev et al., 2018).

Exploratory Graph Analysis was conducted through the EGAnet package (Golino & Christensen, 2020) in R. EGA identifies clusters in multidimensional and unidimensional date for both Likert and dichotomous data. As discussed in the methodology chapter, EGA has advantages over more traditional methods such as exploratory factor analysis. As the analysis is data driven, subjective interpretation is removed due to the formation of clusters, which are displayed in the network graph. EGA has been found to perform as well as the most accurate traditional methods of exploratory analysis (Christensen & Golino, 2020).

Mokken Scale Analysis (MSA), determined whether hierarchically ordered (Mokken) scales, existed within scale items. Hierarchy is determined when assumptions of monotone homogeneity (MH) and invariant item ordering (IIO) are met (Watson. Wang & Thompson, 2014). The method ensured that the items are stochastically independent, indicating that responses to any two items by the same participant only related to their positioning on the latent trait continuum rather than any other consideration (van Schuur, 2003). Each item received a coefficient H^i value and a criterion (crit) value, whereby values of >80 violated MH (Watson, Deary & Shipley, 2008). IIO, which is indicated by H_{trans} (H^T) and inverts the role of participants and items, displayed the extent participants agreed on item ordering, finally the H coefficient was provided for the scale. The H and H^T coefficients denote the strength of the scale. H > 0.3 indicates a weak scale, H > 0.4, a moderate scale and H_s > 0.5, a strong scale, with the same ranges for H^T (Watson, Deary & Shipley, 2008).

4.2 Results

Mean scores and alpha reliabilities are shown in Table 1. Males 18-25 were generally found to score a higher mean than the other samples, whereas females 26 and over mainly scored the lowest mean. The alpha reliability for the Mach IV morality subfactor, was exceptionally low, which could provide a partial explanation as to why researchers use the scale as a unidimensional measure (e.g., Abell & Brewer, 2018).

Table 1

Mean Scores and Alpha Reliability for the Widely Used DT Measures

Scale	F18 Mean (*SE*)	α	M18 Mean (*SE*)	α	F26 Mean (*SE*)	α	M26 Mean (*SE*)	α
LSRP Total	51.34 (0.84)	.85	54.31 (0.88)	.86	45.63 (0.75)	.83	49.23 (0.87)	.86
Primary	29.97 (0.63)	.87	33.01 (0.68)	. 88	25.69 (0.57)	.87	29.17 (0.65)	.87
Secondary	21.37 (0.38)	.70	21.30 (0.36)	.66	19.94 (0.38)	.71	20.06 (0.40)	.75
Mach IV Total	73.97 (1.13)	.79	76.85 (1.03)	.75	69.39 (1.12)	.80	68.73 (1.24)	.85
Tactics	29.89 (0.58)	.67	30.94 (0.55)	.67	27.85 (0.56)	.66	28.90 (0.66)	.75
Views	32.93 (0.56)	.67	37.87 (0.52)	. 52	34.09 (0.62)	.69	36.18 (0.64)	.70
Morality	7.81 (0.15)	.04	8.04 (0.17)	-.29	7.45 (0.15)	.08	7.49 (0.18)	-.02
NPI Total	11.04 (0.54)	.84	12.54 (0.53)	.83	10.07 (0.56)	.87	10.79 (0.57)	.87
Authority	2.63 (0.17)	.72	2.85 (0.15)	.57	2.48 (0.17)	.73	2.72 (0.18)	.74
Self-Sufficiency	1.74 (0.11)	.44	2.15 (0.11)	.38	2.03 (0.11)	.35	2.19 (0.12)	.39
Superiority	1.28 (0.11)	.58	1.78 (0.11)	.52	1.36 (0.11)	.64	1.35 (0.11)	.64
Exhibitionism	1.33 (0.13)	.64	1.56 (0.14)	.69	1.13 (0.12)	.70	1.26 (0.13)	.71
Exploitativeness	1.34 (0.10)	.51	1.53 (0.11)	.52	1.08 (0.10)	.54	1.18 (0.10)	.47
Vanity	1.20 (0.09)	.70	0.95 (0.09)	.73	0.60 (0.07)	.63	0.68 (0.08)	.70
Entitlement	1.51 (0.11)	.29	1.80 (0.11)	.41	1.32 (0.10)	.48	1.35 (0.10)	.36
DD Total	47.66 (0.62)	.85	46.26 (0.66)	.85	51.41 (0.54)	.86	48.01 (0.67)	.88
DD Mach	16.27 (0.25)	.81	15.85 (0.25)	.78	17.38 (0.21)	.83	16.20 (0.28)	.87
DD Psychopathy	16.73 (0.25)	.77	15.58 (0.29)	.79	17.40 (0.21)	.70	16.17 (0.25)	.73
DD Narcissism	14.66 (0.29)	.79	14.82 (0.29)	.80	16.64 (0.25)	.82	15.65 (0.31)	.88
SD3 Total	68.20 (1.01)	.83	73.78 (1.00)	.83	64.53 (1.10)	.87	69.57 (1.09)	.85
SD3 Mach	27.59 (0.47)	.79	29.40 (0.45)	.78	26.23 (0.50)	.82	27.71 (0.47)	.79
SD3 Psychopathy	18.56 (0.43)	.72	20.98 (0.43)	.69	20.98 (0.43)	.76	16.79 (0.43)	.73
SD3 Narcissism	22.05 (0.44)	.73	23.39 (0.46)	.74	21.52 (0.46)	.76	22.27 (0.49)	.79
ASP Total	15.24 (0.48)	.82	17.42 (0.47)	.77	13.17 (0.38)	.80	15.66 (0.44)	.78
SSIS Total	1.21 (0.11)	.58	1.56 (0.16)	.77	0.60 (0.11)	.77	1.10 (0.15)	.79
VAST Total	29.30 (0.65)	.80	36.23 (0.69)	.78	26.95 (0.61)	.81	33.23 (0.71)	.79
Vicarious	1.95 (0.05)	.75	2.62 (0.06)	.67	1.75 (0.05)	.74	2.44 (0.06)	.71
Direct	1.74 (0.04)	.72	1.98 (0.05)	.77	1.64 (0.04)	.73	1.80 (0.05)	.77

Note: F18 = Females 18-25; M18 = Males 18-25; F26 = Females 26 and over; M26 = Males 26 and over; LSRP = Levenson Self-Report Psychopathy Scale; NPI = Narcissistic Personality Inventory; DD = Dirty Dozen; SD3 = Short Dark Triad; ASP = Assessment of Sadistic Personality; SSIS = Short Sadistic Impulse Scale; VAST = Varieties of Sadistic Tendencies.

Outliers were assessed with Mahalonobis distance and the following were found: ASP; females 18-25 (5); males 26 and over (4); females 26 and over (10); DD; females 18-25 (3); males 18-25 (1); females 26 and over (8); males 26 and over (6); LSRP; females 18-25 (1); males 18-25 (3); females 26 and over (1); males 26 and over (2); VAST; females 18-25 (6); males 18-25 (2); females 26 and over (12); males 26 and over (5); SSIS; females 18-25 (16); males 18-25 (18); females 26 and over (12); males 26 and over (16); Mach IV; males 18-25 (2); females 26 and over (3); males 26 and over (1); NPI; females 26 and over (1); males 26 and over (1).

Although researchers may view outliers as problematic and remove them to run a 'cleaner' data analysis (Aguinis et al., 2013), it has been argued that in certain research domains the presence of outliers can lead to important theoretical developments and removing them may compromise findings, thereby diminishing the accuracy of the models which portray the constructs under investigation (Aguinis et al., 2013; Neumann, Kosson & Salekin, 2017). As such, the outliers were considered to represent valid values in the data and were retained due to the nature of this study. Although outliers may impact analyses which relies on maximum likelihood estimators, the methods used in this study were not dependent on this (Aguinis et al., 2013).

Confirmatory Factor Analyses

A CFA was conducted on all scales across age and sex (Appendix 3.3.2), to investigate the model fit in line with the structures determined by the original authors, without applying modifications. For the NPI, a unidimensional model, as used by researchers (e.g., Blinkhorn et al., 2015), was assessed as well as the seven-factor structure. This was also the case with the Mach IV, where the three-factor model as well as the unidimensional scale (Abell & Brewer, 2018), was investigated. A robust method was used, which corrects non-normality of the data, specifically the degree of kurtosis in the variables, which elicits a more precise fit when multivariate non-normality is present (Chou & Bentler, 1995; Neumann, Kosson & Salekin., 2017). Models which were determined as not identifiable by the CFA analysis are not included in the results (Table 2).

Table 2

Confirmatory Factor Analysis For All Scales

Scale	Sample	χ^2	*p*	*df*	CFI	TLI	RMSEA (90%CI)
LSRP	Females 18-25	527.90	<.001	298	0.87	0.86	0.07 (0.06-0.08)
	Males 18-25	553.05	<.001	298	0.87	0.86	0.07 (0.06-0.08)
	Females 26 and over	501.64	<.001	298	0.89	0.88	0.07 (0.06-0.08)
	Males 26 and over	501.64	<.001	298	0.89	0.88	0.08 (0.07-0.09)
NPI	Females 18-25	784.01	0.003	680	0.90	0.89	0.03 (0.02-0.04)
Seven-Factor	Males 18-25	767.75	0.011	680	0.91	0.90	0.03 (0.01-0.04)
	Females 26 and over	738.33	0.060	680	0.96	0.95	0.02 (0.00-0.03)
Mach IV	Females 18-25	757.08	<.001	167	0.56	0.50	0.15 (0.14-0.17)
3 Factor	Males 18-25	568.95	<.001	167	0.59	0.53	0.12 (0.11-0.14)
	Females 26 and over	707.23	<.001	167	0.59	0.54	0.15 (0.14-0.16)
	Males 26 and over	596.29	<.001	167	0.74	0.70	0.13 (0.10-0.14)
Mach IV	Females 18-25	760.22	<.001	170	0.56	0.51	0.15 (0.14-0.16)
	Males 18-25	572.40	<.001	170	0.59	0.54	0.12 (0.11-0.13)
	Females 26 and over	719.78	<.001	170	0.59	0.54	0.15 (0.14-0.16)
	Males 26 and over	606.34	<.001	170	0.73	0.70	0.13 (0.12-0.14)
Dirty Dozen	Females 18-25	126.07	<.001	51	0.95	0.93	0.10 (0.08-0.12)
	Males 18-25	84.27	<.001	51	0.98	0.97	0.06 (0.04-0.09)
	Females 26 and over	130.51	<.001	51	0.95	0.94	0.10 (0.08-0.12)
	Males 26 and over	154.73	<.001	51	0.95	0.94	0.12 (0.10-0.14)
SD3	Females 18-25	571.12	<.001	321	0.84	0.82	0.07 (0.06-0.08)
	Males 18-25	519.69	<.001	321	0.85	0.84	0.06 (0.05-0.07)
	Females 26 and over	696.41	<.001	321	0.86	0.84	0.09 (0.08-0.10)
	Males 26 and over	584.50	<.001	321	0.86	0.85	0.07 (0.06-0.08)
SSIS	Females 18-25	43.92	0.14	35	0.95	0.94	0.04 (0.00-0.08)
	Males 18-25	50.04	0.05	35	0.97	0.97	0.05 (0.01-0.08)
	Males 26 and over	43.03	0.16	35	0.99	0.99	0.04 (0.00-0.07)
ASP	Females 18-25	36.05	0.110	27	0.99	0.99	0.05 (0.00-0.08)
	Males 18-25	63.51	<.001	27	0.97	0.97	0.09 (0.06-0.12)
	Females 26 and over	43.56	0.020	27	0.99	0.99	0.06 (0.02-0.10)
	Males 26 and over	49.76	0.005	27	0.99	0.98	0.07 (0.04-0.11)
VAST	Females 18-25	237.03	<.001	104	0.91	0.90	0.09 (0.08-0.11)
	Males 18-25	230.58	<.001	104	0.91	0.90	0.09 (0.07-0.10)
	Females 26 and over	204.99	<.001	104	0.94	0.93	0.08 (0.06-0.10)
	Males 26 and over	254.91	<.001	104	0.92	0.91	0.10 (0.08-0.11)
VAST	Females 18-25	172.95	<.001	103	0.95	0.95	0.07 (0.05-0.08)
2 Factor	Males 18-25	153.55	<.001	103	0.96	0.96	0.06 (0.04-0.07)
	Females 26 and over	171.12	<.001	103	0.96	0.95	0.07 (0.05-0.08)
	Males 26 and over	171.36	<.001	103	0.96	0.96	0.07 (0.05-0.08)

Note: LSRP = Levenson Self-Report Psychopathy Scale; NPI = Narcissistic Personality Inventory; DD = Dirty Dozen; SD3 = Short Dark Triad; ASP = Assessment of Sadistic Personality; SSIS = Short Sadistic Impulse Scale; VAST = Varieties of Sadistic Tendencies. χ^2 = Chi Square; df=degrees of freedom; CFI=Comparative Fit Index; TLI= Tucker-Lewis Index; RMSEA=Root mean square error of approximation. All χ^2 statistics were significant at p <.005.

Structural problems were found across all samples for the unidimensional model of the NPI. For the seven-factor model of the NPI, a CFA solution could not be found for males 26 and over. This was also the outcome for females over 26 with the SSIS. When a negative error variance is displayed, it implies that the regression may explain more than 100% of the variance, which may indicate a serious problem with the model fit (Wothke,1993).

Measurement Invariance

Only the scales which evidenced a good or adequate CFA fit for both pairs in each age group were eligible for measurement invariance analysis (Appendix 3.3.3). Therefore, the SSIS and VAST two-factor model was investigated for the 18-25 age. For the 26 and over groups, the ASP and the VAST two-factor model (Table 3).

Table 3

Measurement invariance SSIS, VAST AND ASP

Models	χ^2	*p*	*df*	CFI	TLI	RMSEA	$\Delta\chi^2$	ΔCFI	ΔRMSEA
18-25 samples									
SSIS									
Configural	100.91	0.009	70	0.721	0.641	0.054	-	-	-
VAST 2 factor									
Configural	275.25	0.001	206	0.869	0.848	0.047	-	-	-
Over 26 samples									
VAST 2 factor									
Configural	272.69	0.001	206	0.858	0.834	0.047	-	-	-
ASP									
Configural	66.58	0.117	54	0.958	0.944	0.039			
Metric	62.32	0.465	62	0.999	0.999	0.006	4.26	0.041	-0.033

Note: χ^2 = Chi Square; df= degrees of freedom; CFI=Comparative fit index: RMSEA=Root mean square error of approximation. ∆df = df change; ∆CFI = CFI change; ∆RMSEA = RMSEA change; $\Delta\chi^2$ = Chi Square change.

The VAST two-factor model did not achieve configural invariance across both age groups due to the CFI values being outside of the recommended guidelines (Brown, 2015; Hu & Bentler,1998; Kline, 2015), and this was also found with the SSIS for the 18-25 sample. The ASP did not attain metric invariance for the over 26 group as the change in CFI was outside the recommended values (Chen, 2007; Marsh, Nagengast & Morin, 2013). Consequently, the results of the CFA and measurement

invariance indicated that scalar invariance could not be attained, suggesting that mean group comparisons may not be meaningfully made.

Exploratory Graph Analysis

To further investigate why the measures did not provide a good fit, exploratory graph analysis (EGA) was conducted. EGA provided the opportunity to see how the indicators loaded onto the clusters or dimensions for each sample and elucidated the number of dimensions underlying the multivariate and unidimensional data.

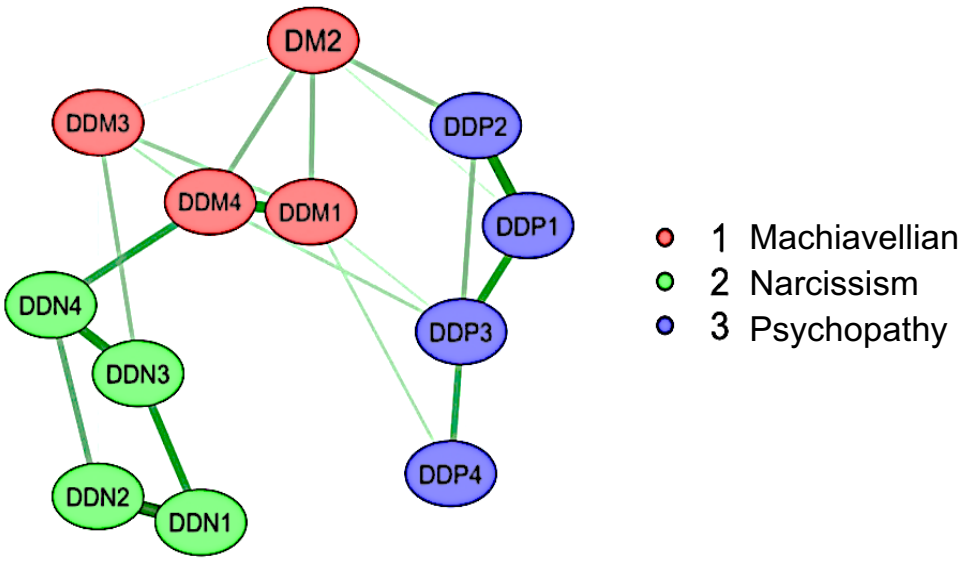
Dirty Dozen

Males and Females 18-25

For the male sample, the structure of the DD fit the data, whereby the four items from each construct loaded onto their respective clusters. For the female sample, the two narcissism items which loaded with the Machiavellian cluster, N3 (I tend to seek prestige or status) and N4 (I tend to expect special favours from others), mediated the divergence from the original structure (Figure 4).

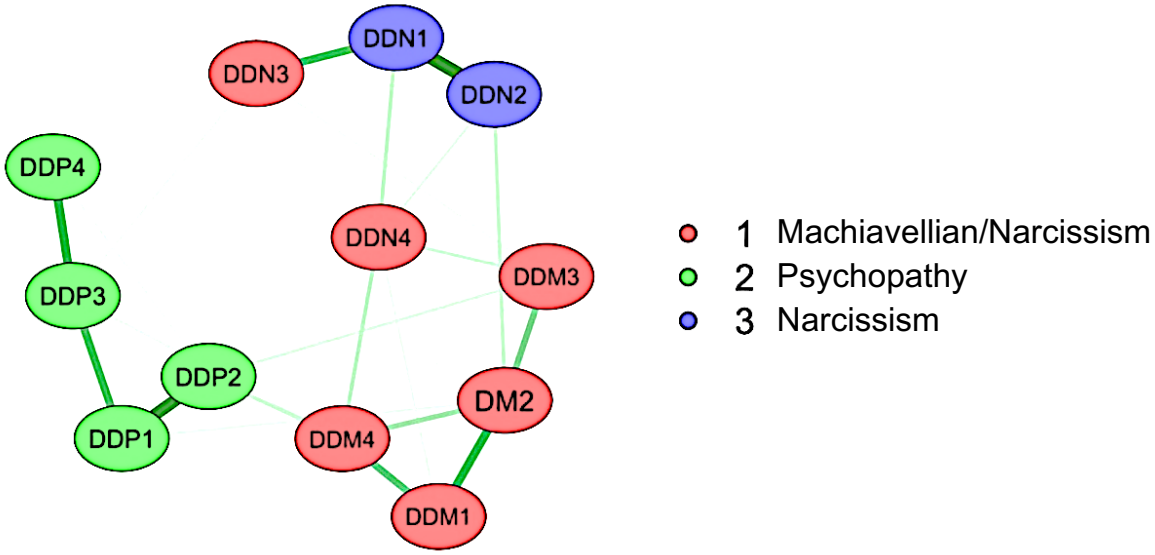
Figure 4

EGA Dirty Dozen Males 18-25



Note: DDN= Dirty Dozen Narcissism; DDP = Dirty Dozen Psychopathy; DDM= Dirty Dozen Machiavellian

EGA Dirty Dozen Females 18-25



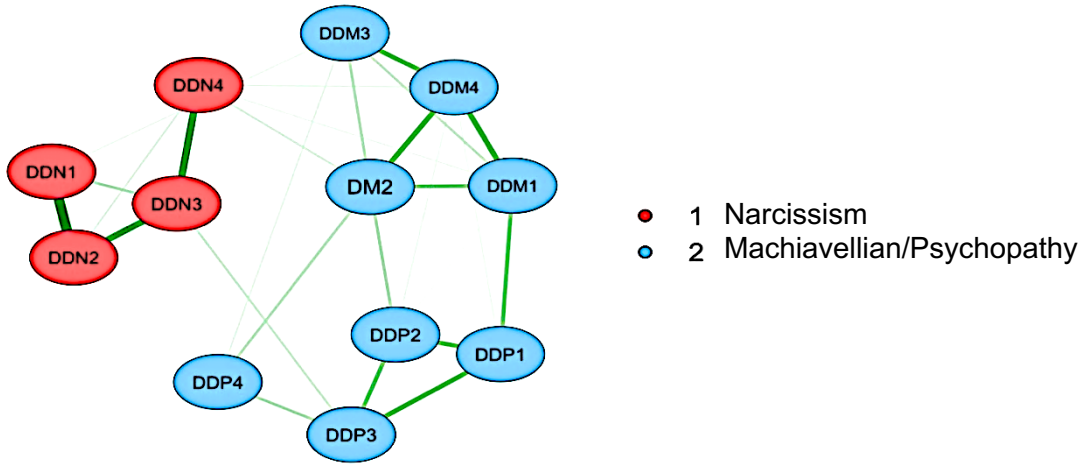
Note: DDN= Dirty Dozen Narcissism; DDP = Dirty Dozen Psychopathy; DDM= Dirty Dozen Machiavellian

Males and Females 26 and Over

In the older age groups, males displayed a structure which aligned with a dark dyad of Machiavellianism/psychopathy with narcissism forming its own cluster. The deviation from this structure in the female sample was the separation of M3 (I have used flattery to get my way) and N4 (I tend to expect special favours from others; Figure 5), which formed a dimension.

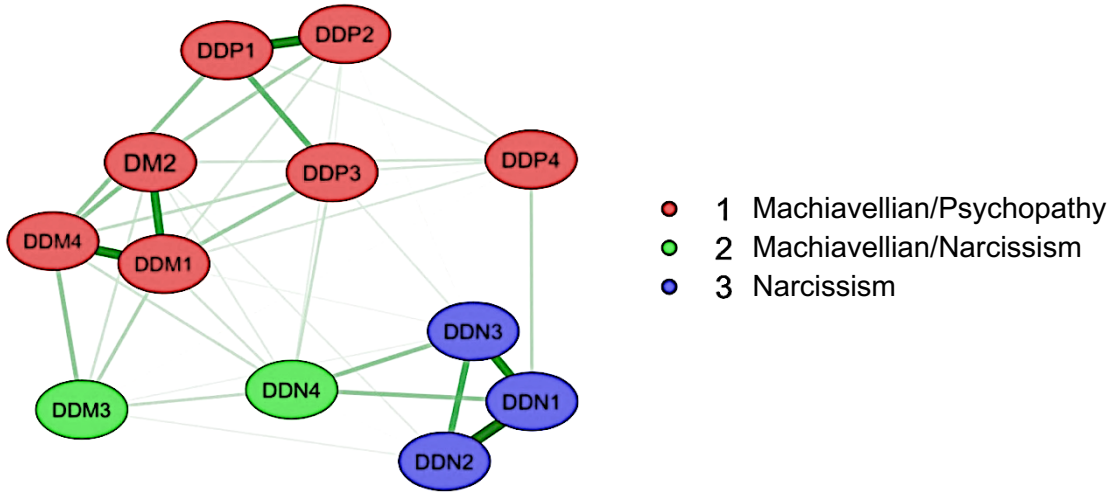
Figure 5

EGA Dirty Dozen Males 26 and Over



Note: DDN= Dirty Dozen Narcissism; DDP = Dirty Dozen Psychopathy; DDM= Dirty Dozen Machiavellian

EGA Dirty Dozen Females 26 and Over



Note: DDN= Dirty Dozen Narcissism; DDP = Dirty Dozen Psychopathy; DDM= Dirty Dozen Machiavellian

Short Dark Triad

Males and Females 18-25

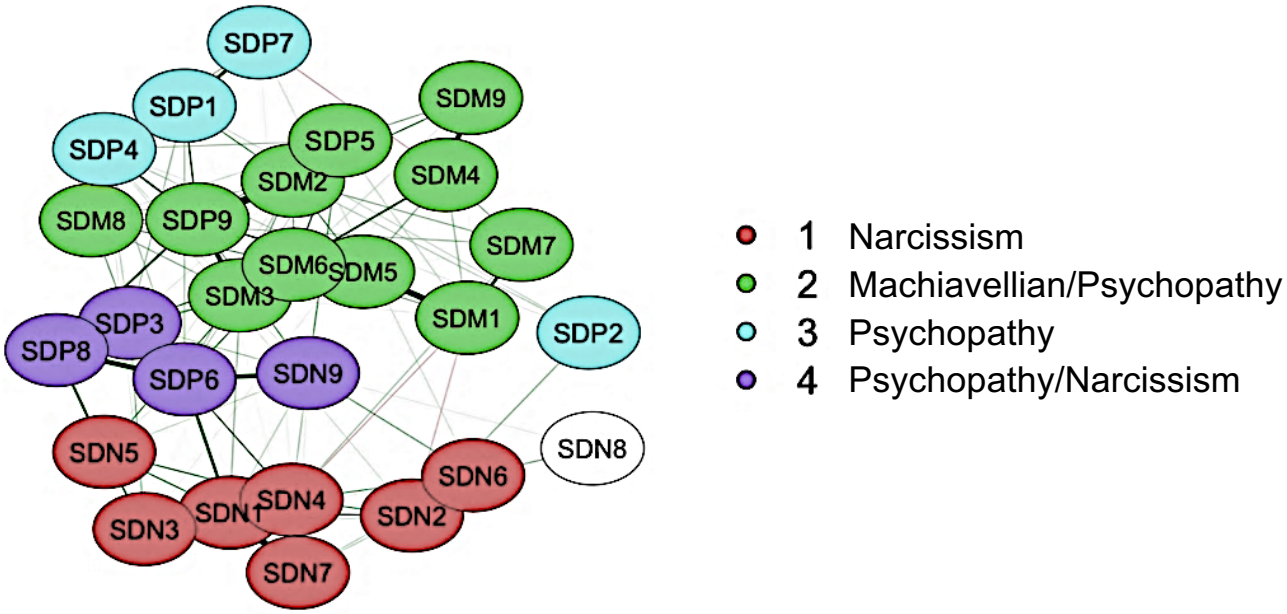
The SD3, developed as a three-factor measure, was not replicated for the younger samples. For the male group, the psychopathy indicators P5 (It's true that I can be mean to others) and P9 (I'll say anything to get what I want), clustered with the Machiavellian indicators. The remaining psychopathy items split between two dimensions appear to divide the antisocial behaviours in cluster 3 from the interpersonal antagonistic attitudes in cluster four. The female structure displayed mainly a dark dyad with narcissism. The separate Machiavellian items on cluster three, M1 (It's not wise to tell your secrets), M4 (Avoid direct conflict with others because they may be useful in the future) and M7 (There are things you should hide from other people to preserve your reputation), related to a strategic-calculating orientation. The white nodes represent indicators that did not load (Figure 6).

Males and Females 26 and Over

For the male 26 and over group, Machiavellian indicators were a problem, where they were split over three clusters. Items M1, M4, M7, formed a cluster as they did for the female 18-25 sample. The remaining indicators mainly loaded on clusters congruent with the construct. For the female sample, the psychopathy indicators were spread over three clusters. As with the younger males, item N9 (I insist on getting the respect I deserve), loaded with the psychopathy indicators. The psychopathy items (P2, P7 and P8) which were found on cluster 1 with the narcissism indicators, did not make conceptual sense (Figure 7).

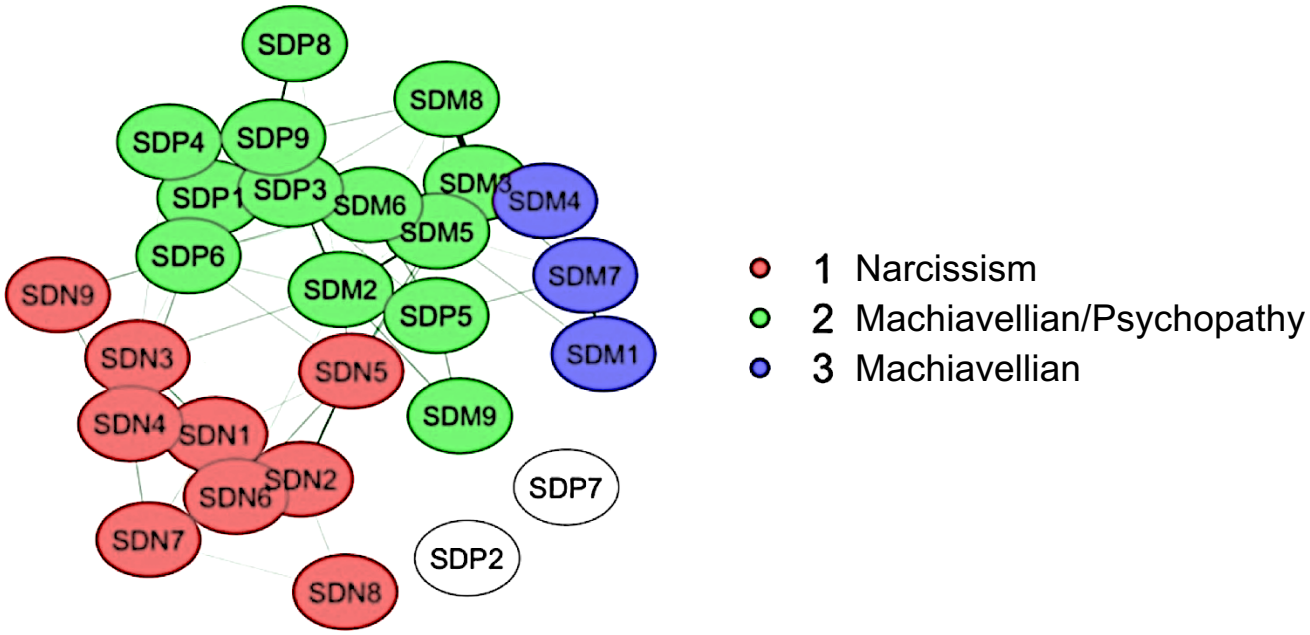
Figure 6

EGA Short Dark Triad Males 18-25



Note: SDN= Short Dark Triad Narcissism; SDP = Short Dark Triad Psychopathy; SDM = Short Dark Triad Machiavellian

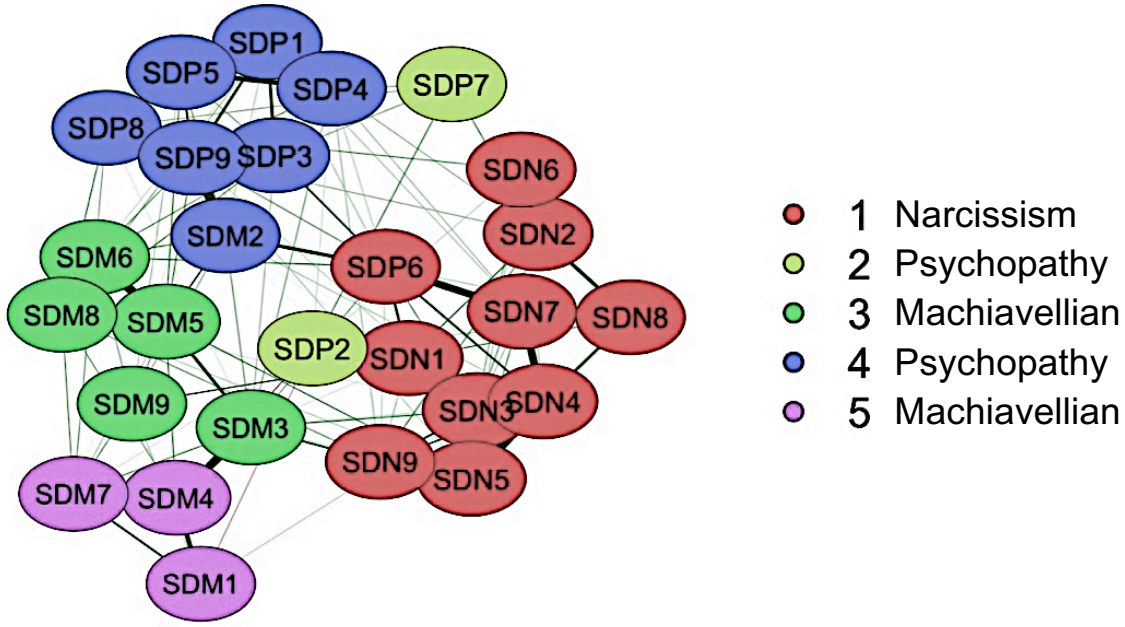
EGA Short Dark Triad Females 18-25



Note: SDN= Short Dark Triad Narcissism; SDP = Short Dark Triad Psychopathy; SDM = Short Dark Triad Machiavellian

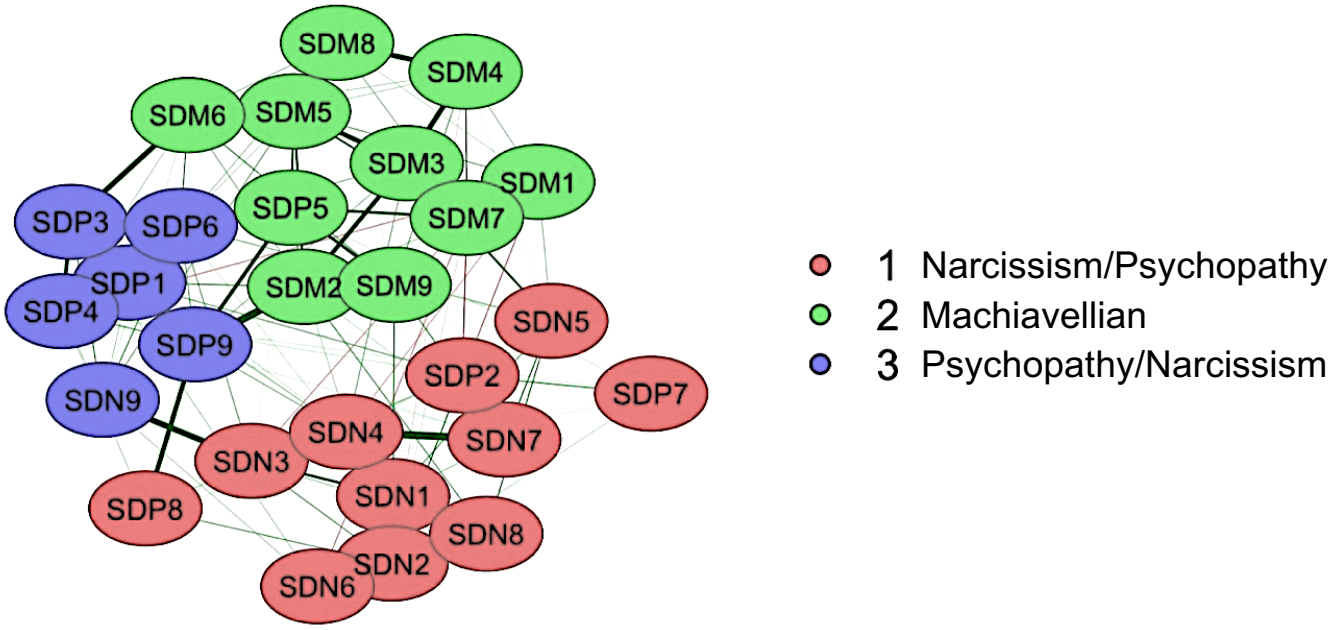
Figure 7

EGA Short Dark Triad Males 26 and Over



Note: SDN= Short Dark Triad Narcissism; SDP = Short Dark Triad Psychopathy; SDM = Short Dark Triad Machiavellian

EGA Short Dark Triad Females 26 and Over



Note: SDN= Short Dark Triad Narcissism; SDP = Short Dark Triad Psychopathy; SDM = Short Dark Triad Machiavellian

Levenson Self-Report Psychopathy Scale

Males and Females 18-25

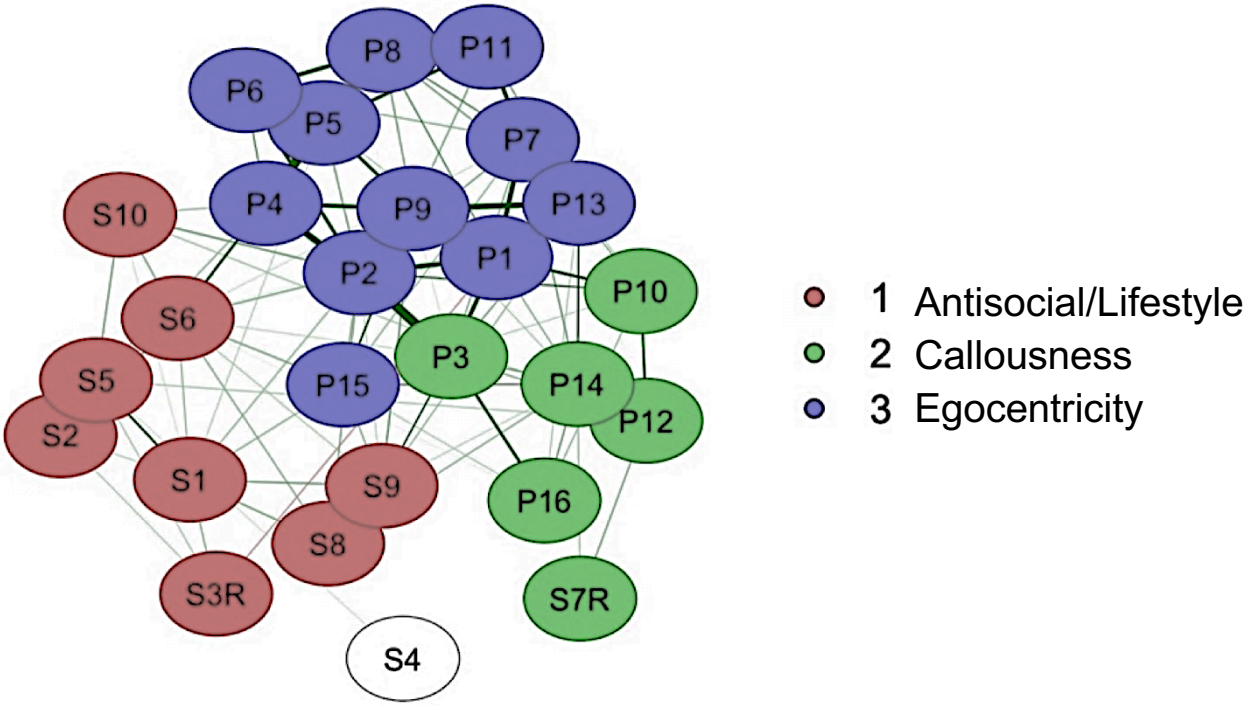
The main noticeable deviation from the original two-factor structure for the male group was the separation of the primary indicators. However, this structure aligned with the three-factor model put forth by Christian and Sellbom (2016). Cluster one addressed the antisocial factor, cluster two represented the callous facet whilst cluster three tapped into egocentricity. The female group evidenced a four-cluster model. This represented: cluster 1: negative affectivity; cluster 2: egocentricity; cluster 3; lack of planfulness; cluster 4; callousness (Figure 8).

Males and Females 26 and Over

The males 26 and over displayed a three-cluster model akin to the younger male group. Cluster one represented the antisocial facet, cluster two tapped into egocentricity and cluster three addressed callousness (Christian & Sellbom, 2016). The female group, although showing a four-cluster model as did the females 18-25, the primary indicators on cluster two addressed both callousness and egocentricity, whilst those on cluster three tapped into egocentricity. The first cluster pertained to lack of planfulness and the fourth cluster to negative affectivity (Figure 9).

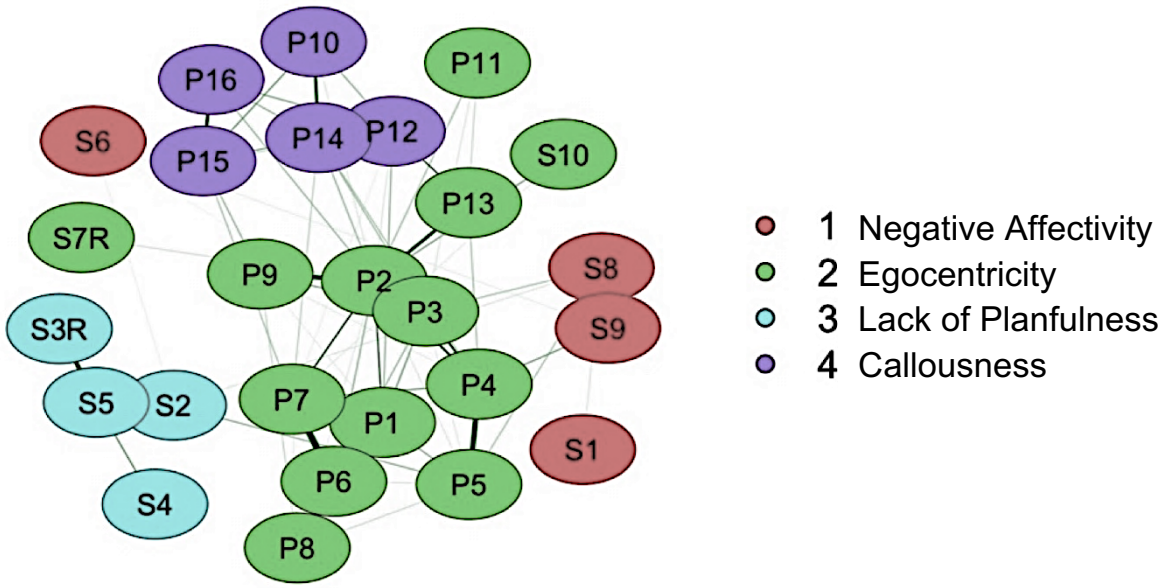
Figure 8

EGA Levenson Self-Report Psychopathy Scale Males 18-25



Note: P = Primary Psychopathy; S = Secondary Psychopathy

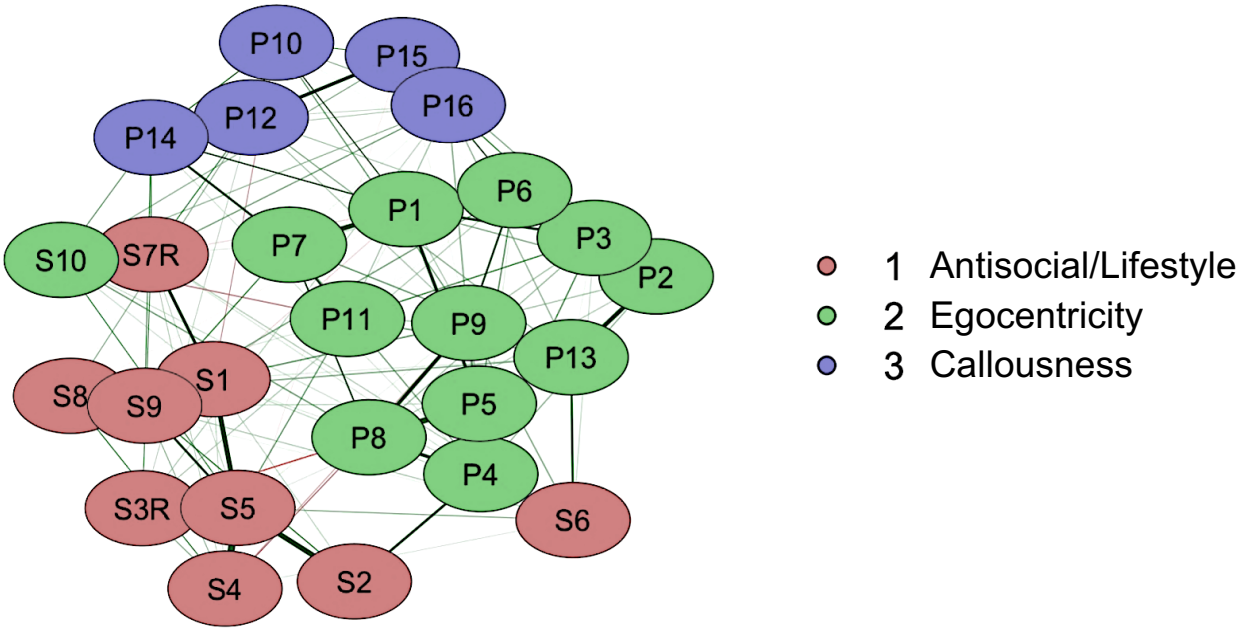
EGA Levenson Self-Report Psychopathy Scale Females 18-25



Note: P = Primary Psychopathy; S = Secondary Psychopathy

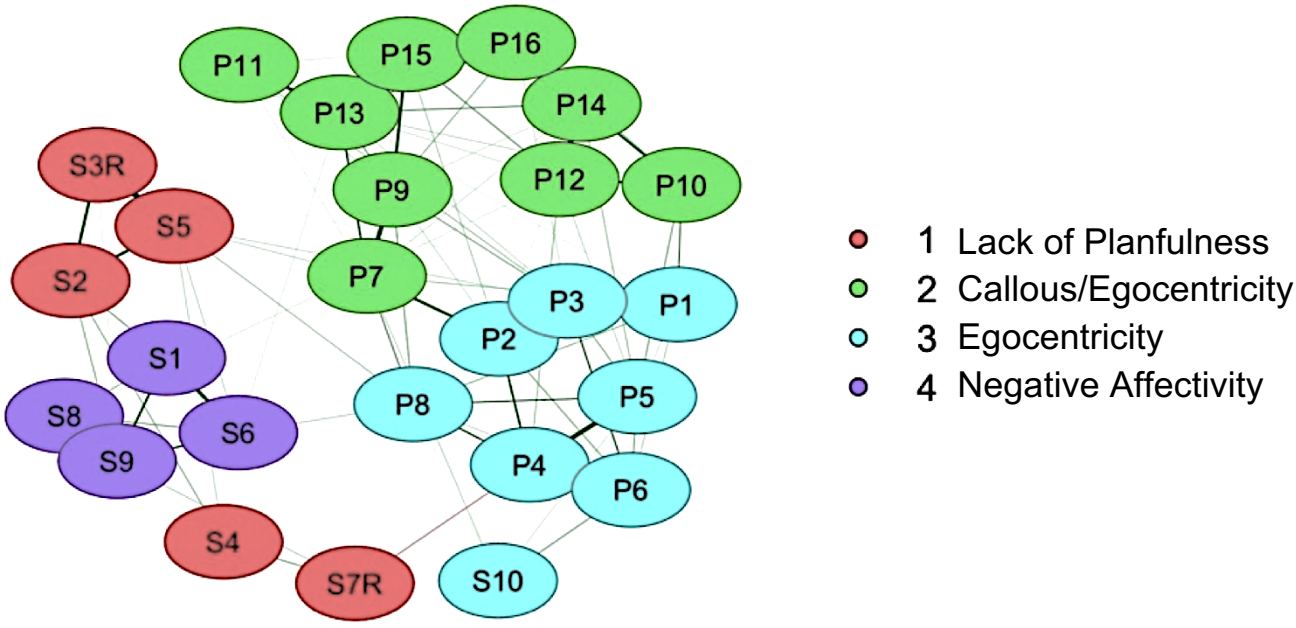
Figure 9

EGA Levenson Self-Report Psychopathy Scale Males 26 and Over



Note: P = Primary Psychopathy; S = Secondary Psychopathy

EGA Levenson Self-Report Psychopathy Scale Females 26 and Over



Note: P = Primary Psychopathy; S = Secondary Psychopathy

Mach IV

Males and Females 18-25

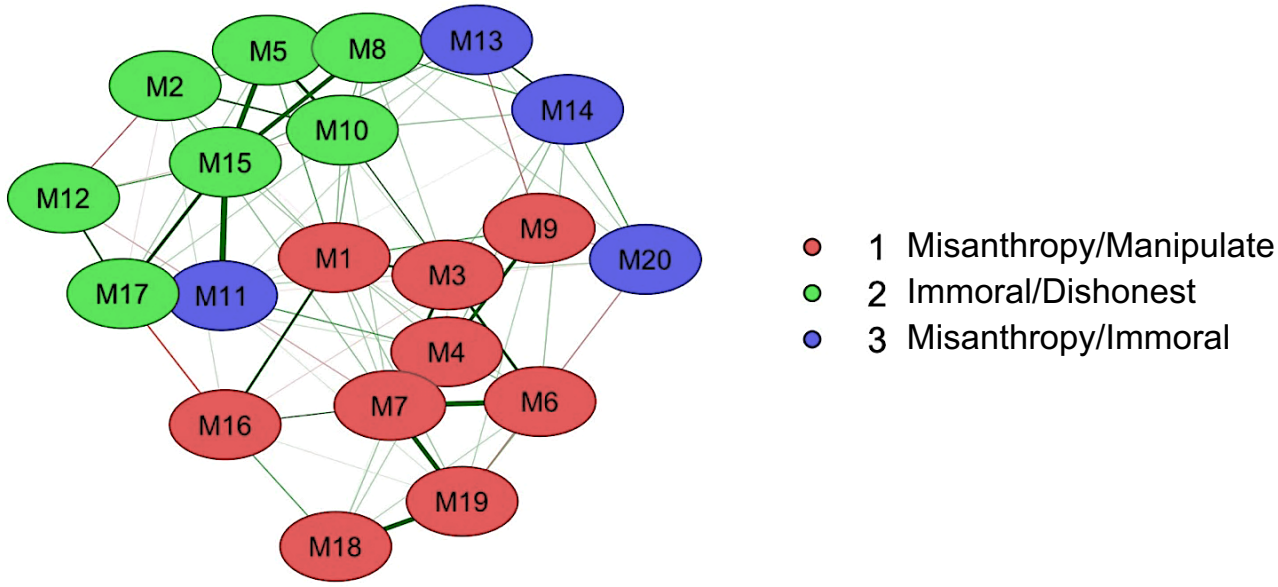
Although the Mach IV was developed with a three-factor structure, the indicators within the three clusters for the males 18-25 group did not align with those of the measure. Focusing on the item content, the male sample showed that cluster one mainly held items pertaining to cynical misanthropy and manipulation, cluster two tapped mostly into immorality and dishonesty, whilst cluster three contained items also addressing misanthropy and immorality. For the females 18-25, cluster one predominantly had items relating to misanthropy and cynicism. Two of the three indicators on cluster two tapped into immorality, whilst most of the items within cluster three addressed immorality. The two items on cluster four pertained to manipulation. Thus, both the structures and the items loading on clusters varied between males and females (Figure 10).

Males and Females 26 and Over

The first cluster for the male sample mainly addressed misanthropy and cynicism whilst the second cluster tapped into immorality and dishonesty. The two items on the third dimension held one item pertaining to manipulation and the other immorality. The female group showed a unidimensional structure (Figure 11).

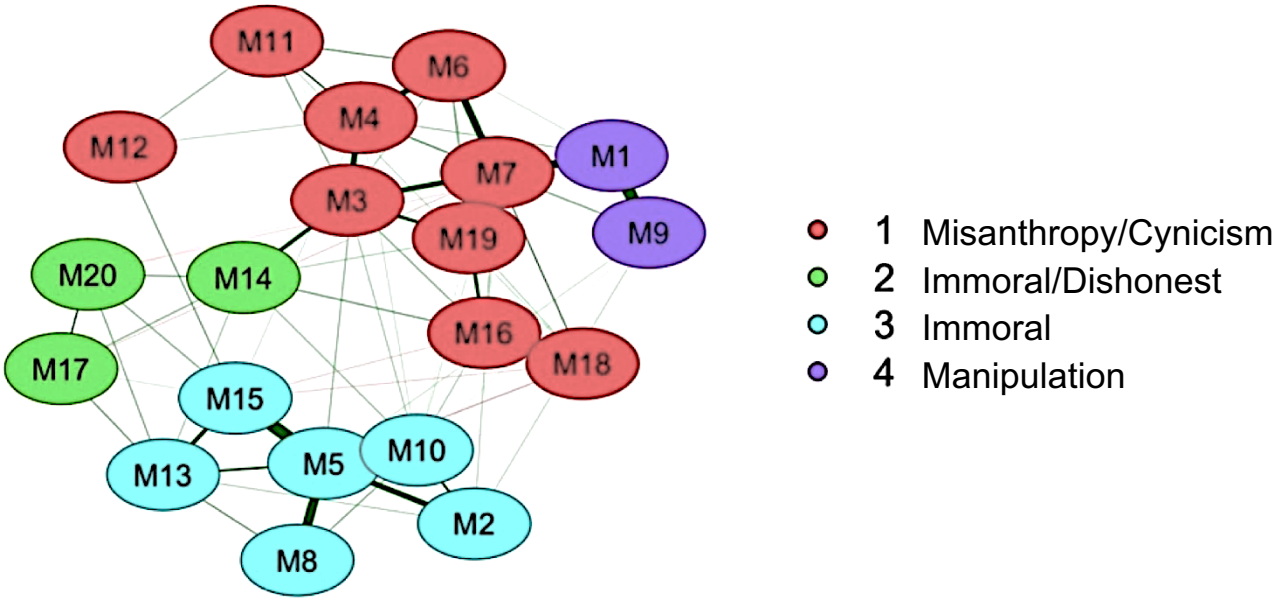
Figure 10

EGA Mach IV Males 18-25



Note: M = Machiavellianism

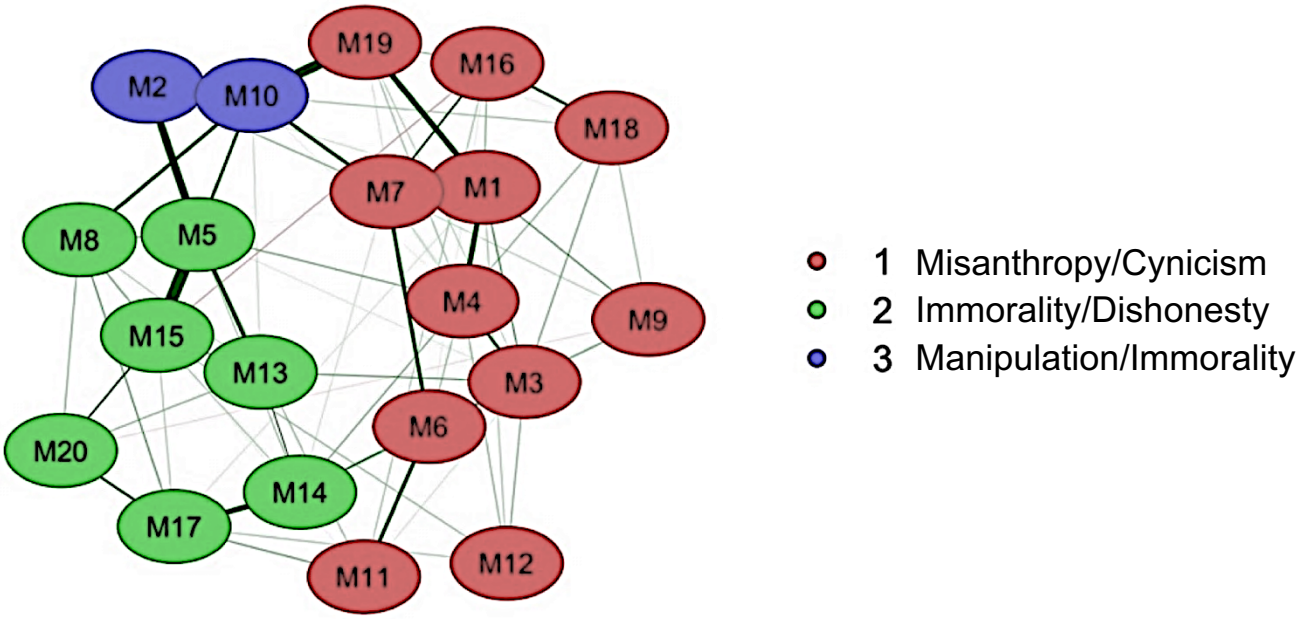
EGA Mach IV Females 18-25



Note: M = Machiavellianism

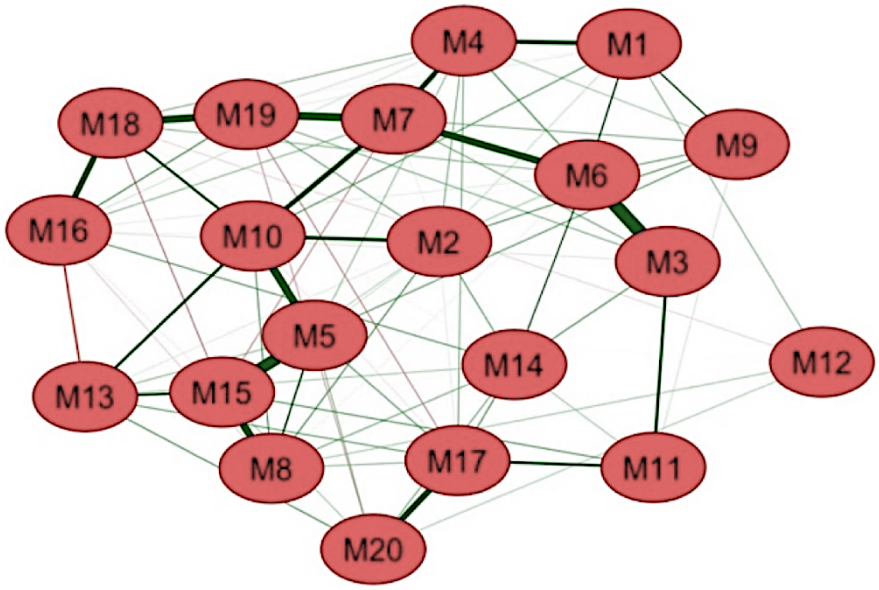
Figure 11

EGA Mach IV Males 26 and Over



Note: M = Machiavellianism

EGA Mach IV Females 26 and Over



Note: M = Machiavellianism

Consequently, across groups, excluding females 26 and over, although item loadings and cluster structures differed, the traits that emerged from the Mach IV are cynical misanthropy, immorality and manipulation.

Narcissistic Personality Inventory

Males and Females 18-25

The eight clusters in the male sample did not clearly align with the facets of the NPI. The first dimension mainly addressed authority and entitlement, cluster two contained items pertaining to authority, exploitativeness and self-sufficiency, whilst the third cluster held indicators tapping into exhibitionism, exploitativeness, self-sufficiency and authority. Dimension four aligned with superiority and authority, cluster five held items relating to exhibitionism and self-sufficiency and the sixth cluster was clearly, vanity. The two items in cluster seven were from the superiority facet, whilst indicators in cluster eight pertained to exhibitionism (Figure 12).

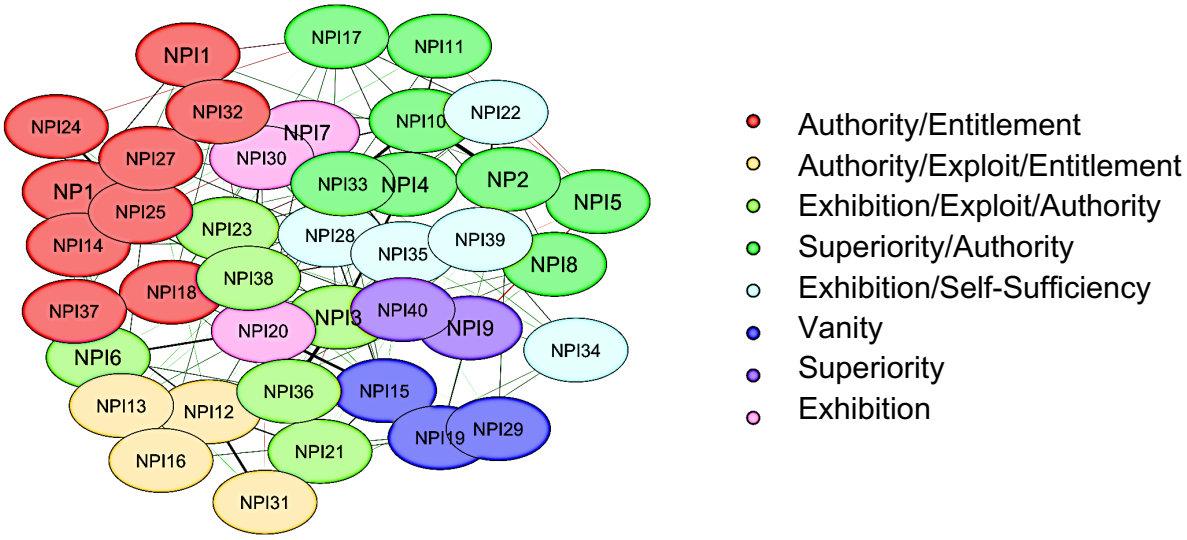
The female group showed that cluster one items mainly tapped into authority, whilst with the second dimension the prevalence was exploitativeness and entitlement. Vanity was a clear focus for the third cluster, whilst exhibitionism and exploitativeness were mostly found in the fifth dimension. The remaining two clusters, four and six, held items across the differing subfactors of the construct, with no definitive focus (Figure 12).

Males and Females 26 and Over

The male sample showed cluster one addressing superiority and vanity, cluster two mainly aligned with exhibitionism, whilst cluster three aligned with both exploitativeness and exhibitionism. The fourth dimension held items mainly tapping into authority and the two items in cluster five related to self-sufficiency. The female group displayed a divergent structure, with the first cluster pertaining to authority and self-sufficiency. The second dimension addressed exploitativeness and entitlement, whilst cluster three contained indicators from across the NPI facets. The three items in the fourth cluster related to self-sufficiency and cluster five mainly tapped into authority. The sixth and final dimension held items of superiority and exhibitionism (Figure 13).

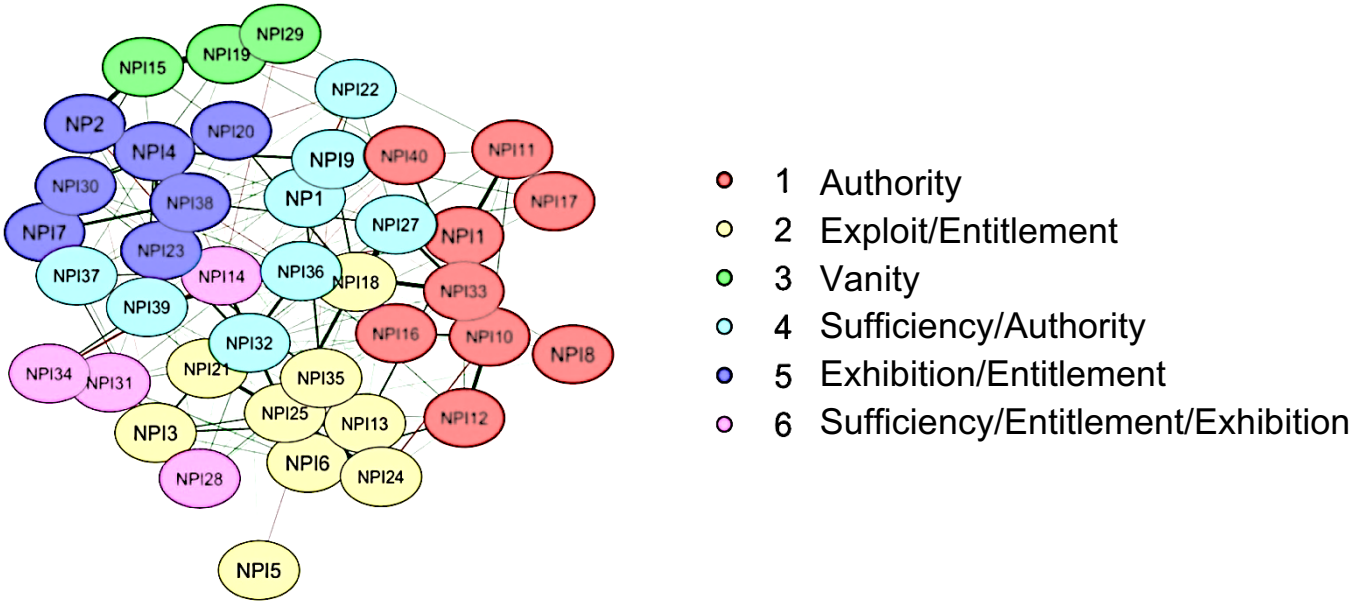
Figure 12

EGA Narcissistic Personality Inventory Males 18-25



Note: NP = NPI Narcissism

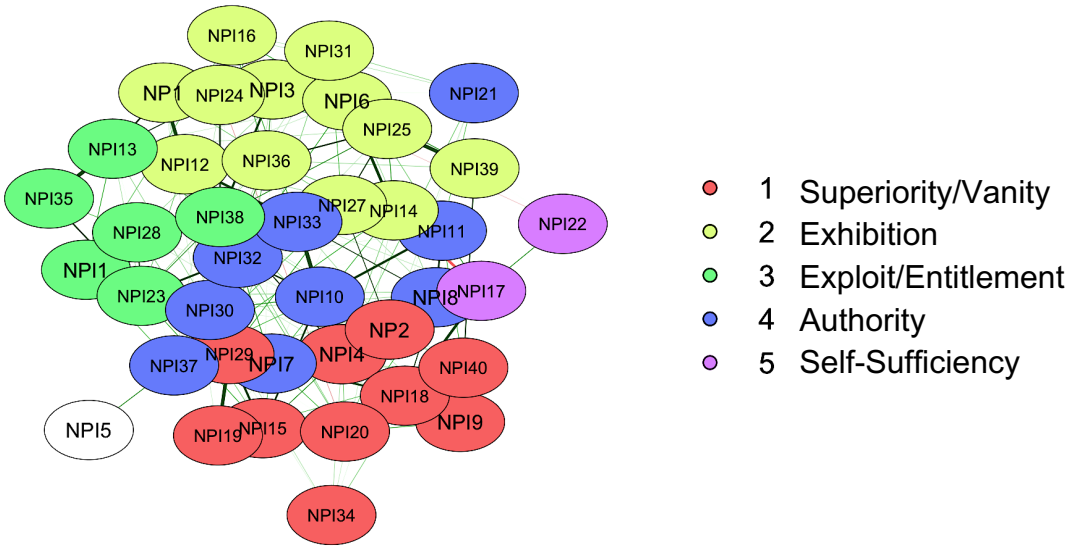
EGA Narcissistic Personality Inventory Females 18-25



Note: NP = NPI Narcissism

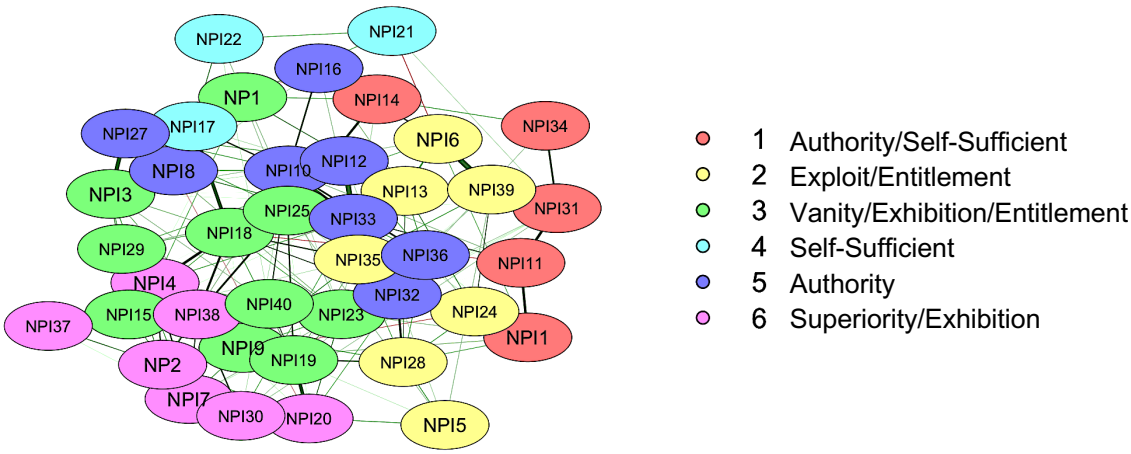
Figure 13

EGA Narcissistic Personality Inventory Males 26 and Over



Note: NP = NPI Narcissism

EGA Narcissistic Personality Inventory Females 26 and Over



Note: NP = NPI Narcissism

Even though there have been various factor-structures suggested for the NPI, across groups, none of the models across groups aligned with them (Ackerman, Donnellan & Robins, 2012; Corry et al., 2008; Emmons, 1987; Kubarych et al., 2004; Raskin & Terry, 1988).

Varieties of Sadistic Tendencies

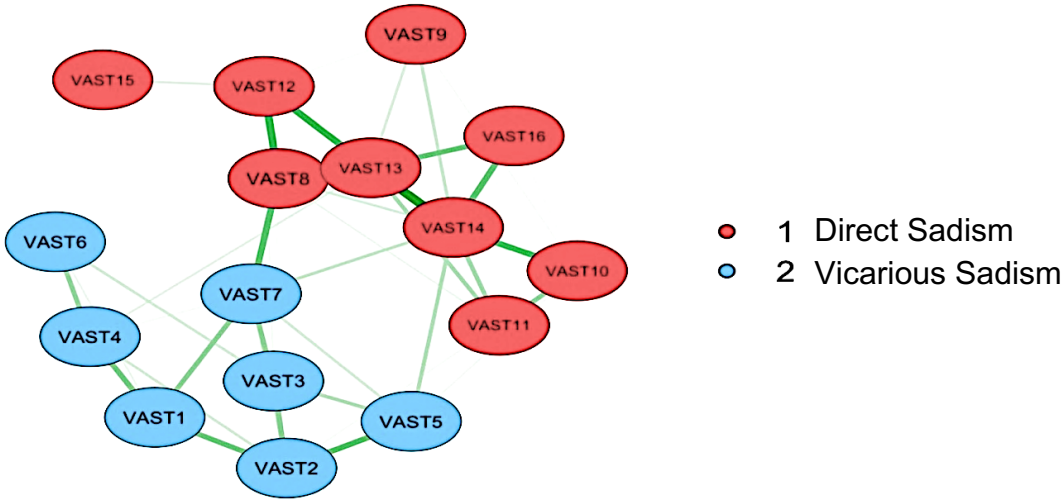
For a two-factor model vicarious sadism is defined by items V1 -V7 and direct sadism by items V8 – V16 (Figure 14).

Males and Females 18-25

As can be seen, the two factors in the male sample aligned with the two-factor model, whilst for females there was no distinction between vicarious and direct sadism

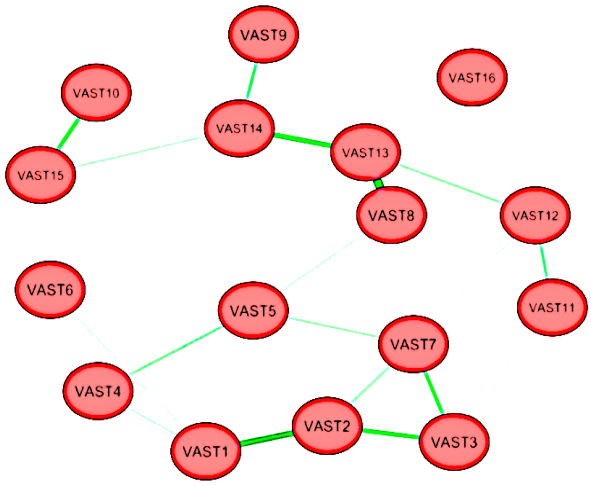
Figure 14

EGA Varieties of Sadistic Tendencies Males 18-25



Note: VAST = Varieties of Sadistic Tendencies

EGA Varieties of Sadistic Tendencies Females 18-25



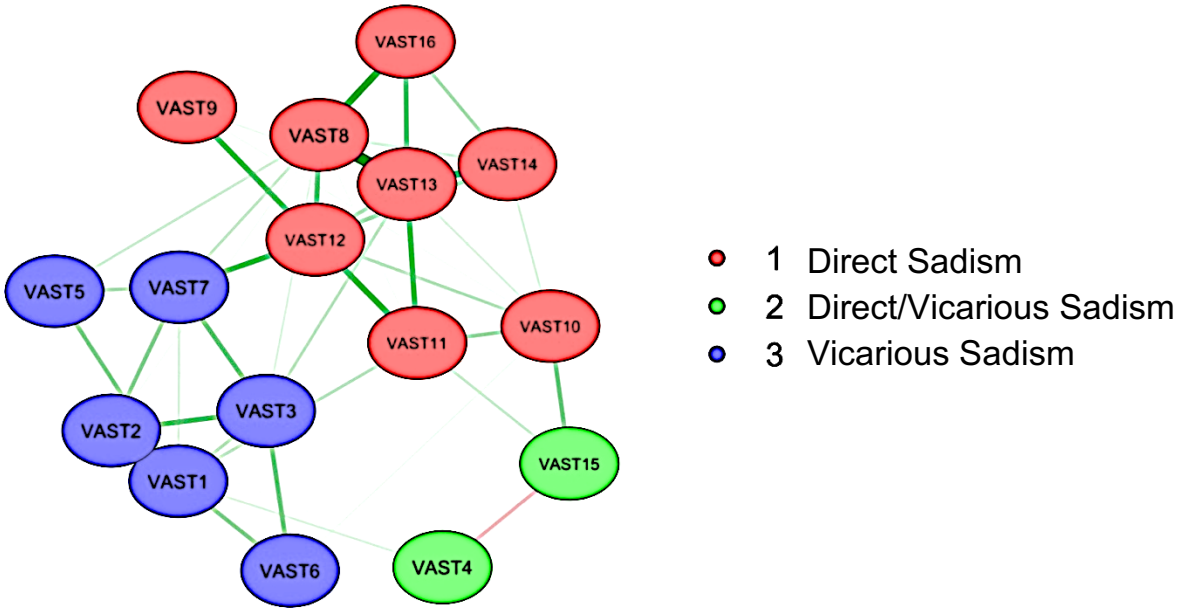
Note: VAST = Varieties of Sadistic Tendencies

Males and Females 26 and Over

For the male sample, two indicators, which are both reverse scored (V4; I sometimes look away in horror movies; V15; I never said mean things to my parents), mediated the deviation from the two-factor structure. As can be seen the female sample as with the younger female group was unidimensional (Figure 15).

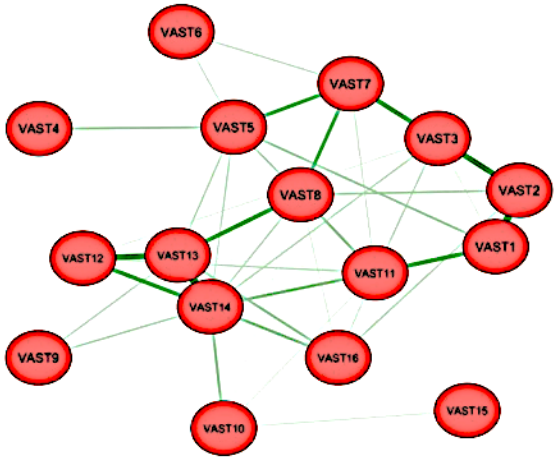
Figure 15

EGA Varieties of Sadistic Tendencies Males 26 and Over



Note: VAST = Varieties of Sadistic Tendencies

EGA Varieties of Sadistic Tendencies Females 26 and Over



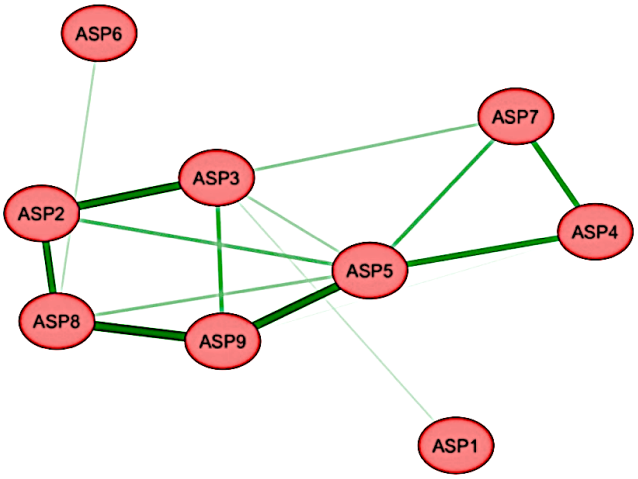
Note: VAST = Varieties of Sadistic Tendencies

Assessment of Sadistic Personality

For both males and females 18-25 and males and females 26 and over, the EGA aligned with the original model (Figures 16 and 17).

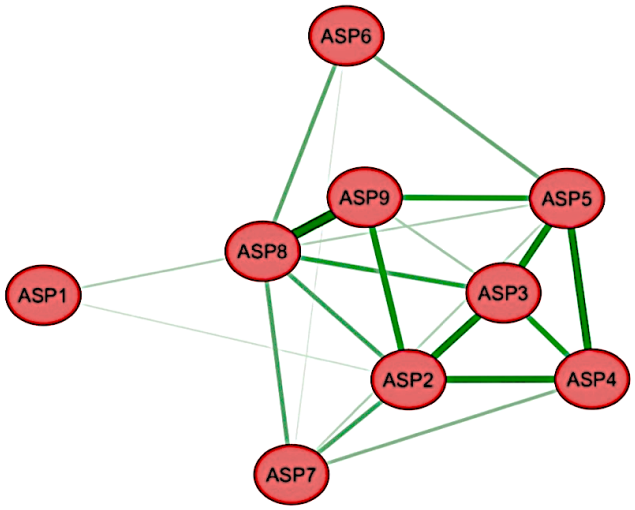
Figure 16

EGA Assessment of Sadistic Personality Males 18-25



Note: ASP = Assessment of Sadistic Personality

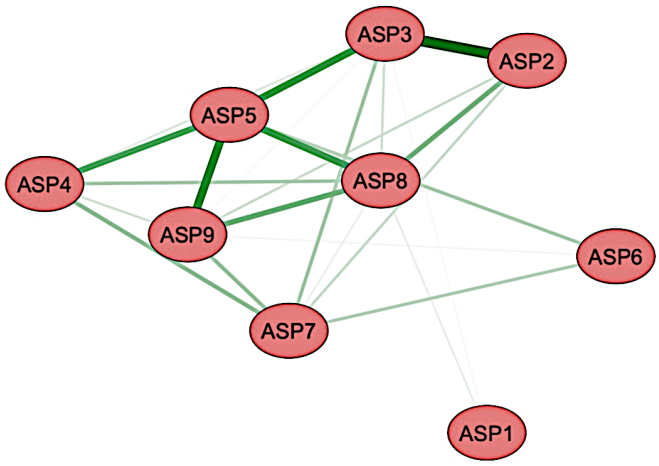
EGA Assessment of Sadistic Personality Females 18-25



Note: ASP = Assessment of Sadistic Personality

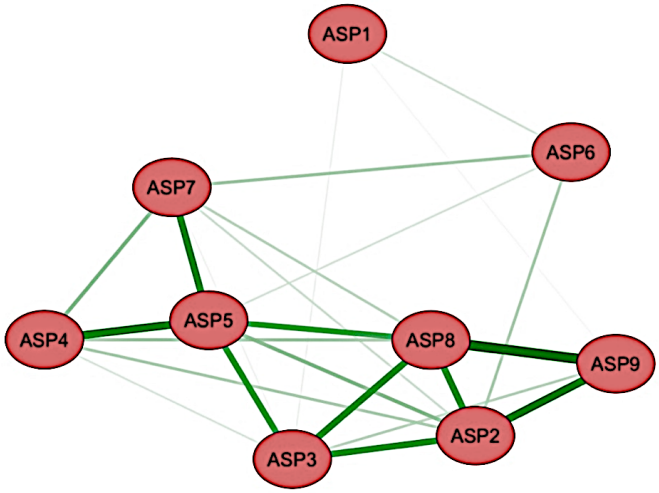
Figure 17

EGA Assessment of Sadistic Personality Males 26 and Over



Note: ASP = Assessment of Sadistic Personality

EGA Assessment of Sadistic Personality Females 26 and Over



Note: ASP = Assessment of Sadistic Personality

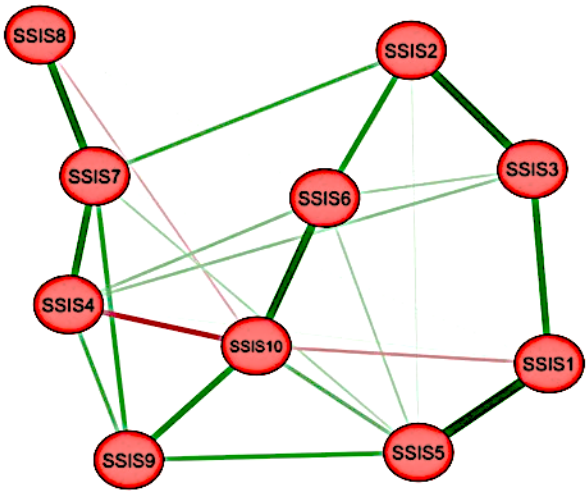
Short Sadistic Impulse Scale

Males and Females 18-25

Both groups were congruent with the original model (Figure 18)

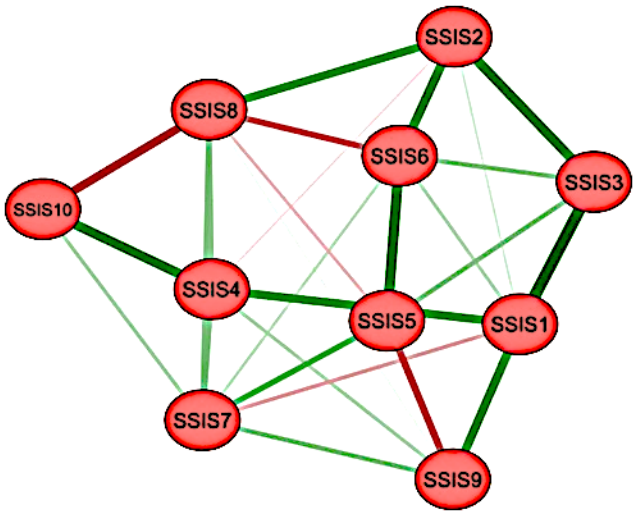
Figure 18

EGA Short Sadistic Impulse Scale Males 18-25



Note: SSIS = Short Sadistic Impulse Scale

EGA Short Sadistic Impulse Scale Females 18-25



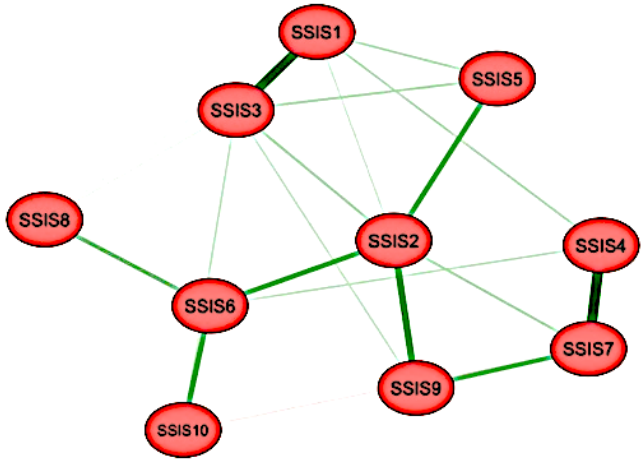
Note: SSIS = Short Sadistic Impulse Scale

Males and Females 26 and Over

It is interesting to note that although there were problems with the CFA for the females over 26 group, whereby the model was not identified. The EGA was an excellent fit to the structure of the measure, aligning with the original model (Figure 19).

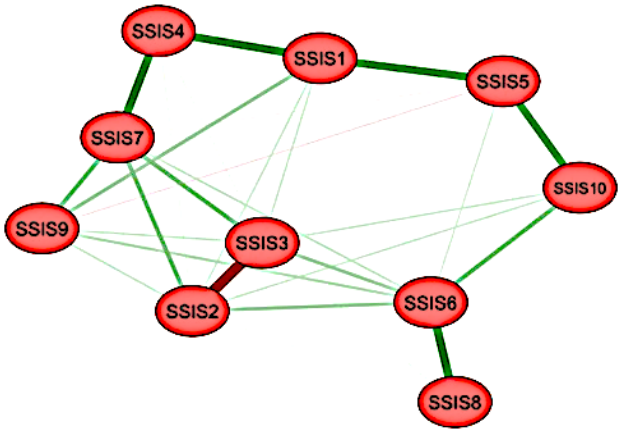
Figure 19

EGA Short Sadistic Impulse Scale Males 26 and Over



Note: SSIS = Short Sadistic Impulse Scale

EGA Short Sadistic Impulse Scale Females 26 and Over



Note: SSIS = Short Sadistic Impulse Scale

Across measures and groups, excluding the ASP and SSIS, heterogeneity was shown. This suggested that the scales may not be measuring the same construct

across groups, whereby qualitative differences in how males and females respond to indicators, affect the structures. As such, Mokken Scale Analysis (MSA; Appendix 3.3.4), was conducted to investigate the scales from the bottom-up (Van Schuur, 2003). As well as Mokken scales being calculated, item difficulty is assessed through the mean scores for each indicator, such that a higher threshold of the construct is needed for the participant to endorse a more difficult item. For example, the item 'I am often bored', may be easier to endorse, requiring a lower level of the construct than the item, 'I often admire a really clever scam' (Shenkin et al., 2014; Watson, Deary & Shipley, 2008; Table 4 and 5). A discussion concerning the method can be found in Chapter Three.

Mokken Scale Analysis

Levenson Self-Report Psychopathy Scale

Females and males 18-25

The automated item selection algorithm (AISP), initially returned four scales for females and six for males. However, in both samples, one scale only had two items clustered, which were discarded. The female sample evidenced seven unscalable items (P5, P11, S1, S4, S6, S7 and S10), whereas nine items were unscalable for the males (P10, P15, P16, S1, S3, S4, S6, S7 and S10). Invariant item ordering (IIO) in the female sample found four items (P2, P3, P7, and P13), with criterion (crit) values outside accepted ranges. Item P13, with the highest value (130), was removed, resulting in all scales displaying no significant violations and all crit values at zero. For the male sample, IIO found item P6 had a crit value of 211 and significant violations, removal of this item left all scales with crit values at zero and no significant violations (Table 4).

Table 4

H-coefficients for Levenson Self-Report Psychopathy Scale Females and Males 18-25

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
Scale 1			Scale 1		
P8. Myself	0.48 (0.06)	2.57	P5. Money	0.51 (0.06)	2.33
P9. Tell others	0.45 (0.06)	2.13	P1. Success	0.52 (0.05)	2.31
P1. Success	0.48 (0.06)	1.88	P9. Tell others	0.49 (0.05)	2.13
P6. Bottom line	0.46 (0.06)	1.85	P3. Justified	0.51 (0.05)	2.06
P4. Goodies	0.47 (0.05)	1.82	P4. Goodies	0.56 (0.05)	2.05
P2. Get away with	0.61 (0.04)	1.74	P2. Get away with	0.62 (0.04)	1.97
P3. Justified	0.56 (0.04)	1.72			
P7. Ripped off	0.47 (0.06)	1.62			
Scale H = 0.49 (0.05)			Scale H = 0.54 (0.04)		
Scale H^T= 0.29			Scale H^T= 0.07		
α = 0.86			α = 0.85		
Scale 2					
P10. Upset	0.43 (0.07)	2.02			
P15. Lie	0.45 (0.07)	2.00			
P12. Hurt others	0.49 (0.07)	1.67			
P16. Cheating	0.44 (0.07)	1.61			
P14. Feel bad	0.44 (0.07)	1.55			
Scale H = 0.46 (0.06)					
Scale H^T= 0.18					
α = 0.76					
Scale 3					
S2. Bored	0.48 (0.07)	2.94			
S5. Interest	0.58 (0.05)	2.51			
S3. Pursue	0.43 (0.07)	2.10			
Scale H = 0.50 (0.06)					
Scale H^T= 0.47					
α = 0.69					

Note: Minimum loading used =.40. Scale H should exceed .30. H_T = coefficient H of items

The Mokken scales showed diverse loadings for males and females. For the female group scale one represented egocentricity, scale two, callousness and scale three tapped into lack of planfulness. For the males, the single Mokken scale addressed egocentricity. The females found the item S2 (I am often bored), the easiest to endorse and P14 (I feel bad if my words or actions cause someone else to feel emotional pain), the most difficult, whilst for the males P5 (Making a lot of money is my most important goal), was the easiest, with P2 (For me what's right is whatever I can get away with), the more difficult item to endorse.

Females and males over 26

Five scales were returned by the AISP for females and four for males. In the male sample, one scale only had two items clustered and was discarded. The female sample evidenced four unscalable items (P11, S4, S7 and S10), whilst there were nine for the males (P7, P10, P11, P14R, S2, S3, S4, S6 and S10). IIO showed on scale four of the female sample that two items (S9 and S1), held crit values of 80 and 85. Removal of item S9, resulted in no significant violations and all scales with crit values of zero. For the males, scale two, which contained three items, displayed crit values of 114 and 105 for items P12 and P15 respectively. Removing one of these items would have resulted in a two-item scale, therefore this scale was discarded. The remaining scales returned no significant violations with all crit values at zero (Table 5).

Table 5

H-coefficients for Levenson Self-Report Psychopathy Scale Females and Males 26 and over

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
Scale 1			Scale 1		
P8. Myself	0.41 (0.06)	2.14	P8. Myself	0.49 (0.05)	2.27
P1. Success	0.46 (0.05)	1.80	P9. Tell others	0.52 (0.05)	1.95
P5. Money	0.44 (0.05)	1.75	P5. Money	0.49 (0.05)	1.92
P6. Bottom line	0.45 (0.06)	1.65	P6. Bottom Line	0.43 (0.05)	1.83
P4. Goodies	0.48 (0.06)	1.50	P1. Success	0.47 (0.05)	1.81
P3. Justified	0.55 (0.05)	1.48	P4. Goodies	0.51 (0.05)	1.73
P2. Get away	0.48 (0.05)	1.40	P3. Justified	0.53 (0.04)	1.71
			P2. Get away	0.54 (0.05)	1.61
			P13.Manipulating	0.42 (0.06)	1.40
Scale *H* = 0.46 (0.04)			Scale *H* = 0.49 (0.04)		
Scale H^T= 0.25			Scale H^T= 0.21		
α = 0.82			α = 0.87		
Scale 2			Scale 2		
P9. Tell others	0.52 (0.05)	1.67	S5. Lose interest	0.41 (0.08)	2.14
P15. Lie	0.44 (0.06)	1.59	S9. Frustrated	0.51 (0.06)	1.81
P7. Stupid	0.47 (0.06)	1.55	S8. Shouting	0.47 (0.07)	1.73
P13.Manipulating	0.51 (0.06)	1.33			
Scale *H* = 0.49 (0.05)			Scale *H* = 0.46 (0.06)		
Scale H^T= 0.13			Scale H^T=0.17		
α = 0.74			α = 0.68		
Scale 3					
P10. Upset	0.52 (0.08)	1.77			
P12. Hurt others	0.48 (0.08)	1.48			
P16. Cheating	0.48 (0.09)	1.42			
P14. Feel bad	0.58 (0.07)	1.36			
Scale *H* = 0.51 (0.07)					
Scale H^T= 0.24					
α = 0.75					
Scale 4					
S6. Don't understand	0.41 (0.06)	1.96			
S1. Trouble	0.40 (0.06)	1.85			
S8. Shouting	0.39 (0.07)	1.75			
Scale *H* = 0.40 (0.06)					
Scale H^T= 0.30					
α =0.63					
Scale 5					
S2. Bored	0.44 (0.07)	2.49			
S5. Lose interest	0.45 (0.06)	2.12			
S3. Pursue	0.45 (0.06)	2.07			
Scale *H* = 0.44 (0.06)					
Scale H^T= 0.18					
α =0.66					

Note: Minimum loading used =.40. Scale *H* should exceed .30. H_T = coefficient *H* of items.

The Mokken structure for the female sample showed scale one and two addressing egocentricity, scale three, callousness, scale four negative emotionality and scale five, a lack of planfulness. However, for the male sample, scale one tapped into egocentricity whilst scale two pertained to negative affectivity. As with the younger group, the females found the item S2 (I am often bored), the easiest to endorse, whilst for males, it was P8 (Looking out for myself is my top priority). The most difficult indicator to endorse for both groups was P13 (I enjoy manipulating other people's feelings).

Narcissism

Narcissistic Personality Inventory

Females and males 18-25.

The female sample evidenced seven scales, with four of these only having two items clustered, which were discarded. The data from the male sample elicited eight scales, however only one of these had three items clustered, with the remaining scales containing only two items, therefore, these were discarded. For the females, eighteen items were unscalable, (2, 5, 6, 8, 14, 16, 17, 18, 20, 21, 22, 28, 31, 32, 34, 37, 38 and 39), and the males displayed twenty-three unscalable items (1, 2, 3, 5, 6, 8, 11, 12, 13, 14, 16, 17, 20, 22, 23, 24, 25, 28, 31, 34, 35, 38 and 39). The scales for both samples met the assumptions of MH and IIO (Table 6).

Table 6

H-coefficients for Narcissistic Personality Inventory Females and Males 18-25

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Female			Male		
NPI			NPI		
Scale 1			Scale 1		
26. Complimented	0.49 (0.08)	0.42	29. Mirror	0.71 (0.09)	0.43
15. Show off body	0.45 (0.07)	0.32	19. Look at body	0.76 (0.07)	0.25
30. I really like.. centre	0.51 (0.06)	0.26	15. Show off body	0.75 (0.09)	0.17
7. I like to be the centre	0.44 (0.07)	0.19			
4. I am good	0.44 (0.07)	0.17			
23. Stories	0.50 (0.08)	0.15			
Scale *H* = 0.47 (0.06)			Scale *H* = 0.74 (0.07)		
Scale H^T= 0.21			Scale H^T= 0.52		
α = 0.75			α = 0.73		
Scale 2					
10. Good leader	0.49 (0.08)	0.32			
12. Authority	0.43 (0.08)	0.28			
33. Prefer to be a leader	0.55 (0.06)	0.26			
27. Power	0.43 (0.09)	0.16			
36 Born leader	0.56 (0.08)	0.13			
Scale *H* = 0.49 (0.06)					
Scale H^T= 0.17					
α = 0.75					
Scale 3					
3. Dare	0.43 (0.13)	0.13			
13. Manipulate	0.52 (0.08)	0.24			
35. Believe	0.49 (0.09)	0.26			
Scale *H* = 0.49 (0.09)					
Scale H^T= 0.15					
α = 0.64					

Note: Minimum loading used =.40. Scale *H* should exceed .30. H_T = coefficient H of items

On comparison with the item loadings of the NPI (Raskin & Terry, 1988), only scale two, which tapped into authority for the female group aligned. For the males, the one Mokken scale contained three items pertaining to vanity. In relation to item endorsement, for females 26 (I like to be complimented), was the easiest and for males it was 29 (I like to look at myself in the mirror). The female group found 36 (I am a born leader) and 3 (I would do almost anything on a dare), the more difficult items to endorse, whilst for the male group it was 15 (I like to show off my body).

Females and males over 26

The female sample evidenced five Mokken scales, although scales four and five only had two items and were discarded. The data from the male sample returned eleven scales, with six scales containing three or more items, the remaining scales were discarded. For the female sample, seventeen items were unscalable (1, 2, 5, 8, 11, 14, 16, 21, 22, 24, 27, 28, 31, 34, 35, 37 and 40), whilst for the males, ten items were unscalable (5, 8, 10, 21, 22, 24, 27, 34, 35 and 37). In the male sample, one scale containing three items had crit values outside the accepted limits, removal of one item would result in a scale with only two indicators, thus, this scale was discarded. Crit values for item 4, were also outside of the range in the female sample, resulting in removal of the indicator. All remaining scales for both samples held crit values of zero with no significant violations (Table 7).

Table 7

H-coefficients for Narcissistic Personality Inventory Females and Males 26 and over

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
Scale 1			Scale 1		
10. Good leader	0.58 (0.09)	0.44	10. Good leader	0.41 (0.09)	0.43
32. Recognise authority	0.47 (0.06)	0.25	29. Mirror	0.46 (0.08)	0.33
12. Like authority	0.49 (0.07)	0.25	19. Look at body	0.56 (0.08)	0.21
33. Be a leader	0.57 (0.06)	0.22	15. Show off body	0.72 (0.07)	0.14
36. Born leader	0.44 (0.07)	0.16			
25. Deserve	0.49 (0.08)	0.13			
38. Upset	0.49 (0.10)	0.09			
Scale H = 0.50 (0.05)			Scale H = 0.52 (0.07)		
Scale H^T= 0.31			Scale H^T= 0.31		
α = 0.78			α = 0.67		
Scale 2			Scale 2		
26. Complimented	0.56 (0.08)	0.44	26. Complimented	0.62 (0.08)	0.44
9. Special	0.44 (0.07)	0.25	4. Good	0.52 (0.09)	0.25
30. Really like attention	0.48 (0.06)	0.22	28. Fads	0.46 (0.10)	0.15
7. Like attention	0.48 (0.07)	0.20	38. Upset	0.56 (0.10)	0.13
19. Like to look at body	0.43 (0.07)	0.18			
23. Stories	0.40 (0.08)	0.18			
20. Show off	0.49 (0.08)	0,16			
Scale H = 0.47 (0.06)			Scale H = 0.54 (0.08)		
Scale H^T= 0.20			Scale H^T= 0.41		
α = 0.80			α = 0.66		
Scale 3			Scale 3		
17. Responsibility	0.44 (0.12)	0.48	9. Special	0.54 (0.08)	0.27
18. Be something	0.50 (0.08)	0.24	40. Extraordinary	0.50 (0.08)	0.22
15. Show off body	0.37 (0.11)	0.17	20. Show off	0.40 (0.09)	0.20
Scale H = 0.44 (0.09)			Scale H = 0.48 (0.08)		
Scale H^T= 0.40			Scale H^T= 0.02		
α = 0.52			α = 0.69		
Scale 4			Scale 4		
13. Manipulate	0.46 (0.10)	0.23	12. Have authority	0.55 (0.10)	0.33
6. Talk	0.43 (0.09)	0.16	13. Manipulate	0.46 (0.09)	0.20
39. Capable	0.46 (0.09)	0.15	2. Modesty	0.50 (0.10)	0.15
Scale H = 0.44 (0.08)			Scale H = 0.50 (0.08)		
Scale H^T= 0.06			Scale H^T= 0.23		
α = 0.64			α = 0.62		
			Scale 5		
			33. Leader	0.42 (0.09)	0.29
			32.Recognise authority	0.47 (0.08)	0.28
			28. Stories	0.49 (0.11)	0.15
			Scale H = 0.46 (0.08)		
			Scale H^T= 0.17		
			α = 0.62		

Note: Minimum loading used =.40. Scale H should exceed .30. H_T = coefficient H of items.

Compared to the structure of the NPI subfactors, scale one of the female group aligned mainly with leadership and authority, traits from the remaining sub-scales were spread amongst the remaining Mokken scales, addressing superiority, exhibitionism, superiority, exploitativeness, self-sufficiency and vanity. For the male group, the first Mokken scale mainly tapped into vanity, with the remaining scales incorporating items from the NPI facets. Females found the item 17 (I like to take responsibility for making decisions), the easiest item to endorse, for the males it was 26 (I like to be complimented). Both groups showed 38 (I get upset when people don't notice how I look when I go out in public), the most difficult.

Machiavellianism

Mach IV

Females and males 18-25

For both samples, two Mokken scales were formed. The female sample evidenced six unscalable items (2, 11, 12, 14, 17 and 20), whilst the males showed thirteen items (1, 2, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19 and 20). For both samples, the scales were within the accepted ranges (Table 8).

Table 8

H-coefficients for Mach IV Females and Males 18-25

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females 18-25			Males 18-25		
Scale 1			Scale 1		
6. Viscous	0.47 (0.06)	3.89	4. Get ahead	0.39 (0.06)	4.19
4. Get ahead	0.45 (0.05)	3.87	3. Trust	0.41 (0.06)	4.10
3. Trust	0.46 (0.05)	3.74	6. Viscous	0.44 (0.06)	4.09
7. Never tell	0.55 (0.04)	3.15	7. Never tell	0.50 (0.05)	3.53
19. Stupid	0.46 (0.05)	2.93			
Scale H = 0.47 (0.04)			Scale H = 0.44 (0.05)		
Scale H^T= 0.17			Scale H^T= 0.08		
α = 0.80			α = 0.74		
Scale 2			Scale 2		
15.Lying	0.45 0.06)	4.51	16. Lying	0.43 (0.07)	4.47
5.Honesty	0.56 (0.05)	3.24	8. Moral	0.44 (0.07)	3.61
10.Humble	0.40 (0.07)	2.19	5. Honesty	0.45 (0.07)	3.31
Scale H = 0.47 (0.06)			Scale H = 0.44 (0.06)		
Scale H^T= 0.75			Scale H^T= 0.35		
α = 0.69			α = 0.68		

Note: Minimum loading used =.40. Scale H should exceed .30. H_T = coefficient H of items.

For both groups, scale one mainly addressed cynical misanthropy whilst scale two tapped into immorality and dishonesty. Both groups found indicator 15 (There is no excuse for lying to someone else), the easiest to endorse, whilst the females evidenced item 10 (All in all, it is better to be humble and honest than important and dishonest), the most difficult and for the male group, item 5 (Honesty is the best policy in all cases).

Females and males over 26

Both samples returned two viable Mokken scales. The female sample had eight items which were unscalable (1, 2, 9, 11, 12, 13, 14 and 16), as did the males (2, 6, 9, 11, 12, 16, 18 and 20). The Mokken scales had no significant violations with all crit values at zero (Table 9).

Table 9

H-coefficients for Mach IV Females and Males 26 and over

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
Scale 1			Scale 1		
15. Lying	0.50 (0.06)	4.03	15. Lying	0.46 (0.06)	4.26
8. Moral	0.44 (0.06)	3.22	13. Be good	0.44 (0.06)	3.58
5. Honesty	0.51 (0.05)	3.08	8. Moral	0.42 (0.07)	3.34
10. Humble	0.42 (0.09)	1.83	5. Honesty	0.55 (0.05)	3.16
			10. Humble	0.44 (0.07)	2.03
Scale H = 0.47 (0.06)			Scale H = 0.46 (0.05)		
Scale H^T= 0.58			Scale H^T= 0.47		
α = 0.75			α = 0.78		
Scale 2			Scale 2		
3. Trust	0.53 (0.06)	3.83	4. Get ahead	0.45 (0.06)	3.99
6. Viscous	0.57 (0.05)	3.50	1. Want to hear	0.52 (0.05)	3.83
7. Never tell	0.51 (0.06)	2.84	3. Trust	0.42 (0.06)	3.77
			7. Never tell	0.45 (0.06)	3.25
			19. Stupid	0.41 (0.06)	2.67
Scale H = 0.54 (0.05)			Scale H = 0.45 (0.05)		
Scale H^T= 0.26			Scale H^T= 0.23		
α = 0.75			α = 0.79		

Note: Minimum loading used =.40. Scale H should exceed .30. H_T = coefficient H of items.

The first scale for both groups pertained to dishonesty and immorality, whilst the second scale for the females addressed cynical misanthropy. For the males, the indicators tapped into immorality, manipulation, and cynical misanthropy. Both

groups found item 15 (There is no excuse for lying to someone else), the easiest to endorse, as with the younger samples. The more difficult item for males and females was 10 (All in all, it is better to be humble and honest than important and dishonest).

Dirty Dozen

Females and males 18-25

The data from both groups formed two Mokken scales. For the females, item P4 was unscalable, whilst for the males, items P4 and M3 were unscalable. All Mokken scales when assessed for MH and IIO evidenced no significant violations and all crit values at zero (Table 10).

Table 10

H-coefficients for Dirty Dozen Females and Males 18-25

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
Scale 1			Scale 1		
M4. Exploit	0.56 (0.06)	4.51	M4. Exploit	0.56 (0.05)	4.19
P1. Remorse	0.51 (0.05)	4.43	P1. Remorse	0.56 (0.05)	4.13
P2. Morality	0.49 (0.07)	4.41	M1. Manipulate	0.54 (0.06)	4.10
N4. Favours	0.45 (0.07)	4.34	P2. Morality	0.55 (0.05)	4.07
P3. Callous	0.40 (0.08)	4.24	P3. Callous	0.54 (0.06)	3.98
M1. Manipulate	0.53 (0.06)	4.20	M2. Deceit	0.52 (0.05)	3.67
M2. Deceit	0.54 (0.05)	3.79			
M3. Flattery	0.42 (0.07)	3.77			
Scale H = 0.48 (0.05)			Scale H = 0.55 (0.04)		
Scale H^T= 0.22			Scale H^T= 0.09		
α = 0.86			α = 0.86		
Scale 2			Scale 2		
N3. Status	0.51 (0.07)	3.59	N4. Favours	0.53 (0.07)	4.20
N2. Attention	0.58 (0.05)	3.44	N3. Status	0.54 (0.06)	3.65
N1. Admire	0.65 (0.04)	3.28	N2. Attention	0.54 (0.06)	3.56
			N1. Admire	0.54 (0.06)	3.40
Scale H = 0.58 (0.05)			Scale H = 0.54 (0.05)		
Scale H^T= 0.06			Scale H^T= 0.27		
α = 0.78			α = 0.80		

Note: Minimum loading used =.40. Scale H should exceed .30. H_T = coefficient H of items.

For both 18-25 groups a dark dyad of Machiavellianism/psychopathy is shown with narcissism as a separate facet. Both groups found that M4 (I tend to exploit others towards my own end), was the easiest item to endorse, whilst N1 (I tend to want others to admire me), was the most difficult.

Females and males over 26

Two scales were returned for the female sample, whilst the data from the males showed a unidimensional scale. For both samples, item P4 (I tend to be cynical), was unscalable and IIO evidenced that item M4 was outside of the accepted ranges, therefore, it was removed. The final scales were within the accepted boundaries (Table 11).

Table 11

H-coefficients for Dirty Dozen Females and Males 26 and over

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
Scale 1			Scale 1		
P2. Morality	0.52 (0.08)	4.66	P1. Remorse	0.42 (0.07)	4.29
N4. Favours	0.44 (0.10)	4.62	P2. Morality	0.48 (0.06)	4.27
P1. Remorse	0.44 (0.10)	4.61	N4. Favours	0.49 (0.06)	4.26
M1. Manipulate	0.57 (0.07)	4.32	M1. Manipulate	0.53 (0.06)	4.17
M2. Deceit	0.59 (0.07)	4.28	M2. Deceit	0.57 (0.05)	4.01
M3. Flattery	0.45 (0.09)	4.04	N3. Status	0.43 (0.07)	3.95
			M3. Flattery	0.50 (0.06)	3.74
Scale H = 0.50 (0.07)			Scale H = 0.49 (0.05)		
Scale H^T= 0.25			Scale H^T= 0.11		
α = 0.82			α = 0.85		
Scale 2					
N3. Status	0.67 (0.06)	4.21			
N2. Attention	0.68 (0.05)	3.99			
N1. Admire	0.71 (0.06)	3.81			
Scale H = 0.69 (0.05)					
Scale H^T= 0.20					
α = 0.83					

Note: Minimum loading used =.40. Scale *H* should exceed .30. H_T = coefficient *H* of items.

In relation to item difficulty, the females found P2 (I tend to be unconcerned with the morality of my actions), the easiest to endorse, whilst for the males it was P1 (I tend to lack remorse). For item difficulty, the females found N1 (I tend to want others to admire me), the hardest to endorse and the males, M3 (I have used flattery to get my way).

Short Dark Triad

Females and males 18-25

Three scales were returned for the females and two for the male sample. Eleven items were unscalable (M4, M8, M9, N1, N7, N8, N9, P2, P5, P7 and P8) for the females and seventeen items for the males (M1, M4, M7, M8, M9, N2, N3, N5, N6, N8, P1, P2, P3, P4, P5, P7 and P8). Item M6 was removed from the male scale following IIO analysis. All scales for both samples then evidenced no significant violations with all crit values at zero (Table 12).

Table 12

H-coefficients for Short Dark Triad Females and Males 18-25

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
Scale 1			Scale 1		
M3. Important	0.45 (0.06)	2.82	M5. Wise	0.46 (0.06)	3.06
M6. Get back	0.49 (0.05)	2.73	M3. Important	0.47 (0.05)	2.96
M5. Wise	0.53 (0.05)	2.72	M2. Manipulation	0.47 (0.06)	2.61
M2. Manipulation	0.44 (0.06)	2.19	P9. Say anything	0.50 (0.05)	2.27
P9. Say anything	0.42 (0.06)	2.06			
P3. Payback	0.50 (0.05)	1.98			
Scale *H* = 0.47 (0.05)			Scale *H* = 0.48 (0.05)		
Scale H^T= 0.25			Scale H^T= 0.23		
a = 0.83			a = 0.75		
Scale 2			Scale 2		
P6. Regret	0.48 (0.05)	2.15	N1. Leader	0.51 (0.06)	2.48
P1. Revenge	0.43 (0.06)	1.92	N7. Famous	0.49 (0.06)	2.31
P4. Out of control	0.40 (0.07)	1.71	N4. Special	0.52 (0.06)	2.26
Scale *H* = 0.44 (0.05)			Scale *H* = 0.50 (0.05)		
Scale H^T= 0.14			Scale H^T= 0.02		
a = 0.66			a = 0.72		
Scale 3					
N5. Acquainted	0.42 (0.06)	2.98			
N2. Attention	0.44 (0.06)	2.49			
N6. Embarrassed	0.44 (0.06)	2.41			
Scale *H* = 0.43 (0.05)					
Scale H^T= 0.18					
a = 0.65					

Note: Minimum loading used =.40. Scale *H* should exceed .30. H_T = coefficient *H* of items.

The female sample provided three Mokken scales whereby the first scale contained items pertaining to a dark dyad with the combination of Machiavellian and psychopathy indicators. The second scale addressed antisocial behaviour whilst the third scale tapped into narcissism. For the male sample, a dark dyad with narcissism as a separate facet was shown.

The female group found the item N5 (I like to get acquainted with important people), the easiest to endorse and for the males, M5 (It's wise to keep track of information that you can use against people later). P4 (People say that I'm out of control), was the most difficult indicator for females and N4 (I know that I am special because everyone keeps telling me so), for the males.

Females and males over 26

Two viable scales were returned for the females by the AISP and four for the males. There were seven items which were unscalable for the females (N5, N7, N8, N9, P2, P7 and P8), and eight from the male data (M1, M9, N1, N8, N9, P2, P7 and P8). In the female sample, item M3 when assessed for IIO was outside of the accepted range and was removed. All scales for both samples were within the accepted ranges (Table 13).

Table 13

H-coefficients for Short Dark Triad Females and Males 26 and over

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
Scale 1			Scale 1		
M8. Manipulated	0.45 (0.05)	3.25	N5. Important	0.41 (0.07)	2.87
M6. Get back	0.50 (0.04)	2.56	N7. Famous	0.45 (0.05)	2.33
P5. Mean	0.47 (0.05)	2.54	N3. Dull	0.44 (0.06)	2.19
M5. Wise	0.49 (0.05)	2.42	P6. Regret	0.43 (0.05)	2.16
M2. Manipulation	0.52 (0.04)	2.02	N4. Special	0.54 (0.05)	2.01
P6. Regret	0.53 (0.04)	2.00			
P1. Revenge	0.49 (0.05)	1.80			
P9. Say anything	0.52 (0.04)	1.73			
P3. Payback	0.50 (0.05)	1.73			
P4. Out of control	0.42 (0.07)	1.53			
Scale *H* = 0.49 (0.04)			Scale *H* = 0.46 (0.05)		
Scale H^T= 0.45			Scale H^T= 0.23		
a = 0.88			α = 0.78		
Scale 2			Scale 2		
N1. Leader	0.45 (0.06)	2.74	M6. Get back	0.50 (0.06)	2.94
N4. Special	0.51 (0.05)	2.05	M5. Wise	0.49 (0.05)	2.85
N3. Dull	0.53 (0.06)	2.05	M8. Benefit	0.42 (0.06)	2.63
			M2. Manipulation	0.46 (0.06)	2.18
			P9. Say anything	0.48 (0.05)	2.03
Scale *H* = 0.49 (0.05)			Scale *H* = 0.47 (0.05)		
Scale H^T= 0.42			Scale H^T= 0.30		
a = 0.70			α = 0.79		
			Scale 3		
			P5. Mean	0.42 (0.06)	2.72
			P1. Revenge	0.44 (0.06)	2.20
			P3. Payback	0.40 (0.05)	1.91
			P4. Out of control	0.47 (0.06)	1.69
			Scale *H* = 0.43 (0.05)		
			Scale H^T= 0.38		
			α = 0.70		
			Scale 4		
			M7. Preserve	0.40 (0.09)	3.73
			M4. Conflict	0.47 (0.07)	3.33
			M3. On your side	0.42 (0.07)	2.79
			Scale *H* = 0.43 (0.07)		
			Scale H^T= 0.41		
			α = 0.65		

Note: Minimum loading used =.40. Scale *H* should exceed .30. H_T = coefficient *H* of items.

As can be seen, the female sample evidenced a dark dyad and narcissism structure. However, the males deviated from this. The item P6 (People who mess with me always regret it), which was incorporated with the narcissism indicators on scale one, could semantically be perceived as tapping into narcissistic rage (Pincus, Ansell

et al., 2009). P9 (I'll say anything to get what I want), which is found on scale 2 with the Machiavellian items, aligns with the theme of the indicators in relation to manipulation. Scale three addresses antisocial behaviour, whilst the final scale refers to cynical manipulation. The easiest item for the female group to endorse was M9 (Most people can be manipulated) and for the males, M7 (There are things you should hide from other people to preserve your reputation). Both groups found P4 (People often say I'm out of control), the most difficult indicator to endorse.

Everyday Sadism

Females and males 18-25

In relation to the SSIS, the female data returned two scales, whilst the items clustered on one scale for the males. Three items were unscalable for the females (S5, S7 and S8), and one for the males (S10). For the ASP, both samples evidenced one scale, with items A1 and A6 being unscalable. The VAST also produced one Mokken scale for both samples. For the females, three items were unscalable (V6, V9 and V11), whilst for the males there were eight items (V3, V4, V5, V6, V9, V10, V11 and V15). All scales had crit values of zero and no significant violations (Table 14).

Table 14

H-coefficients for Short Sadistic Impulse Scale, Assessment of Sadistic Personalities And Varieties of Sadistic Tendencies; Females and Males 18-25

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
SSIS Scale 1			SSIS Scale 1		
S6. Fantasies	0.70 (0.13)	0.11	S8. Intentionally	0.09 (0.09)	0.19
S3. Exciting	0.66 (0.13)	0.05	S9. Humiliated	0.23 (0.09)	0.13
S2. Enjoy hurting	0.58 (0.16)	0.05	S6. Fantasies	0.30 (0.08)	0.11
S1. Enjoy seeing	0.64 (0.22)	0.03	S5. Enjoy	0.18 (0.08)	0.07
			S7. Hurt	0.18 (0.06)	0.05
			S4. Enjoyment	0.21 (0.09)	0.05
			S3. Exciting	0.36 (0.09)	0.05
			S2. Enjoy hurting	0.33 (0.08)	0.05
			S1. Enjoy seeing	0.42 (0.10)	0.03
Scale *H* = 0.65 (0.13)			Scale *H* = 0.24 (0.06)		
Scale H^T= 0.32			Scale H^T= 0.14		
α = 0.74			α = 0.62		
SSIS Scale 2					
S10. Angry	0.46 (0.12)	0.47			
S9. Humiliated	0.37 (0.14)	0.13			
S4. Enjoyment	0.55 (0.12)	0.05			
Scale *H* = 0.45 (0.13)					
Scale H^T= 0.70					
α = 0.36					
ASP Scale 1			ASP Scale 1		
A7. Irritate	0.51 (0.06)	2.07	A4. Excites	0.40 (0.06)	2.27
A4. Excites	0.54 (0.04)	1.81	A7. Irritate	0.46 (0.06)	2.19
A3. Control	0.62 (0.05)	1.80	A3. Control	0.49 (0.05)	2.01
A2. Mock	0.64 (0.04)	1.59	A2. Mock	0.45 (0.05)	2.01
A9. Pushing	0.60 (0.05)	1.51	A5. Meant control	0.58 (0.04)	1.82
A5. Hurt	0.62 (0.05)	1.48	A8. Mocking people	0.53 (0.05)	1.68
A8. Pleasure	0.60 (0.05)	1.42	A9. Pushing	0.56 (0.04)	1.63
Scale *H* = 0.59 (0.04)			Scale *H* = 0.49 (0.04)		
Scale H^T= 0.22			Scale H^T= 0.15		
α = 0.88			α = 0.84		
VAST Scale 1			VAST Scale 1		
V7. Fighting	0.45 (0.07)	1.67	V7. Fighting	0.44 (0.06)	2.38
V2. Gory	0.47 (0.07)	1.61	V12. Dominate	0.43 (0.06)	1.99
V12.Dominate	0.44 (0.07)	1.60	V13. Suffer	0.62 (0.04)	1.56
V14.Mocking	0.46 (0.07)	1.35	V14. Mocking	0.52 (0.05)	1.54
V13 Suffer	0.61 (0.06)	1.27	V8. Hurting	0.58 (0.04)	1.51
V8. Physically hurting	0.61 (0.06)	1.24	V16. Tormenting	0.43 (0.06)	1.31
V16. Tormenting	0.43 (0.12)	1.10			
Scale *H* = 0.49 (0.06)			Scale *H* = 0.50 (0.04)		
Scale H^T= 0.27			Scale H^T= 0.42		
α = 0.82			α = 0.81		

Note: Minimum loading used =.40. Scale *H* should exceed .30. H_T = coefficient *H* of items.

The SSIS is a unidimensional scale as evidenced with the male sample. It was surprising to see the second scale for the female sample. Although scale one mainly relates to vicarious or sadistic fantasies, scale two appears to be associated with felt emotions both at an inter and intra-personal level. The ASP for both groups align with the unidimensional nature of the measure, whereas the VAST has been significantly reduced, with indicators for both males and females tapping into vicarious and direct sadism. For the SSIS, the females found S10 (Sometimes I get so angry I want to hurt people), the ASP, A7 (I think about hurting people who irritate me) and for VAST, V7 (I love the YouTube clips of people fighting), the easiest items to endorse. The male group also found V7 the easiest to endorse for the VAST, but diverged on the SSIS, with item S8 (I wouldn't intentionally hurt anyone) and ASP (Watching people get into fights excites me). In relation to item difficulty, for the females and the ASP, A8 (I get pleasure from mocking people in front of their friends), was the hardest item to endorse, whilst for males it was A9 (I never get tired of pushing people around). Both groups evidenced V16 (I enjoy tormenting animals – especially the nasty ones) from the VAST and S1 (I enjoy seeing people hurt), from the SSIS, as the most difficult.

Females and males over 26

For both the SSIS and ASP, one Mokken scale was suggested for both samples. In relation to the SSIS, items S5 and S8 were unscalable for the females and S8 and S10 for the males. For the ASP, items A1 and A6 were unscalable for both samples. The VAST comprised one scale for the females and two for the males. Four items were unscalable for the females (V4, V6, V9 and V15) and five for the males (V4, V5, V6, V9 and V15). The Mokken scales had all crit values at 0 with no significant violations as determined by MH and IIO (Table 15).

.

Table 15

H-coefficients for Short Sadistic Impulse Scale, Assessment of Sadistic Personalities, Varieties of Sadistic Tendencies; Females and Males 26 and over.

Scale	H_i (*SE*)	Mean	Scale	H_i (*SE*)	Mean
Females			Males		
SSIS Scale 1			SSIS Scale 1		
S10 Angry	0.69 (0.12)	0.21	S6. Fantasies	0.60 (0.11)	0.18
S9. Humiliated	0.52 (0.15)	0.07	S9. Humiliated	0.56 (0.10)	0.10
S7. Hurt people	0.62 (0.13)	0.05	S7. Hurt	0.58 (0.10)	0.10
S6 Fantasies	0.51 (0.17)	0.04	S4. Enjoyment	0.51 (0.11)	0.08
S4 Enjoyment	0.57 (0.17)	0.03	S5. Hurting others	0.46 (0.13)	0.07
S3. Exciting	0.75 (0.11)	0.02	S3. Exciting	0.57 (0.13)	0.05
S2. Enjoy hurting	0.75 (0.11)	0.02	S2. Enjoy hurting	0.73 (0.08)	0.04
S1. Enjoy seeing	0.54 (0.21)	0.02	S1. Enjoy seeing	0.62 (0.14)	0.03
Scale *H* = 0.61 (0.12)			Scale H = 0.57 (0.09)		
Scale H^T= 0.54			Scale H^T= 0.25		
α = 0.77			α = 0.83		
ASP Scale 1			ASP Scale 1		
A7. Irritate	0.53 (0.06)	1.83	A7. Irritate	0.50 (0.06)	2.01
A3. Control	0.59 (0.07)	1.47	A2. Control	0.60 (0.05)	1.83
A4. Excites	0.48 (0.07)	1.39	A2. Mock	0.54 (0.06)	1.71
A9. Pushing	0.49 (0.07)	1.36	A5. Hurt	0.58 (0.05)	1.57
A2. Mock	0.61 (0.06)	1.35	A8. Mocking people	0.58 (0.05)	1.45
A8. Mocking people	0.66 (0.05)	1.34	A9. Pushing	0.60 (0.06)	1.38
A5. Control	0.66 (0.05)	1.26			
Scale H = 0.57 (0.05)			Scale H = 0.57 (0.05)		
Scale H^T= 0.26			Scale H^T= 0.25		
α = 0.87			α = 0.85		
VAST Scale 1			VAST Scale 1		
V10. Cruel	0.45 (0.08)	2.05	V10. Cruel	0.42 (0.06)	2.41
V12. Dominate	0.46 (0.06)	1.47	V12. Dominate	0.53 (0.05)	1.74
V11. Sex	0.44 (0.08)	1.46	V11. Sex	0.48 (0.06)	1.62
V14. Mocking	0.60 (0.06)	1.27	V14. Mocking	0.47 (0.06)	1.49
V13. Suffer	0.57 (0.06)	1.16	V13. Suffer	0.61 (0.05)	1.35
V8. Physically hurt	0.58 (0.10)	1.12	V8. Physically hurt	0.57 (0.06)	1.26
V16. Tormenting	0.47 (0.10)	1.09	V16. Torment	0.48 (0.07)	1.18
Scale H = 0.50 (0.06)			Scale H = 0.50 (0.05)		
Scale H^T= 0.55			Scale H^T= 0.52		
α = 0.78			α = 0.81		
			Scale 2.		
			V1. Blood	0.47 (0.05)	2.77
			V3. No escape	0.42 (0.06)	2.16
			V7. Fighting	0.41 (0.06)	2.00
			V2. Gory	0.47 (0.06)	1.77
			Scale H = 0.44 (0.05)		
			Scale H^T= 0.29		
			α = 0.72		

Note: Minimum loading used =.40. Scale *H* should exceed .30. H_T = coefficient *H* of items.

The SSIS and ASP both displayed a unidimensional structure in line with the measures. The VAST for the female group addressed direct sadism, whilst for the males, scale one tapped into direct sadism and scale two pertained to vicarious sadism, which is analogous to the two-factor structure of the measure. In relation to item difficulty, both groups showed A7 (I think about hurting people who irritate me), from the ASP as the easiest. For the VAST, females found V10 (I was purposely cruel to someone in high school), and S10 (Sometimes I get so angry I want to hurt people), from the SSIS, as the easiest to endorse. For males, the SSIS indicator S6 (I have fantasies which involve hurting other people), and V1 (In video games, I like the realistic blood spurts), from the VAST were the easiest. The more difficult items to endorse for both groups from the SSIS was S1 (I enjoy seeing people hurt) and V16 (I enjoy tormenting animals – especially the nasty ones), from the VAST. For the ASP, females showed A5 (I would hurt somebody if it meant that I would be in control), to be the most difficult, whilst for males it was A9 (I never get tired of pushing people around).

All Mokken scales across samples evidenced either moderate (>.40) or strong (>.50) scales, denoted by the *H* coefficient value (Ligtvoet et al., 2011). In relation to IIO, the only scales which evidenced strong ordering (>0.5) was the first scale of the NPI for males 18-25, SSIS for females 18-25 and VAST for males 26 and over. However, all H^T values were positive and therefore IIO could be assumed for each sample (Ligtvoet et al., 2011).

4.4 Discussion

The advancement in psychometric techniques enabled an in-depth view of the widely available Dark Tetrad (DT), measures from the top-down and bottom-up (Chiarotto et al., 2018; Van Schuur, 2003). Across multiple methods of statistical analyses, the questions mediated by the introductory discussion were answered. To recap, the present study explored whether the scales are sex and age invariant. If they are congruent with the constructs they seek to measure or whether the definitions of the constructs, as determined by the authors, are incompatible across and between age and sex. Considerations also focused on secondary psychopathy and the suggestion that its inclusion is disproportionate to general population samples (e.g., Boduszek, Dhingra, et al., 2016; Boduszek & Debowska, 2016; Cooke & Logan, 2015), as well as the assertation that Machiavellianism and psychopathy are parallel constructs (e.g., McHoskey et al., 1998; Miller, Hyatt et al., 2016; Persson. 2019;

Persson, Kajonius & Garcia, 2017a; Vize, Lynam et al., 2016). The findings, as will be demonstrated, revealed insights for the measures and the constructs.

Beginning with sex and age invariance, it was apparent from the confirmatory factor analyses and assessment of measurement invariance, that apart from the Assessment of Sadistic Personality (ASP), which achieved metric invariance in the 26 and over group, the scales are neither sex nor age invariant (Klimstra et al., 2020; Kawamoto et al., 2020). Therefore, moving forwards with the EGA, for psychopathy and the Levenson Self-Report Psychopathy Scale (LSRP), it was interesting to see both male groups aligning with the three-factor structure of egocentricity, callousness and antisocial behaviour. This structure was initially suggested by Brinkley et al., (2008), mediated by a sample of institutionalised females and that of Christian and Sellbom (2016), who used both student and general population samples. Conversely, the female groups evidenced a four-factor structure. Although in the older group, items pertaining to egocentricity were split across two dimensions, the indicators still aligned with the structure of egocentricity, callousness, lack of planfulness and negative emotionality. This is akin to the model found within a mixed sex institutionalised sample (Walters, Brinkley et al., 2008), and although the criminality aspect is not addressed in the LSRP, it corresponds with the four-factor structure determined by Hare and Neumann (2006), of interpersonal, affective, antisocial lifestyle and antisocial behaviour.

The MSA further elucidated the hierarchical ordering of the indicators relative to the levels of the latent construct the participants embodied (Shenkin et al., 2014). This reduced the female 18-25 structure to a three Mokken scale model comprising, egocentricity, callousness and lack of planfulness. Although the older females displayed five Mokken scales, the content still aligned with the four-factor model. The male groups were quite different, whereby males 18-25 had one scale tapping into egocentricity, which was echoed in the older group. Yet the latter group had a second scale mainly addressing boredom and short-temperedness. Previous research, using a student sample, has attributed these items more so to males (Gummelt et al., 2012). In the current study, it can be seen that apart from the younger males, these items were endorsed across age and sex. Yet, it can be questioned whether boredom and short temperedness are core features of the psychopathy construct. Indeed, evidence suggests that antisocial and/or criminal tendencies are a consequence rather than

characteristics of psychopathy (Lilienfeld & Andrews, 1996; Boduszek & Debowska, 2016).

Nevertheless, these findings did not clarify whether secondary psychopathy is necessarily redundant in the general population. The construct is not reflected in the DD, with the psychopathy indicators pertaining to primary characteristics of callousness, lack of morality and remorse (Kajonius et al., 2016). Conversely, the psychopathy items in the SD3 are analogous to secondary psychopathy (Persson, Kajonius & Garcia, 2017b; Vize, Lynam et al., 2018), mainly omitting the primary traits. There were many SD3 psychopathy items in the MSA which were unscalable, with seven out of the nine in the factor, being discarded in the males 18-25 group. Thus, the only psychopathy item from the SD3 which was homogenous across groups was, 'I'll say anything to get what I want', which can be aligned with manipulation and deception, a primary psychopathic characteristic (Kajonius et al., 2016). Although the SD3 psychopathy items formed Mokken scales or loaded onto scales with indicators from the remaining constructs, from an item level perspective, it would appear that secondary psychopathy as conceptualised by the SD3 and LSRP is not germane across age and sex in the general population.

A lack of invariance was evident for the Narcissistic Personality Inventory (NPI), whereby the younger females and both over 26 groups showed a clear subfactor of exploitativeness/entitlement and authority, the remaining facets were heterogenous across groups. Many clusters of the EGA contained indicators from various facets of the measure, which was reflected in the quantity of unscalable items in the MSA across samples. This provided further evidence as to why the seven-factor structure of the NPI across methods in past research has not been replicated (Barelds & Dijkstra, 2010).

The methods used revealed how the latent constructs manifest at both qualitative and quantitative levels (Sijtsma & van der Ark, 2017). As previously discussed in the introduction, assessments of narcissism based on self-report may mediate sex differences, which can lead some researchers to assume sex role stereotypes (Baez et al., 2017). However, the findings of the MSA elucidated that sex role stereotypes are not present in the NPI. For example, the measure is stated as mainly being a measure of grandiose narcissism which reflects the sex roles of men (Corry et al., 2008). Yet, the MSA displayed one Mokken scale for males 18-25, which addressed vanity. Indeed, items determined to be unscalable for the younger males

tapped into aspects of the construct discussed as being male oriented, including that of authority (Corry et al., 2008; Tschanz et al.,1998; Weiser, 2015). Although it is contended that manifestations of narcissism can incorporate vanity (Vazire et al., 2008), self-admiration does not determine the darker and more sinister side of the construct, which is defined by an inflated sense of superiority and entitlement (Baumeister, Bushman & Campbell, 2000). The results across methods therefore question whether the NPI is wholly congruent with the grandiose form of narcissism or whether some of the indicators more readily address self-esteem and positive psychological health (Brown, Budzek & Tamborski, 2009; Rosenthal & Hooley, 2010).

Clarity on how everyday sadism manifests across sex and age was provided. Initial examination of the means scores showed that males 18-25 embodied higher levels of the construct across groups. However, the measures did not achieve scalar invariance and although it is discussed that mean scores can accurately reflect sex differences in the construct (Chiorri et al., 2017), they cannot be assumed. Although there were problems with the SSIS with the CFA for the female over 26 sample, as discussed in the results section, MSA displayed a hierarchical scale for this group which was comparable with the counterpart male group. Indeed, for the everyday sadism measures overall, Mokken scales suggested that the item ordering and the coefficient *H* values were similar enough across groups to imply comparable levels of the construct across the continuum.

The measure of Machiavellianism, the Mach IV, has previously been stated as not being sensitive to the female manifestation (Brown & Guy, 1983; O'Connor & Simms, 1990). Yet, the model fit for the CFA across groups was poor, whilst EGA and MSA clarified the homogenous characteristics of cynical misanthropy, immorality, deceit and manipulation across clusters and Mokken scales, albeit with a reduced number of indicators. This aligns with the conceptualisation of the construct by Christie and Geis (1970) and was further corroborated by Rauthmann (2013).

However, one problem which may contribute to the instability of the Mach IV may be the third-person indicators, drawn from Machiavelli's book 'The Prince' (Christie & Geis, 1970), which may not demonstrate a good enough semantic equivalence with the construct (Shen et al., 2009). When participants endorse scale items, they do so through subjective interpretations. Yet, for third-person indicators, their responses may be based on hypothetical attitudes, which may not be reflected in their own behaviours (Rauthmann, 2013). The characteristics that the Mach IV

indicators tap into are construct congruent, yet the measure is psychometrically unsound.

A problem still persists though, whereby these characteristics are also pertinent to primary psychopathy. Looking at the findings for the Dirty Dozen (DD), over the Short Dark Triad (SD3), where, as discussed, the latter scale's psychopathy indicators are mainly analogous with secondary psychopathy, the EGA showed that apart from the younger male group, Machiavellian and psychopathy items mainly clustered together. Mokken scales for both 18-25 groups and females 26 and over, displayed a dark dyad of Machiavellianism/psychopathy, with narcissism as a separate facet. The older male group showed a unidimensional scale incorporating indicators from Machiavellianism, psychopathy and narcissism. The indicators from the SD3, merged with the psychopathy items when their focus related to primary traits. Thus, the concerns raised by previous researchers, whereby Machiavellian and primary psychopathy are considered parallel constructs and are not empirically separable in general population samples, is mainly supported (McHoskey et al., 1998; Miller, Hyatt et al., 2016; Persson. 2019; Vize, Lynam et al., 2016).

Cumulatively, deceit, immoral behaviour, egocentricity, manipulation, strategic planning and interpersonal cynicism are reflective of the successful psychopath (Babiak et al., 2010; Mayer et al., 2019; Palmen et al., 2018). When synthesised, primary psychopathy and Machiavellianism can be seen to represent this construct within general population samples (Babiak et al., 2006; Gao & Raine, 2010; Mullins-Sweatt et al., 2010). Comparisons of successful and unsuccessful psychopaths have shown that the former group have better executive functioning, increased autonomic reactivity and normal frontal and amygdala volumes and consequently, good impulse control. This is considered to safeguard them from extreme deviant behaviour (Gao & Raine, 2010; Palmen et al., 2018; Sellbom & Drislane, 2020). Outwardly, these individuals may be overtly charismatic, yet this masks their misanthropy, which mediates the effective manipulation and deceit of others in pursuit and achievement of their goals (Palmen et al., 2018).

The findings from this study have shown there are no repercussions from the male oriented aetiologies of the constructs, as discussed in Chapter Two. Across measures and methods of analyses, the narrative found within the DT literature, which suggests that the traits are embodied by males more so than females, in line with western cultural stereotypical associations (e.g., Corry et al., 2008; Jonason &

Webster, 2010; Levenson et al., 1995), was refuted. As has been evidenced, the measures do not hold scalar invariance and therefore the inter-groups mean scores cannot be meaningfully compared (Dong & Dumas, 2020). Psychometric problems were found with most of the widely available DT measures, with the exception, to an extent, of everyday sadism. As previously discussed, the more sophisticated means of statistical analyses enabled a more robust scrutiny of the measures, which were not available for the researchers who developed scales such as the LSRP, Mach IV and NPI. Thus, the narrative which puts forth stereotypical associations within the literature has been shown to be unjustified.

Limitations

This study was not without its limitations. All research is dependent on the data provided by the participants and when based on self-report, it may be considered a snapshot in time and sample dependent, which may address the inconsistencies in research. General population samples are likely to have lower levels of dark personality traits than institutionalised populations, however the scales were designed for use with non-institutionalised populations and therefore should capture the range of dark personality constructs within the samples used.

4.4.1 Conclusion

The prevalent scales which measure the dark personality constructs of Machiavellianism, psychopathy, narcissism and everyday sadism are mainly variant over age and sex as evidenced by the diverse statistical methods used in this study. As such, the findings imply that intergroup mean score comparisons may not necessarily be relied upon (Dong & Dumas, 2020). MSA elucidated that albeit with reduced items, the scales for everyday sadism were on the whole, homogenous across groups, whilst the NPI does not tap into the latent construct across males and females from both age groups. The Mach IV was shown to be psychometrically unsound, whilst Machiavellianism as a construct, when measured with the DD and to an extent the SD3, was mainly subsumed under a dark dyad. These findings corroborated previous research (e.g., McHoskey et al.,1998; Rauthmann, 2013). Yet it should be ascertained whether moving away from the widely used measures and using alternative indicators produces the same results. This would confirm that although Machiavellianism and psychopathy originated from divergent aetiologies, in general population samples, they merge to form the successful psychopath (e.g., Mullins-Sweatt et al., 2010).

Relative to psychopathy, the findings indicated a prevalence of the interpersonal and affective facet of the construct more so than the antisocial and lifestyle aspect. Further consideration will be given as to the relevance of secondary psychopathy, with recent findings suggesting this facet is a redundant construct (Cooke, Michie et al., 2004; Dinić, Wertag et al., 2020), and as mainly corroborated by the MSA in this study. It may be that outside of targeted or institutionalised samples, people in the general population embody more covert, primary traits and behaviours, which further suggests the concept of the successful psychopath is more germane.

Finally, it was evidenced that the male oriented aetiologies of the measures did not have any effect on the female endorsement of indicators. The narrative which has been present in previous research, whereby traits are embodied by males more so than females, in line with western cultural stereotypical associations (Corry et al., 2008; Jonason & Webster, 2010; Levenson et al., 1995), was refuted. The dark tetrad traits embodied by males and females eclipse the boundaries of stereotypical associations, which may be better elucidated when scales are developed which are both sex and age invariant.

Chapter Five

Reconstructing the Dark Tetrad: Preliminary Development of The Dark Side of Humanity Scale

5.1 Introduction

The results from the preceding chapter provided evidence which highlighted psychometric deficiencies and group differences within the widely available Dark Tetrad (DT) measures. It was interesting to note however, that the most recently developed measure investigated, the Assessment of Sadistic Personality (ASP; Plouffe, Saklofske & Smith, 2017), was the only scale which was mainly homogenous across and between age and sex as shown through exploratory graph analysis and Mokken scale analysis (MSA). The interpersonal and affective characteristics of primary psychopathy were consistent for males and females across ages, yet the secondary psychopathy facet, which address antisocial and lifestyle behaviours was mainly redundant (Cooke, Michie, Hart & Clark, 2004; Dinić, Wertag et al., 2020), as shown through MSA. Narcissism as measured by the Short Dark Triad (SD3; Jones & Paulhus, 2014) and Dirty Dozen (DD; Jonason & Webster, 2010), was primarily constant, yet the Narcissistic Personality Inventory (NPI: Raskin & Hall, 1979), did not address the complexities of the construct and a stable structure was not found. The findings from the Mach IV (Christie & Geis, 1979), elucidated the core features of Machiavellianism, however, these are seemingly analogous with those of the successful psychopath (Babiak, Neumann & Hare, 2010; Mayer et al., 2019; Palmen et al., 2018), which may explain why Machiavellian and primary psychopathy indicators are mainly subsumed under one factor, as seen in Chapter Four and previous research (McHoskey et al., 1998: Miller, Hyatt et al.,2017; Persson, 2019; Persson, Kajonius & Garcia, 2017; Siddiqi et al., 2020).

In light of these findings and to address the problems shown within the widely available DT measures, the aim of this study is to begin development of the Dark Side of Humanity Scale (DSHS). This scale will seek to clarify how the constructs of Machiavellianism, psychopathy, narcissism and everyday sadism manifest invariantly across and between age and sex, in the general population.

Consideration of theory and expert ratings of the constructs will mediate the indicators chosen for the item pool of the DSHS, whilst the items which were homogenous across groups from the MSA will also be incorporated. Theory

determines that a Machiavellian yields power, is misanthropic and manipulative, whereby, power, cynicism, deceit and manipulation are inextricably entwined (Christie & Geis, 1970). Expert raters (*n*=39), who had published multiple Dark Triad studies, contended that the key characteristics of a Machiavellian were of one who is strategic, cautious, persistent, ambitious, and organised (Miller, Hyatt et al., 2017). When comparing this profile to existing measures, such as the Dirty Dozen (Jonason & Webster, 2010), the experts declared the scales and their profile to be incongruent, whereby the measures held positive relationships with a wide range of impulsivity related traits. Indeed, the profiles were suggested as more analogous to expert conceptualisations of psychopathy than Machiavellianism (Miller, Hyatt et al., 2017). This has been corroborated in previous research (McHoskey et al., 1998; Persson, Kajonius & Garcia, 2017b), as well as in the study conducted in Chapter Four. To further ascertain whether there is a distinction between primary psychopathy and Machiavellianism, the indicators to measure the construct in the item pool will reflect the definition of: An individual who embodies cynicism, immorality, deceitfulness, manipulation and misanthropy, seeking power in organisational and interpersonal relationships through strategic planning (Rauthmann, 2013)

Turning to psychopathy, the core primary traits were historically determined from clinical and forensic samples by expert raters which incorporated a lack of remorse, deception, a charming demeanour but inept at forming close relationships, fearless with low anxiety, manipulative, callous and selfish (Cleckley, 1941; Karpman, 1948; Lykken, 1995; McCord & McCord, 1964). It was also suggested that antisocial behaviour and criminality were not core facets of psychopathy (Cleckley, 1941: Karpman, 1948). More recently, experts which included psychologists (*n*=82), criminal lawyers (*n*=25) and clinical psychology professors (*n*=41), did not endorse traits of impulsivity or irresponsibility, found in secondary psychopathy measures (Mullins-Sweatt et al., 2010). These secondary behaviours were considered to be embodied by the unsuccessful psychopath, an individual whose behaviours are a consequence of psychopathy, operating outside of social norms and legal boundaries, which may contribute to criminal convictions (Sellbom & Drislane, 2020). Conversely, and more specifically to primary psychopathy traits, those who operate within legal boundaries, found in general population samples, may be defined as the ‘successful psychopath’ (Cooke, Michie, Hart & Clark, 2004; Hall & Benning 2006; Mullins-Sweatt et al., 2010; Widom, 1977).

Research has provided evidence that both successful and unsuccessful psychopaths do not differ on the core primary traits (Ishikawa et al., 2001). Yet, antisocial behaviours prevalent in those high in secondary psychopathy, have been found to have significant relationships with mental health conditions such as antisocial and borderline personality disorders as well as anxiety. These behaviours can also be entwined with comorbidities of substance and/or alcohol abuse (Lilienfeld & Andrews, 1996; Lykken, 1995: Sellbom & Drislane, 2020; Smith, 1985; Sturek et al., 2008). The evidence therefore implies that antisocial behaviours and criminality do not determine the embodiment of psychopathy (Cooke & Logan, 2015; Cunliffe & Gacono, 2005; Klein, Tuente et al., 2014). In Chapter Four, it was found that the only items from the LSRP, which were homogenous across groups, were those from primary facet. Research has shown from general population samples, that individuals who display the core features of primary psychopathy have no significant relationships with mental health conditions (Sturek et al., 2008). Indeed, the literature has discussed many professions, such as politics, business, the military and police, where success in a given field may be due to the embodiment of primary psychopathic traits (Cleckley, 1941; Babiak, Hare & McLaren, 2006; Lilienfeld, Watts & Smith. 2015; Stevens et al., 2012).

Consequently, the DSHS will move away from the conceptualisation of psychopathy which incorporates the secondary facet. Instead, the focus will be In line with expert raters', theory and previous research, whereby the item pool for psychopathy, will be based on primary traits, to tap into the core of the construct. Concentrating on these key aspects will facilitate the ability to disentangle antisocial behaviours from those of the psychopath, thereby providing a more robust way of investigating external correlates (Lilienfeld & Andrews, 1996). The individual with psychopathic traits is therefore defined in the present study as one in society who operates within legal boundaries, propelled by intrapsychic goals, where their interpersonal charm and calm manner masks their traits of callousness, manipulation and lack of remorse (Palmen, Derkson & Kolthoff, 2020; Poythress & Hall, 2011).

Narcissism, as measured by the NPI, produced some concerning findings in the previous study, whereby neither the socially toxic nature of the construct nor the vulnerable facet was substantially addressed. The key characteristics of narcissism which formed a prototypical profile, was determined by expert ratings from researchers and mental health clinicians (Ackerman, Witt et al., 2011; Lynam &

Widiger, 2001; Rosenthal & Hooley, 2010; Samuel & Widiger, 2004; Thomas et al., 2012; Widiger et al., 2002), Cumulatively, grandiose narcissism was epitomised by one who is entitled, callous, manipulative and has an air of superiority, whilst a vulnerable narcissist was denoted by self-absorption, reactive aggression, a negative disposition and distrust of others. As can be seen, facets of the NPI which tap into leadership and authority were not included.

This latter point addresses a debate found within the literature. Some researchers argue that leadership is part of the construct, as determined by Freud (1950), and therefore, the dismissal of it in the DSM-III (APA, 1980), definition was considered neglectful (Campbell & Miller, 2011; Miller, Maples & Campbell, 2011). The flip-side of the debate contends that items pertaining to leadership may conflate non-narcissistic traits with narcissistic ones, whereby indicators which tap into leadership and authority can be analogised with investigations into self-esteem and positive psychological traits (Ackerman, Corretti & Carson, 2018; Peterson et al., 2009). This suggests a lack of discriminant validity between constructs, where the items which tap into leadership and authority are not apt for inclusion in a narcissism scale (Rosenthal & Hooley, 2010). Indeed, the leadership facet was only endorsed by two expert raters, out of a total of 19 clinicians and narcissism researchers, whilst the authority component received ten endorsements. The experts stated that these dimensions were the most adaptive and least relevant to the narcissism construct (Rosenthal & Hooley, 2010). Entitlement, however, received seventeen endorsements from the expert raters, and was determined to be the most pertinent characteristic, tapping into the core maladaptive aspect of narcissism (Dickinson & Pincus, 2003; Rosenthal & Hooley, 2010).

Entitlement and exploitativeness have further been shown as the bridge which joins grandiose and vulnerable narcissism (Dinić, Sokolvksa & Tomašević, 2021). Consequently, the current study defines grandiose narcissism through the traits of superiority, entitlement and exploitativeness. Vulnerable narcissism is defined through the traits and behaviours of reactive aggression, egotistical self-absorption and a negative disposition (e.g., Ackerman, Witt et al., 2011; Lynam & Widiger, 2001; Rosenthal & Hooley, 2010)

Finally, everyday sadism has been described as an ‘elusive construct’ (Foulkes, 2019), which may in part, be due to a universal definition not being reached by researchers (e.g., O’Meara et al., 2011; Plouffe, Saklofske & Smith, 2017). Expert

profiles have not emerged, which may be related to Sadistic Personality Disorder (SPD), being unrecognised in the DSM-V (APA, 2013). Although it was mentioned in the DSM-III-R as a disorder 'needing further study' (APA, 1987), it was never pursued due to its male stereotypical associations and high comorbidities, mainly with narcissistic and antisocial personality disorders (Fiester & Gay, 1991). Transcending from the forensic field to social psychology has elucidated that sadistic traits exist on a continuum in the general population (O'Meara et al., 2011). The core of the definition which is agreed upon is that sadists take pleasure from inflicting physical or psychological suffering on others (Baumeister & Campbell, 1999; Buckels, Jones & Paulhus, 2013; Pfattheicher, Keller & Knezevic, 2019). However, some researchers argue that as well as providing pleasure, the behaviours may also occur as a consequence of wanting to exert dominance or power over others (O'Meara et al., 2011; Plouffe, Saklofske & Smith, 2017).

It is therefore unclear whether everyday sadism as manifested in the general population incorporates power and domination. Taking pleasure from sadistic acts is agreed upon across researchers (e.g., O'Meara et al., 2011). Yet some authors argue that pleasure is derived through cruelty, not power (Book et al., 2015), whilst others incorporate power, whereby power within relationships can be sustained through sadistic behaviours (Paulhus & Dutton, 2015). In light of these disparities, the item pool for everyday sadism, will encompass both power and pleasure to shed further light on the construct and in turn a more definitive definition.

In summary, this study will seek to develop a scale which measures the traits of psychopathy, Machiavellianism, narcissism and everyday sadism, which is invariant across age and sex. This will be achieved by adhering to the way these constructs manifest, as determined by a general population sample. As elucidated in Chapter Four, the widely used DT scales investigated were found to be variant across age and sex. The DSHS seeks to rectify this.

5.2 Method

Scale Development

The pool of items used in the development of the DSHS contained those evidenced as homogenous across sex and age, based on the hierarchical scales of the MSA from the measures used in Chapter Four:

Machiavellian

I tend to manipulate others to get my way.

I have used deceit or lied to get my way.

I like to use clever manipulation to get my way

It's wise to keep track of information that you can use against people

It is hard to get ahead without cutting corners here and there.

Anyone who completely trusts anyone is asking for trouble.

Never tell anyone the real reason you did something unless it is useful to do so.

One should take action only when it is morally right.

Honesty is the best policy in all cases

Psychopathy

I tend to lack remorse.

I tend to be unconcerned with the morality of my actions.

I'll say anything to get what I want.

I tell other people what they want to hear so that they will do what I want them to.

Success is based on survival of the fittest, I am not concerned about the losers.

My main purpose in life is getting as many goodies as I can.

For me, what's right is whatever I can get away with.

In today's world, I feel justified in doing anything I can get away with.

Everyday Sadism

I have fantasies which involve hurting other people.

Hurting people would be exciting.

I would enjoy hurting someone physically, sexually or emotionally.

I enjoy seeing people hurt.

I have humiliated others to keep them in line.

I think about hurting people who irritate me.

I have made fun of people so that they know I am in control.

When I mock someone, it is funny to see them get upset.

I never get tired of pushing people around.

I would hurt somebody if it meant that I would be in control.

I get pleasure from mocking people in front of their friends.

I can dominate others using fear.

I enjoy mocking losers to their face.

I enjoy making people suffer.
I enjoy physically hurting people.
I enjoy tormenting animals – especially the nasty ones.

Narcissism

I like to show off my body.
I tend to seek prestige or status.
I tend to expect special favours from others.

With so few items emerging as homogenous, select items were taken from existing measures (e.g., Boduzsek, Debowska, Dhingra & DeLisi, 2016; Sherman et al., 2015), to address the conceptual gaps. These were chosen based on expert ratings, theory and past research, as previously discussed. Some items were used verbatim, whereas others were adapted although still very similar to the original item (Christian & Sellbom, 2016). Nine items were added to the pool for everyday sadism which related to online trolling. These items written were based on theory (Buckels et al., 2019), as well as discussions with colleagues who instigated the 'Don't feed the trolls' campaign (CCDH, 2020), to assess item construct, semantic and face validity, ensuring psychological relevance (Furr, 2011). The item pool, the scales they originated from and the facet of the construct they address, can be found in Appendix 4.1.

Likert Scale Options for the Dark Side of Humanity Scale

When developing a new scale, a salient consideration is the number of Likert scale options to include. The original conceptualisation of the Likert scale determined five-points with options reflecting strongly agree, agree, undecided/neither, disagree or strongly disagree (Likert, 1932; Simms et al., 2019). However, ongoing debates within the literature have not reached a definitive conclusion as to the ideal number of scale response options. Investigations have considered the neutral or middle point as well as reliability and concluded a four-point scale as optimal (Borgers et al., 2004). A further study examined both a four- and six-point measure, finding that the number of options had no effect on criterion-related validity (Chang 1994), yet it was discussed that seven Likert scale options may reach the peak of reliability (Allen & Seaman, 2007).

However, an alternative view is that a scale with more options may be preferable because it increases the total score variability which would boost validity and accuracy. This is an important consideration, especially when researchers use analysis methods such as CFA, as increases in valid score variance should mediate an increase in measurement accuracy, which maximises validity coefficients. This would provide confidence in the scale, whereby the variance in participants scores reflect reliable and valid differences (Finn et al., 2015; Hilbert et al., 2016).

However, an equivalent reliability has historically been found in scales which have three, five, six or nine scale options, although reliability decreased with eleven options (Bendig, 1953). This was partially corroborated by Leung (2011), yet, further evidence contended that scales with two and three choices, and to a lesser degree, four and five, mostly weakened the psychometric accuracy of the scale (Simms et al., 2019). Indeed, scales with two to four options have been shown as inadequate in relation to reliability, validity and discriminatory power (Preston & Colman, 2000), whilst scales with five options are considered too simple to capture the nuances of the construct being measured (Finstad, 2010; Russell & Bobko, 1992). The implications suggest that measures with few options reduces the accuracy and would require more indicators. This highlights the compromise between developing a short scale and the need for measurement precision (Simms, 2019).

A further decision is whether having an even or odd number of scale points makes a difference. The neutral choice in odd-numbered measures can be argued as being ambiguous. Removing the middle option may reduce social desirability effects however, the scale then becomes a forced choice measure, where participants are compelled to make a cognitive effort by adopting a definitive stance rather than taking a neutral one (Allen & Seaman, 2007; Krosnick et al., 2002). This may culminate in distorted results, if participants do not provide the required information for measures which tap into sensitive issues (Garland, 1981). It further removes the choice of endorsing a neutral option if the question evokes feelings of apathy or conflict (Garland, 1991; Johns, 2005; Nowlis et al., 2002). Yet, a measure with no neutral option may alleviate cognitive passivity and return data from which a more in-depth understanding of the construct being investigated can be achieved. Further, alpha criterion validity showed no advantage when measured by both odd and even numbered scales (Kulas & Stachowski, 2013; Simms et al., 2019).

Therefore, with a focus on a six- or seven-point Likert scale, whereby seven retains the neutral option, some authors have argued that the seven-point scale is the most accurate and easiest to use by participants and has produced stronger correlations with t-test results (Diefenbach et al.,1993; Lewis, 1993), with scales showing improved performance over those with lesser points (Preston & Colman, 2000). Yet, a more recent study by Simms et al., (2019), found that seven options may challenge participants, when faced with options which are semantically similar. Although measures with six and seven options have shown minimal to non-existent psychometric differences and thus, no apparent consequence for retaining the middle option, Simms et al., (2019), stated that a six-point measure is the preferred option on the grounds of parsimony.

On reflection of the evidence, the Dark Side of Humanity Scale will be measured on a six-point Likert scale, in line with previous research (Christian & Sellbom, 2016) and in order to increase the variance at the extremes of the distribution. The options range from *Not at all like me* (1), *Mainly unlike me* (2), *A little unlike me* (3), *A little like me* (4), *Mainly like me* (5), *Very much like me* (6), in response to each item.

5.2.1 Participants

An online study was designed to investigate a pool of 192 items which tapped into the core facets of psychopathy, narcissism, Machiavellianism and everyday sadism, to ascertain which of the items were invariant across age and sex. Ethical approval was received from the University of Derby (Appendix 4.2). Invitations to participate were posted on a crowd-data provider, Prolific (Prolific, 2019), and participants, who were all over the age of 18, were paid £1.67 for completion of the survey. Initially there were 678 recorded responses, eight participants had missing data, which were removed from the study, one participant preferred not to state their sex and two preferred to self-define. As a core focus of the study was the similarities and differences between males and females, these participants were removed from analysis. A final sample of 667 participants (334 females; 333 males), contributed data to this study, with an age range of 18-73 (M_{age} = 28.64, *SD* = 11.01). The mean ages of the sample split by groups; Males 18-25 (*n* = 175; M_{age} = 21.30, *SD* = 2.17); Females 18-25 (*n* = 157; M_{age} = 21.84, *SD* = 2.06); Males 26 and Over (*n* = 158; M_{age} = 36.51,*SD* = 9.67); Females 26 and Over (*n* = 177; M_{age} = 38.68, *SD* = 10.91).

5.2.2 Procedure

The invitation to participate provided a link to the online survey, hosted by Qualtrics (Qualtrics, 2019), thereby ensuring that involvement was voluntary and anonymous. The only exclusion criteria was that participants under the age of 18 could not take part. Prior to commencement of the survey, participants were briefed about the study and its aims, they were then asked for their consent to proceed. Once this was given, participants were presented with demographic questions, followed by the questionnaires, presented in a randomised order, for which they were informed there were no time restrictions for completion. Contact information was given for the researcher and supervisor, should participants wish to ask any questions about the nature of the study. They were also informed that they could withdraw their data within two weeks of taking part, with their responses removed from the study, by providing the unique identification code they had created before beginning the survey.

5.2.3 Data Analysis

Full details are provided in the methodology chapter. In brief, the analyses used in this study were conducted in R Studio, Version 1.2.1335. The structure of the item pool was initially investigated with Exploratory Graph Analysis (EGA) using the EGAnet package (version 0.9.8; Golino & Christensen, 2020). Previous research has shown that network models can determine the number of latent dimensions in psychological and personality data including openness to experience and the Dark Triad (Christensen, Cotter & Silvia, 2019; Golino & Epskamp, 2017; Truhan et al., 2020).

Dimensionality of the data was estimated using the graphical least absolute shrinkage and selection operator (GLASSO; Friedman et al., 2008), which estimates a Gaussian Graphical Model (GGM; Lauritzen, 1996) and the ‘Walktrap’ algorithm (Pons & Latapy, 2006), which is a weighted community detection algorithm (Fortunato, 2010). The indicators or nodes represent the variables and edges represent the conditional dependence (partial correlations) between nodes, given all other nodes in the network and denote statistical relationships (Epskamp & Fried, 2016). The least absolute shrinkage and selection operator (LASSO; Tibshirani, 1996) of the GLASSO, is a regularisation method which reduces parameter estimates, with some becoming exactly zero. The EGA displays formed clusters which represent underlying latent variables.

Item reduction was initially conducted with redundancy analysis, from the EGAnet package (Golino & Christensen, 2020). A similarity measure is calculated between the regularised partial correlations of the nodes using weighted topological (global structure) overlap from the wTO package (Gysi et al., 2018). The adaptive alpha (Pérez & Pericchi, 2014), applies multiple comparison corrections to determine which scale items are redundant (Constantini et al., 2015; Epskamp, Borsboom & Fried, 2017), by focusing on pairwise relationships between the indicators. This quantifies how similar the connections between nodes are. Those which have a large topological overlap are likely to have a shared latent influence. Using a multiple comparison method, node pairs whose *p* values are less than the corrected alpha are considered to be significantly redundant. Items considered redundant are further assessed based on theory and semantic representation (Christensen, Golino & Silvia, 2020). The analysis displays the target node and iteratively identifies other nodes that are redundant with it. This continues until all redundancies are resolved. The package documents the process which provides transparency over the decision-making process (Appendix 4.3.1).

Item Response theory (IRT), specifically the Generalised Partial Credit Model (GPCM; Muraki & Muraki. 2016), was used to determine the degree in which the scale items for each construct captured the latent trait levels in participants. The GPCM estimates two main parameters, the slope or discrimination parameter (*a*), which shows how well the indicator identifies participants at differing levels of the latent trait. A high value denotes a strong item (Adedoyin & Mokobi, 2013; An & Yung, 2014; Yang, 2014). The threshold or step difficulty parameter (*b*), indicates the level of the latent trait where participants move from one scale category to the next. The number of step difficulty parameters is equal to the number of categories minus one (Embretson & Reise, 2000). Based on the guidelines by Baker (2001), the slope values are determined by: very low (0.01-0.34), low (0.35-0.64), moderate (0.65-1.34), high (1.35-1.69) and very high (>1.70). GPCM is useful for both informing scale development and assessing models (Langer et al., 2008).

In relation to confirmatory factor analysis (CFA), the estimator was set as weighted least squared means and variance (WLSMV) and the model fit was assessed using chi square (χ^2), Confirmatory Fit Index (CFI), Tucker-Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA). A good model fit

was determined when CFI and TLI values were more than or equal to .90, in line with Kline's (2015), recommendations. RMSEA values evidenced a good fit if they were .06 or below and an adequate model fit was accepted at .08 or below. Models were rejected if they displayed a value above 1.0 (Brown, 2015; Hu & Bentler, 1998; Marsh, Wen & Hau, 2004).

Measurement invariance through multigroup nested CFA models, were used to determine whether the scales held for configural, metric and scaler invariance. Analysis was conducted in R using lavaan (Rosseel, 2020). The CFI and RMSEA considerations were the same as the CFA criteria already discussed. Research has shown that comparing models based on a chi-squared difference test are impacted by the same issues as the chi-squared goodness of fit test, with the change in CFI being less sensitive to sample size than the chi-square and more sensitive to lack of invariance than chi-square (Meade et al., 2008), which has led to the development of fit indices. If the difference in the fit indices (ΔCFI and ΔRMSEA) between a model and the preceding less constrained model was not larger than 0.01 for ΔCFI and equal or less than 0.015 for ΔRMSEA, then it was considered that the level of measurement invariance was achieved (Chen, 2007; Cheung & Rensvold, 2002; Marsh, Nagengast & Morin, 2013; Rudnev et al., 2018).

5.3 Results

Each construct was investigated with EGA and discrepancies were exposed for each group in relation to the number of clusters that emerged. It will be noted that clusters contained indicators which addressed more than one facet of the construct. Therefore, in this preliminary stage of development, there was, using psychopathy as an example, more than one cluster with items tapping into callousness.

Exploratory Graph Analysis

The EGA for Machiavellianism showed divergence in the cluster structure and item loadings across groups (Figure 25 and 26). The numbers and colours on the legend correspond with the cluster the indicators align with.

Machiavellianism

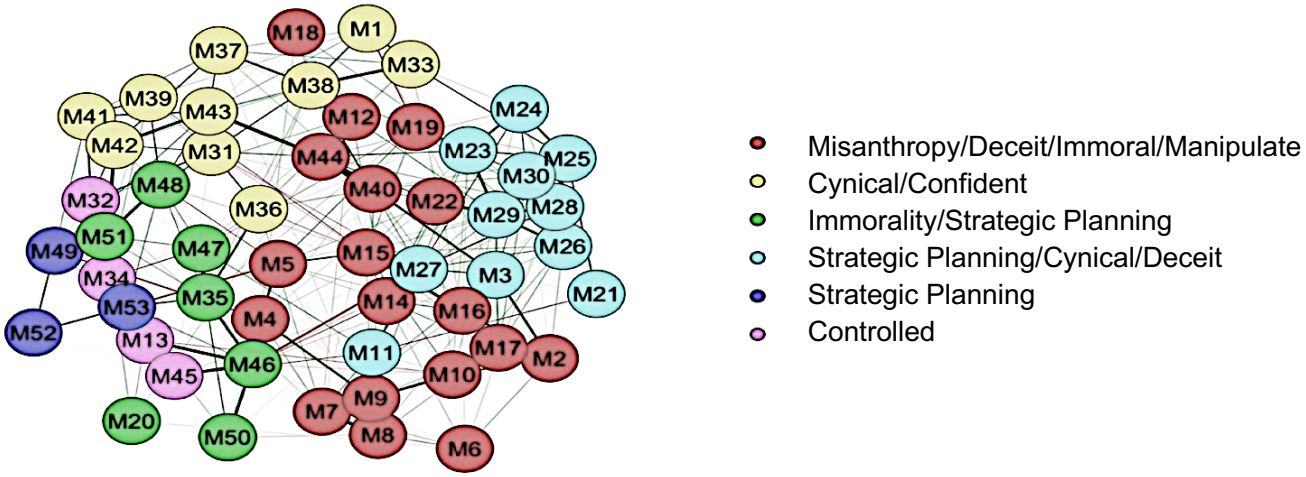
Males and Females 18-25

Strategic planning was the only clear dimension which emerged from the Machiavellian items for the males 18-25, although the characteristics of cynicism, immorality, deceitfulness, misanthropy and power were clearly present. The first two

clusters of the female group combined indicators across facets of the construct, however, the remaining clusters were clearly more defined. Thus, there was not homogeneity between the two 18-25 groups (Figure 20).

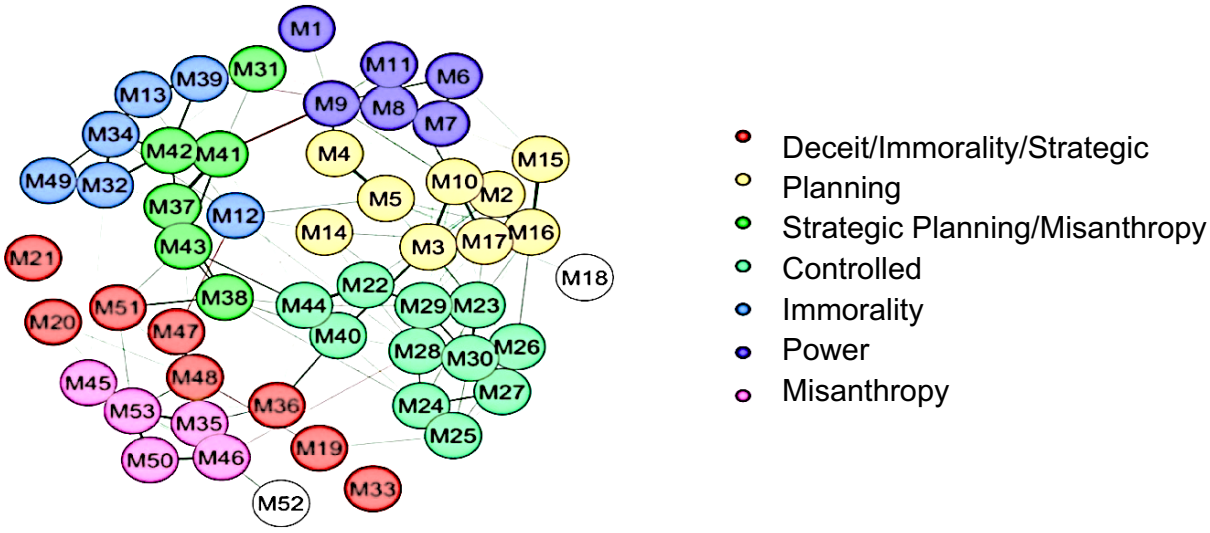
Figure 20

EGA Machiavellianism Males 18-25



Note: M = Machiavellian

EGA Machiavellianism Females 18-25



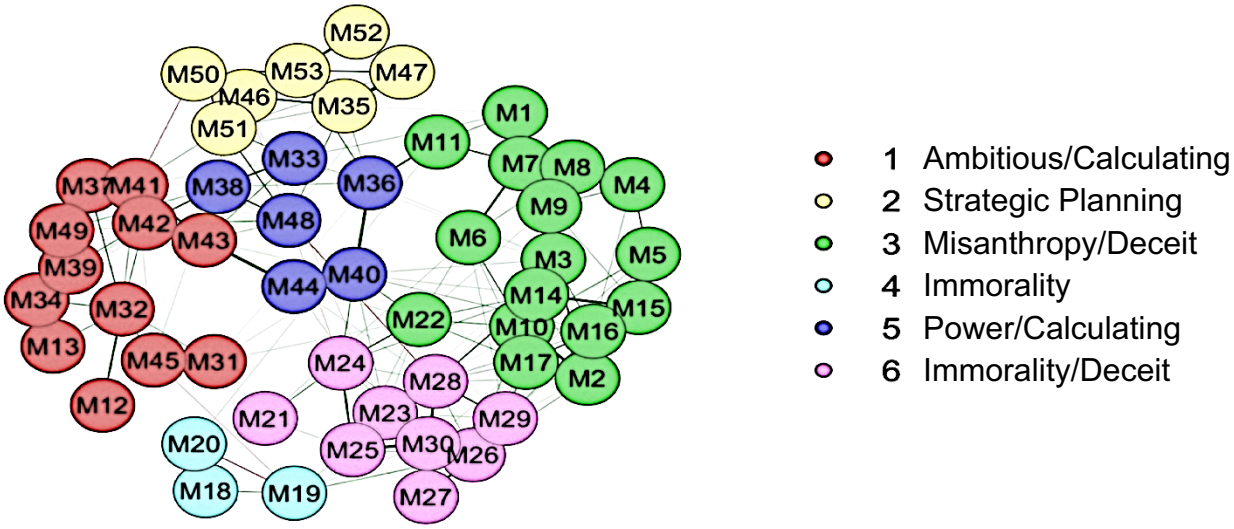
Note: M = Machiavellian

Males and Females 26 and Over

For the older groups, facets of immorality, misanthropy and strategic planning clearly emerged, yet indicators loaded onto different clusters for males and females. Apart from the number of clusters differing, items also loaded onto different dimensions (Figure 21).

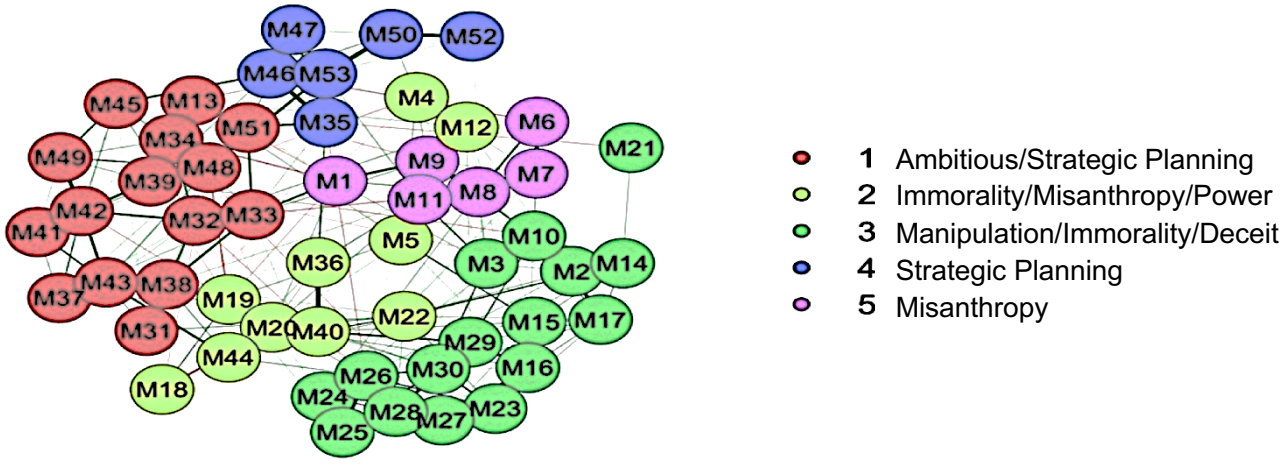
Figure 21

EGA Machiavellianism Males 26 and Over



Note: M = Machiavellian

EGA Machiavellianism Females 26 and Over



Note: M = Machiavellian

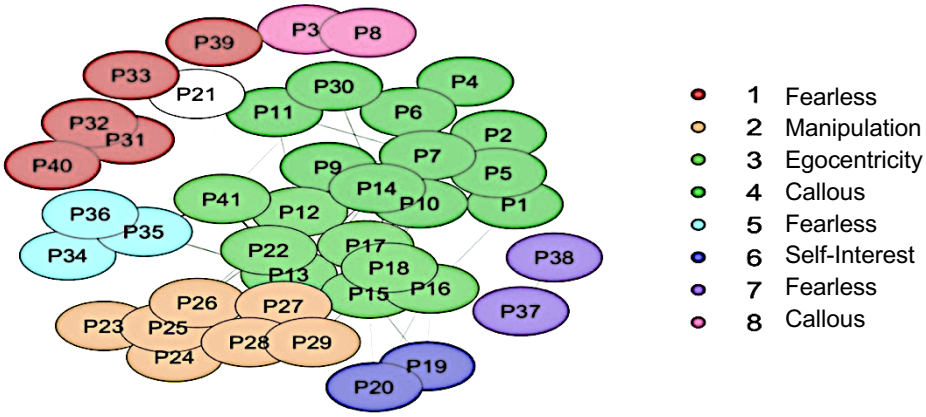
Psychopathy

Males and Females 18-25

Turning to psychopathy, again there was divergence in the number of clusters in this group. However, conversely to Machiavellianism, the clusters for psychopathy were far more clearly defined. Clusters six, seven and eight for the male group only contained two indicators each and therefore needed further refinement. Although there were deviations as to the indicators loading on clusters, these were far less than was found in Machiavellianism (Figure 22).

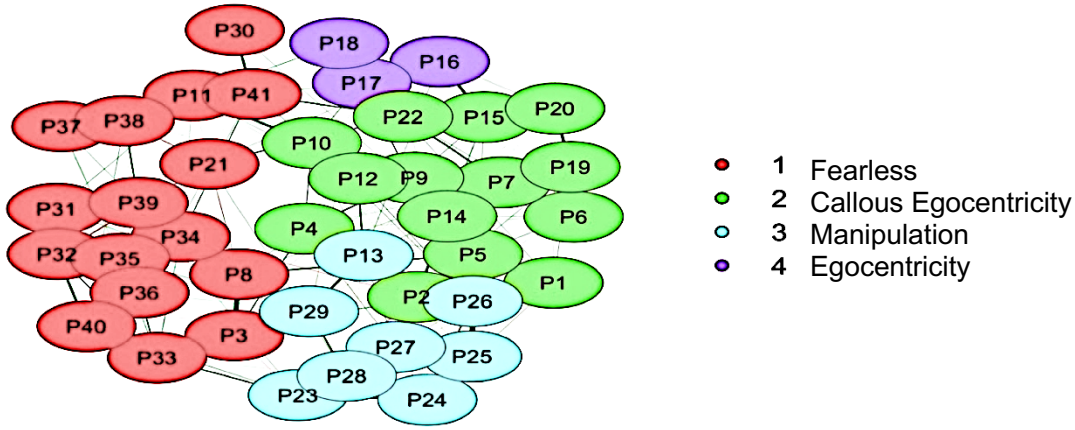
Figure 22

EGA Psychopathy Males 18-25



Note: P = Psychopathy

EGA Psychopathy Females 18-25



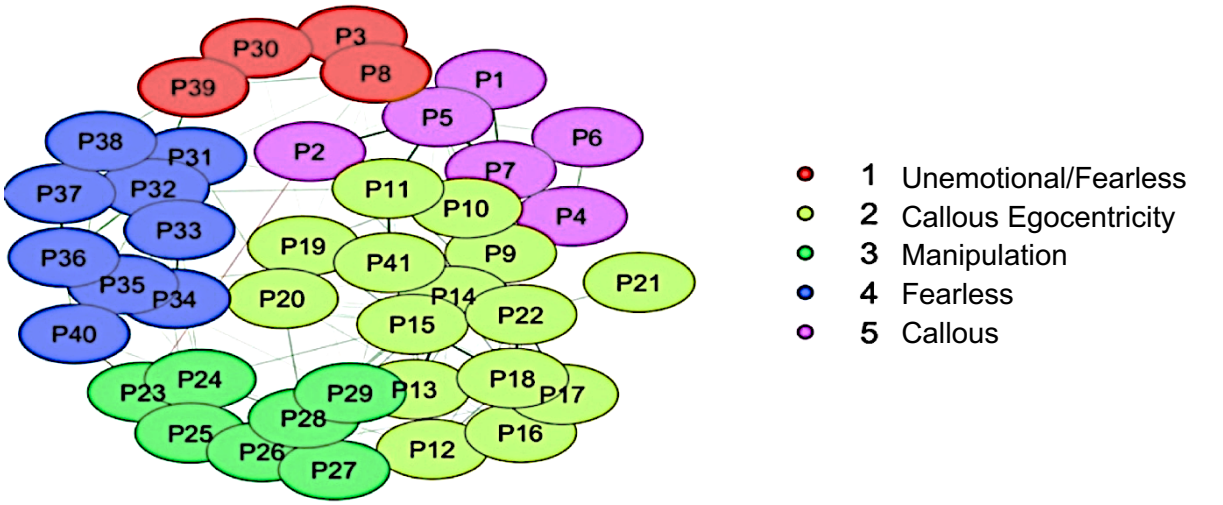
Note: P = Psychopathy

Males and Females 26 and Over

As with the younger group, the clusters aligned with the characteristics of primary psychopathy and the deviations of item loadings were minimal. The construct as shown by the initial EGA was mainly homogenous across age and sex (Figure 23).

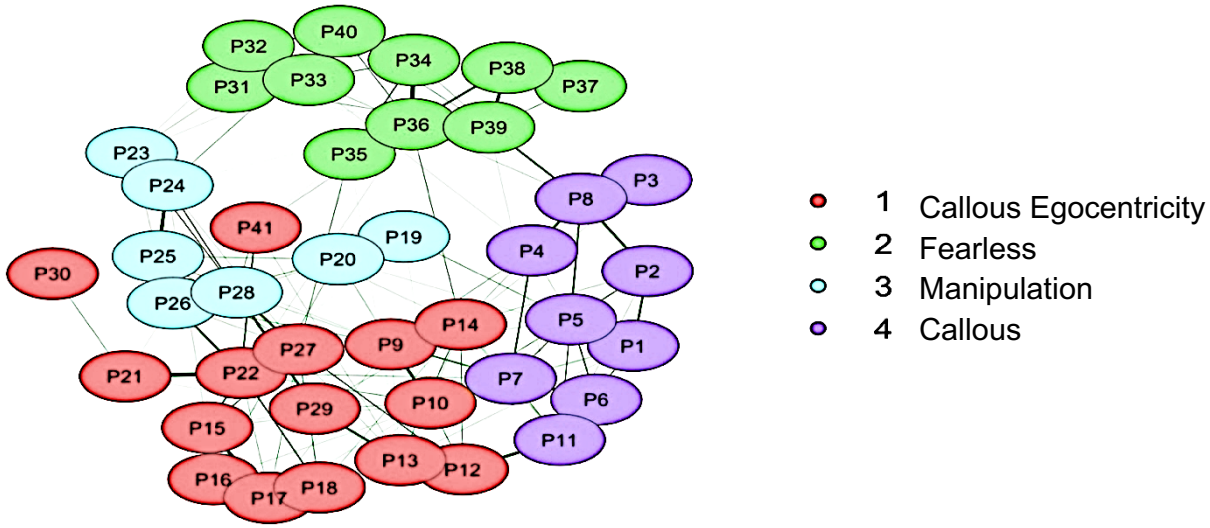
Figure 23

EGA Psychopathy Males 26 and Over



Note: P = Psychopathy

EGA Psychopathy Females 26 and Over



Note: P = Psychopathy

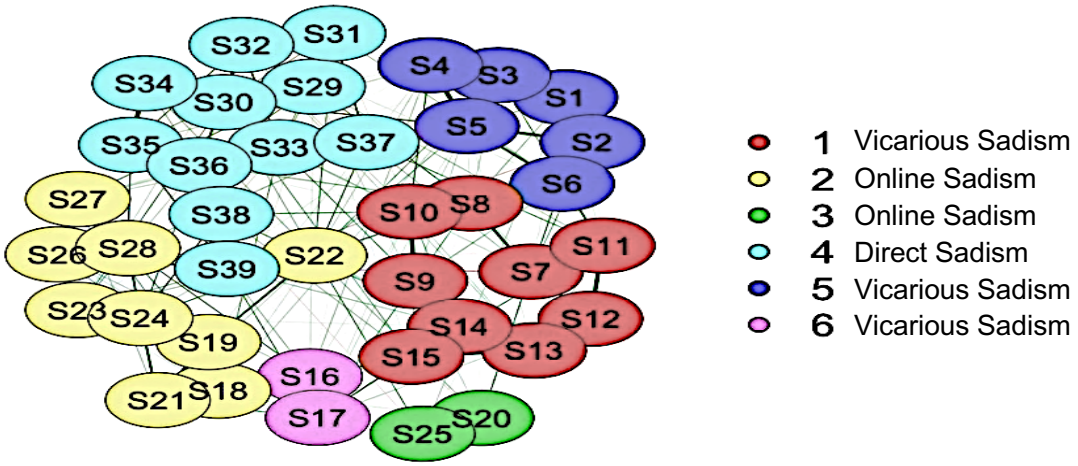
Everyday Sadism

Males and Females 18-25

Although very clear clusters emerged for both groups, cluster three and six for the males and cluster five for the females only contained two indicators. This mediated facets to be split across dimensions suggesting the indicators need further investigation (Figure 24).

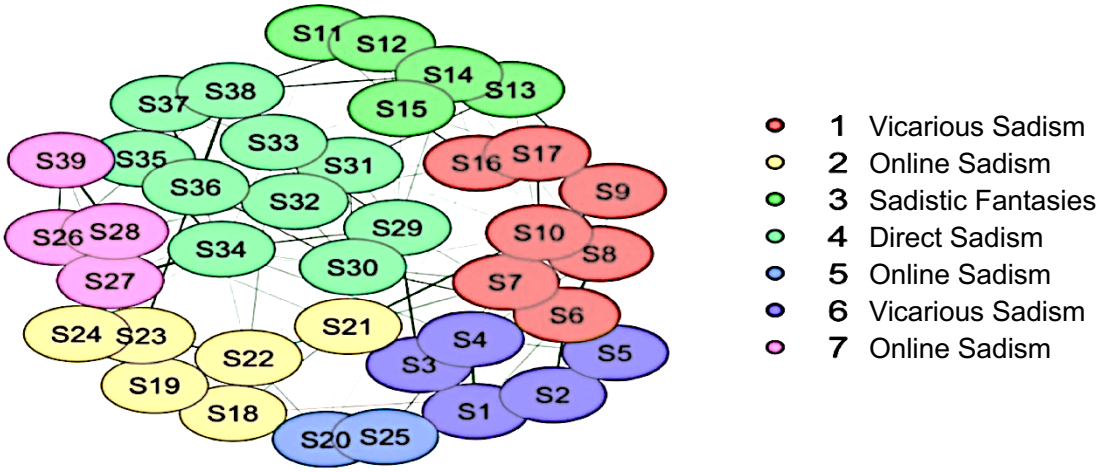
Figure 24

EGA Everyday Sadism Males 18-25



Note: S = Everyday Sadism

EGA Everyday Sadism Females 18-25



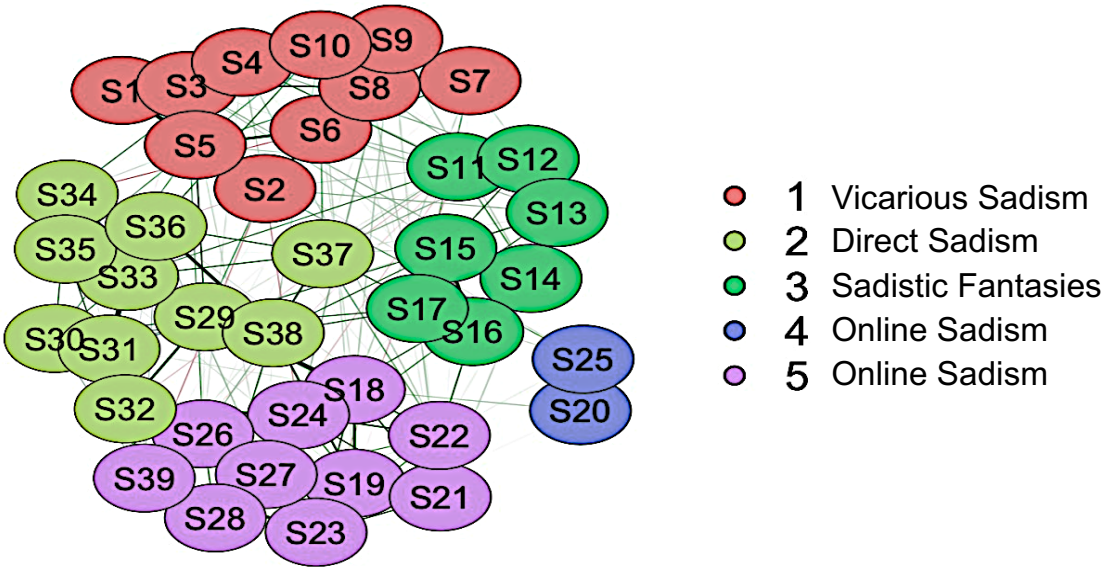
Note: S = Everyday Sadism

Males and Females 26 and Over

The clusters, as with the younger groups, were clearly defined (Figure 25). However, cluster four for the males and cluster five for the females only contained two items, which were the same (S20; Trolling behaviour is both cruel and unnecessary (R), and S25; Being anonymous on social media forums lets me enjoy abusing people with different religious beliefs to my own). Indeed, these two indicators formed a two-item cluster in both 18-25 samples. Therefore, although there is mainly homogeneity across groups, certain indicators mediated clusters which would not be psychometrically viable.

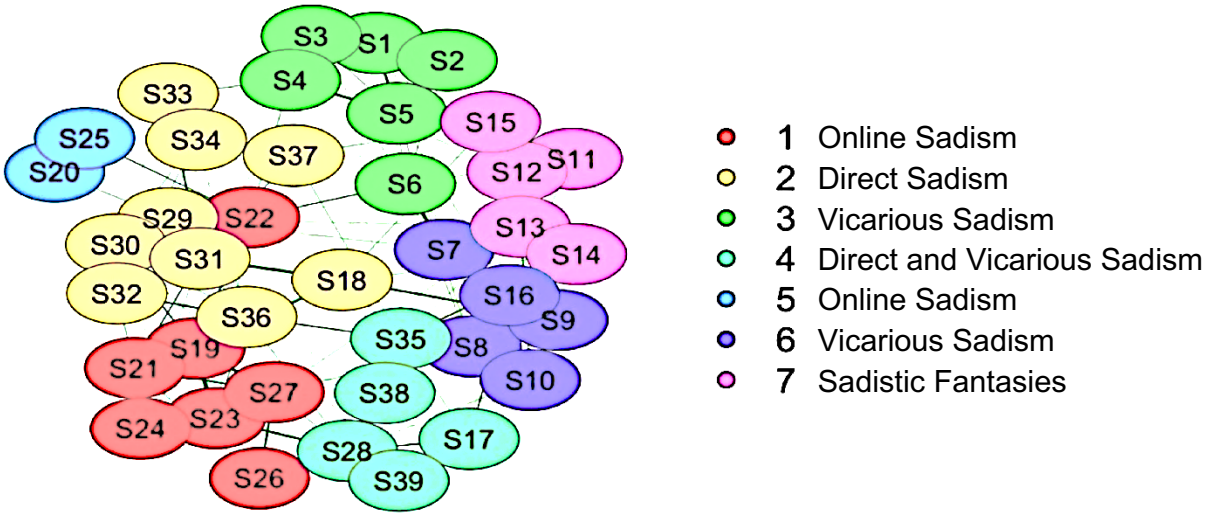
Figure 25

EGA Everyday Sadism Males 26 and Over



Note: S = Everyday Sadism

EGA Everyday Sadism Females 26 and Over



Note: S = Everyday Sadism

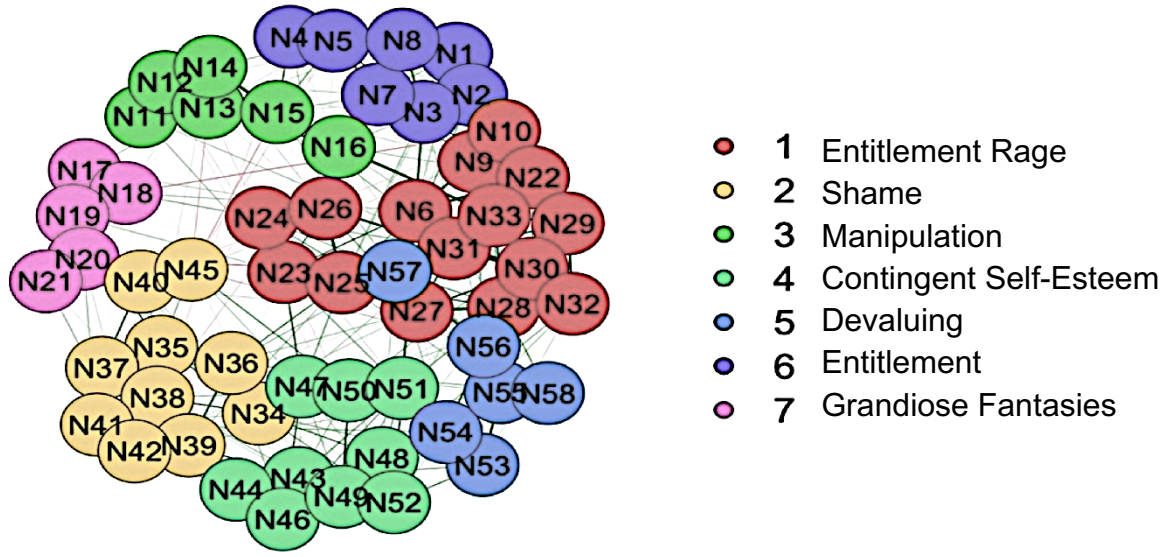
Narcissism

Males and Females 18-25

The two groups differed, with females displaying eight clusters, and males showing seven dimensions (Figure 26).

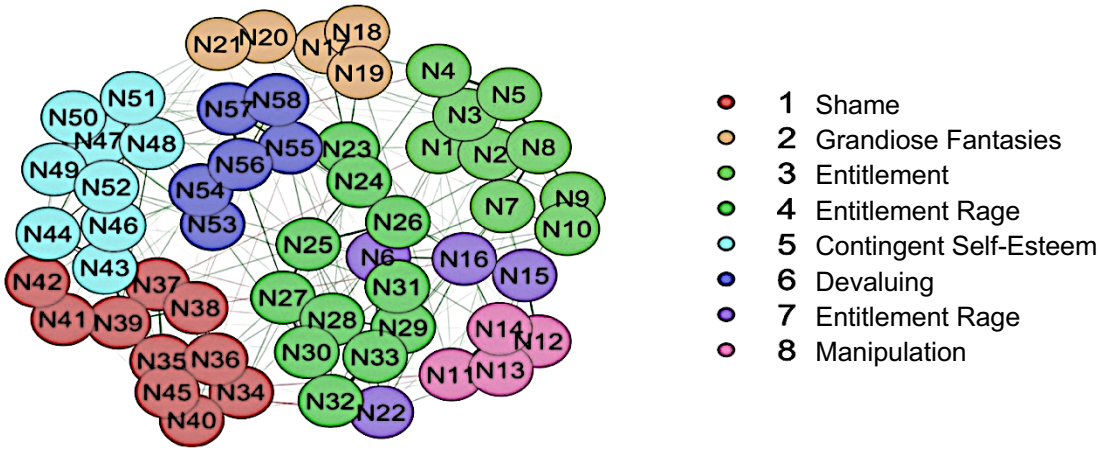
Figure 26

EGA Narcissism Males 18-25



Note: N = Narcissism

EGA Narcissism Females 18-25



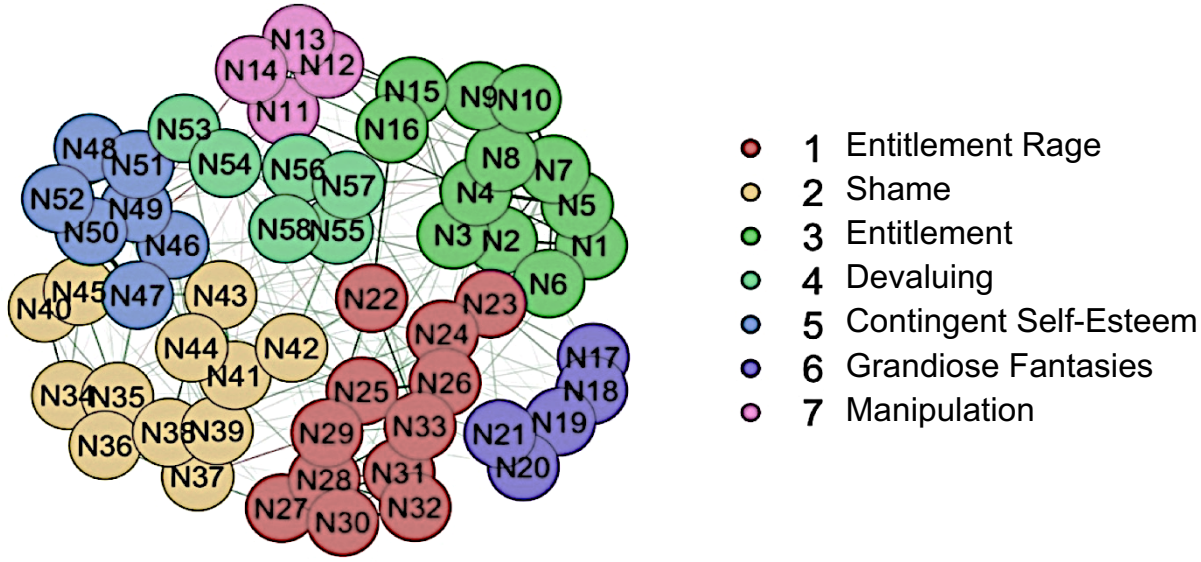
Note: N = Narcissism

Males and Females 26 and Over

Both groups displayed seven clusters and were mainly homogenous (Figure 27).

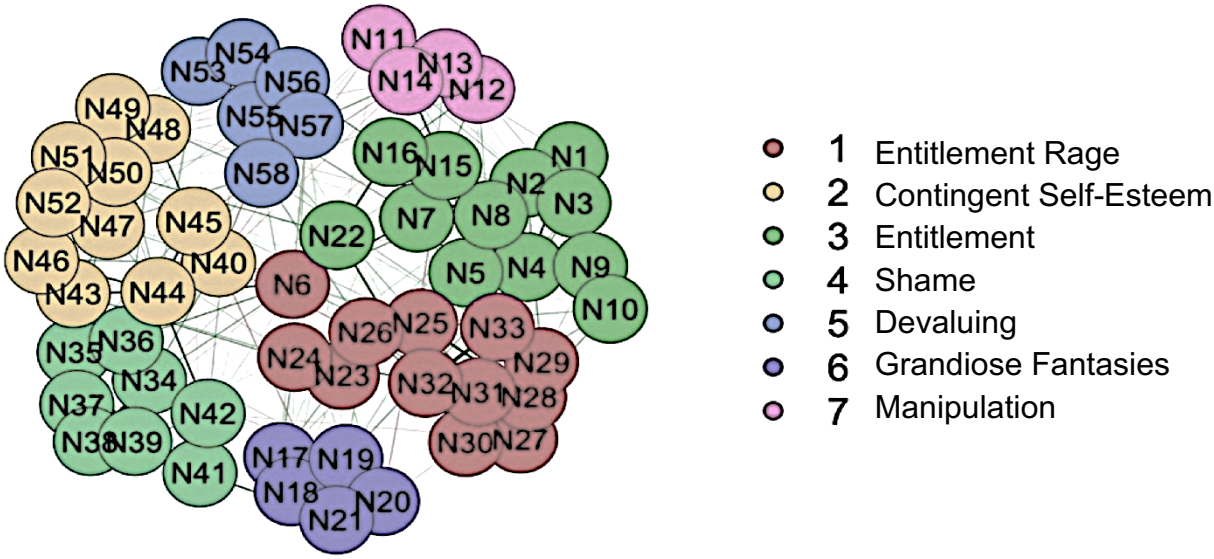
Figure 27

EGA Narcissism Males 26 and Over



Note: N = Narcissism

EGA Narcissism Females 26 and Over



Note: N = Narcissism

There were several indicators that varied on the clusters across groups but overall, the clusters were mainly homogenous. With further refinement and removal of

items which hold similar face validity, not just for narcissism but all the constructs, a clearer picture will emerge.

Item Redundancy Analysis

To facilitate removing items that are redundant and further ascertain which indicators are homogenous across groups, item redundancy analysis was conducted. Target items were presented with potential indicators that were considered for redundancy as seen in Figure 28.

Figure 28

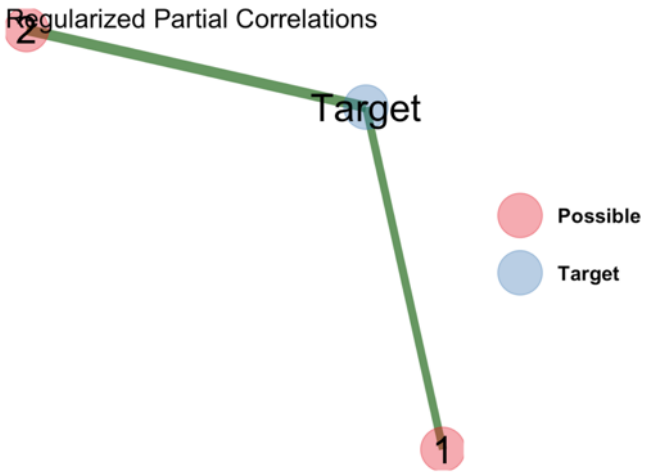
Item Redundancy

```
Target variable: 'M13'

Potential redundancies:

0. None
1. 'M34'
2. 'M39'

Enter numbers of variables redundant with the target variable (separate by commas)
Selection:
```

The semantic content of the suggested indicators were explored and a decision made as to which items were redundant. The example above shows item M13 ‘I get so emotional that I can’t think straight (R)’, from the Machiavellian item pool, which was compared with M34 ‘When I’m under a great deal of stress, sometimes I feel like I’m going to pieces (R)’ and M39 ‘People would describe me as emotionally stable’. The partial correlation plot was also examined (Figure 29). The edges between items denote redundancies determined to be statistically significant. The thickness of the edges correspond to the items connections in the network, mediated by the regularised partial correlations (Christensen, Golino & Silvia, 2020). The item M39 was retained for two reasons. Firstly, items M34 and M13 were reverse scored. Evidence has shown that reverse keyed indicators can negatively affect scales, whereby reliability and validity can be diminished. It has therefore been advised that reverse scored items should not be used in measures where samples may be heterogenous (Menold, 2020). The second reason to retain the item was that it aligns with the characteristics of the Machiavellian construct, with no semantic ambiguities.

Figure 29

Item Redundancy Regularised Partial Correlation Plot



This process continued for all indicators and consequently reduced the item pool for each construct. The remaining items were compared across groups to ascertain which items were homogenous. For Machiavellianism, 17 items remained from 53 (M1, M3, M4, M10, M11, M12, M14, M19, M20, M21, M23, M24, M26, M28, M29, M39, M40). For psychopathy there were 14 items (P2, P7, P9, P10, P12, P13, P14, P15, P17, P18, P22, P27, P30, P39), retained from the original 41. With narcissism, 27 out of 58 indicators were the same across groups (N1, N2, N3, N4, N5, N6, N7, N8, N9, N10, N11, N15, N22, N23, N25, N26, N28, N29, N31, N33, N35, N40, N43, N44, N52, N54, N58), and twelve indicators (S4, S8, S10, S12, S13, S15, S20, S24, S30, S32, S34, S36), from the original 39 items for sadism.

Exploratory Graph Analysis

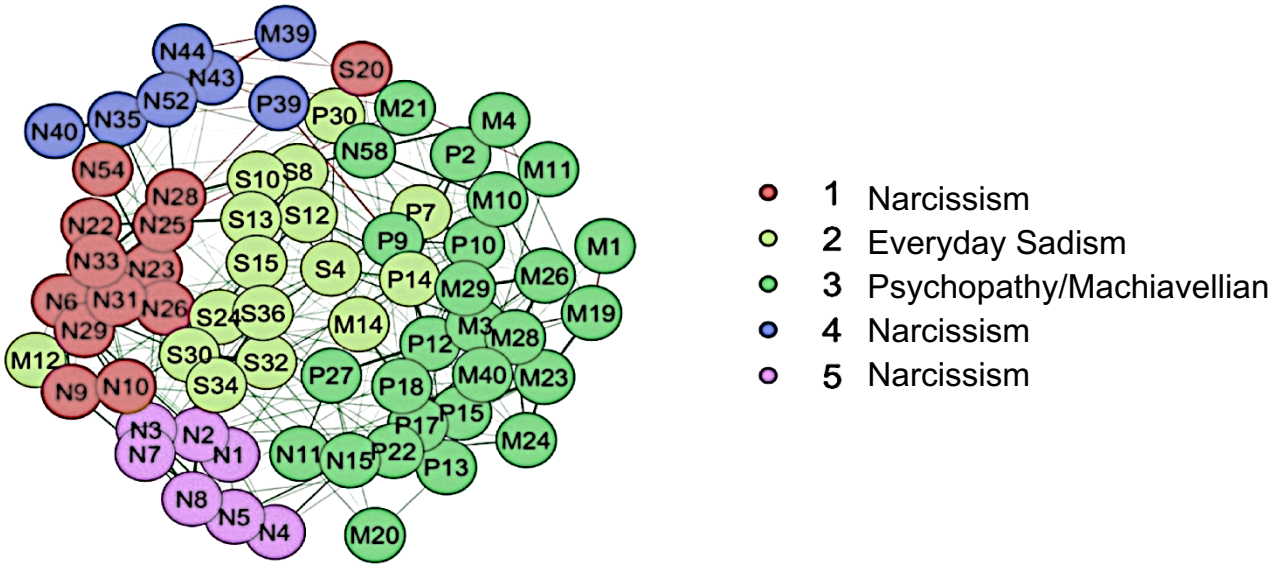
A further EGA, with the gamma set at 0.25 and the minimum lambda ratio at 0.1 was run to investigate the first model of the Dark Side of Humanity Scale with the four constructs. However, discrepancies still remained.

Males and Females 18-25

Although both groups displayed five clusters, as can be seen in Figure 30, items loadings varied across clusters for males and females. Machiavellian and psychopathy indicators mainly loaded on cluster three for males whilst for females this was found in cluster four.

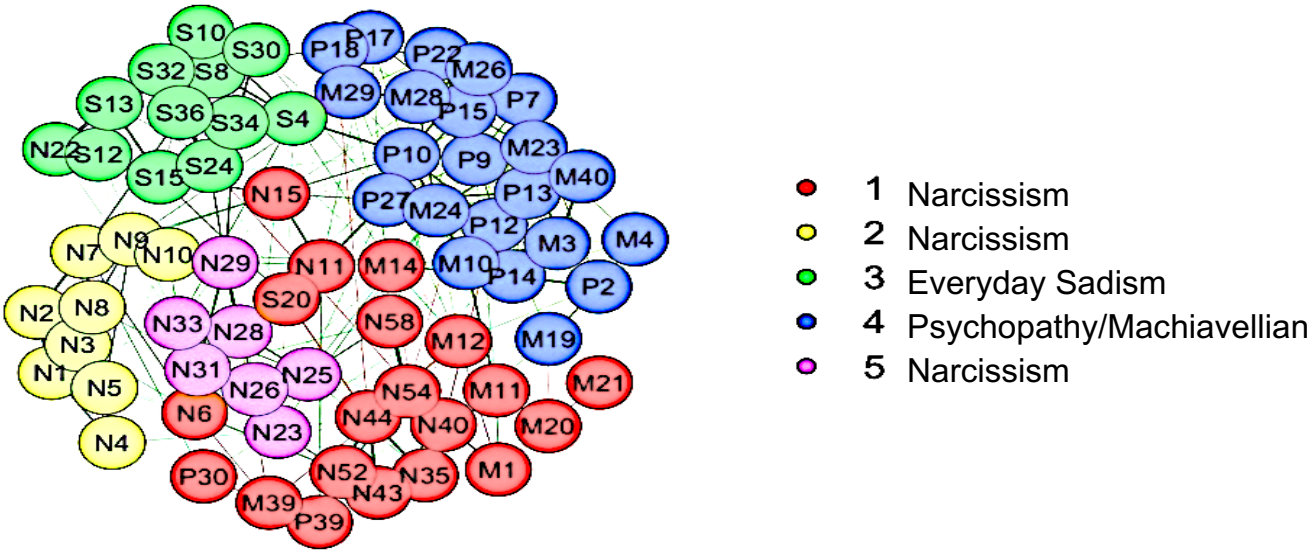
Figure 30

EGA Males 18-25 After Initial Item Reduction



Note: P = Psychopathy; M = Machiavellian; N = Narcissism; S = Everyday Sadism.

EGA Females 18-25 After Initial Item Reduction



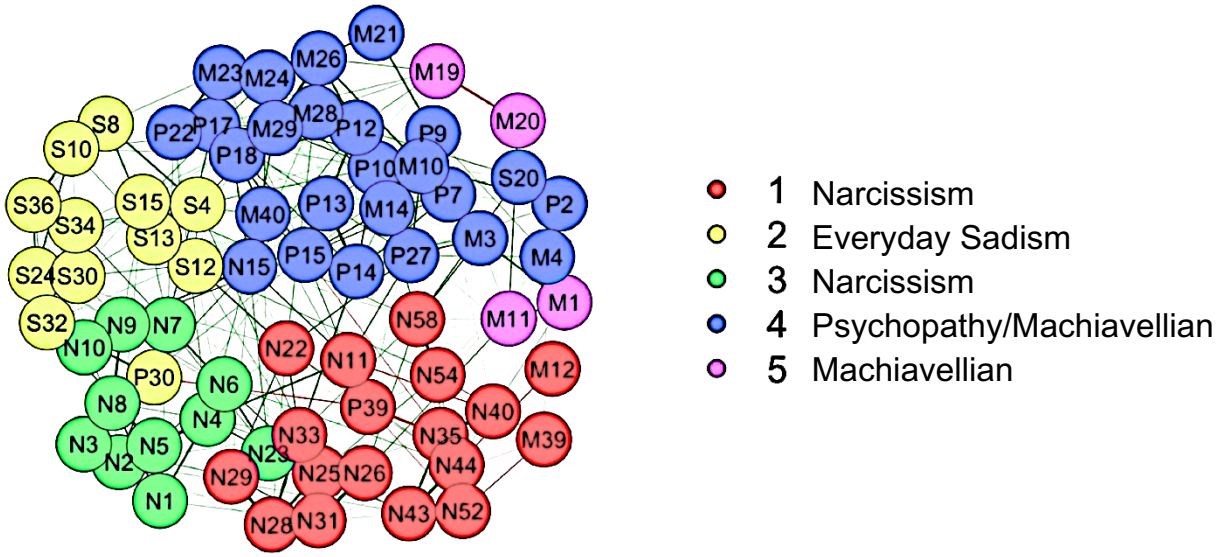
Note: P = Psychopathy; M = Machiavellian; N = Narcissism; S = Everyday Sadism.

Males and Females 26 and Over

Interestingly, both groups showed clusters for everyday sadism and psychopathy/Machiavellian as well as a separate cluster for Machiavellian, which contained four indicators. Narcissism, especially in the male group required further refinement (Figure 31).

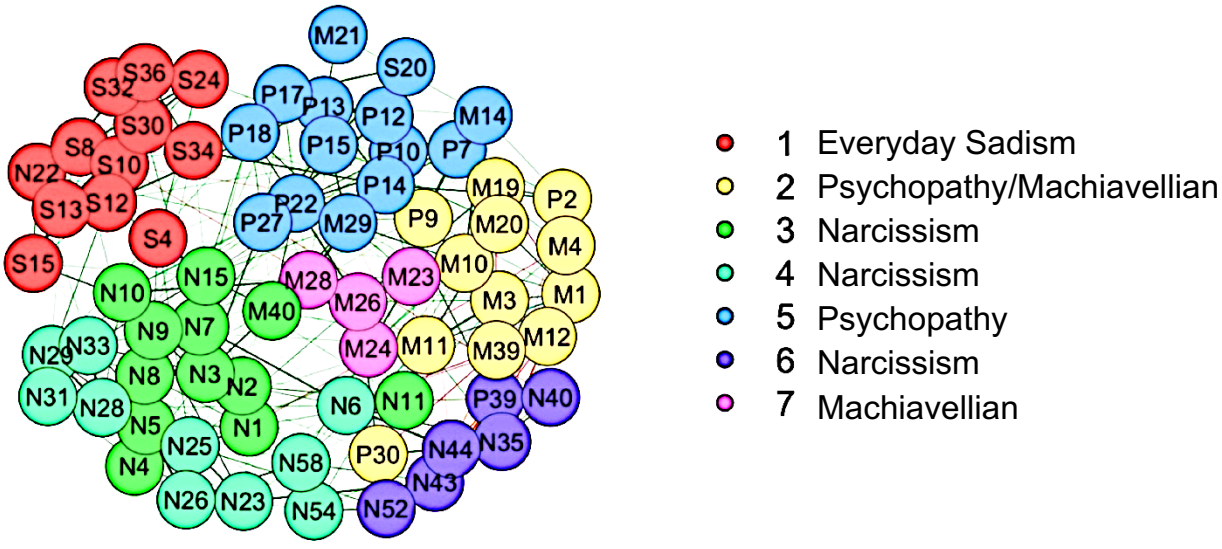
Figure 31

EGA Males 26 and Over After Initial Item Reduction



Note: P = Psychopathy; M = Machiavellian; N = Narcissism; S = Everyday Sadism.

EGA Females 26 and Over After Initial Item Reduction



Note: P = Psychopathy; M = Machiavellian; N = Narcissism; S = Everyday Sadism.

Everyday sadism clearly emerged as a cluster in all groups except the younger males where the cluster contained psychopathy and Machiavellian indicators. Entitlement rage was present in the female groups, yet for the males 18-25, entitlement and entitlement rage were clustered together, whilst for the older males, entitlement rage was clustered with devaluing and shame. The Machiavellian and psychopathy indicators were split over clusters and varied across groups.

Generalised Partial Credit Model

To explore the indicators in further depth, multidimensional item response theory (IRT), using the mirt package (Chalmers 2012), in the form of a Generalised Partial Credit Model (GPCM) was employed. This method can assist in the development of psychometrically robust measures of latent traits (Edelen & Reeve, 2007). The data for narcissism was separated into the grandiose and vulnerable facets, Machiavellianism and psychopathy were retained within their own factors as well as everyday sadism. Firstly the data was investigated for outliers with Mahalonobis distance which revealed; Machiavellianism (females 18-25 (2); males 18-25 (4); females 26 and over (2); males over 26 (4); psychopathy (females 18-25 (6); males 18-25 (4); females over 26 (11); males over 26 (8); everyday sadism (females 18-25 (20); males 18-25 (11); females 26 and over (22); males over 26 (13); grandiose narcissism (females 18-25 (5); males 18-25 (9); females 26 and over (10); males 26 and over (10); vulnerable narcissism (females 18-25 (2); males 18-25 (4); females 26 and over (6); males 26 and over (4)).

Outliers were retained, as they were in the first study, as their removal may comprise findings and diminish the accuracy of the final models, which seek to represent the investigated constructs (Aguinis et al., 2013; Neumann, Kosson & Salekin, 2017). Therefore, due to the nature of this study, the outliers, after consideration, were deemed to represent valid values in the data and were retained. Although outliers may impact analyses which relies on maximum likelihood estimators, the methods used in this study were not dependent on this (Aguinis et al., 2013). Finally, Mardias test for multivariate normality showed that none of the samples across the scales were normally distributed.

The slope parameters (*a)*, found in the GPCM. indicated how well the items discriminated across the different levels of the construct (Hussein, 2010). Values mainly ranged from moderate (0.65-1.34), to very high (>1.70), with the exception of items M3 for the over 26 groups and items P27 and N23 in males over 26 which were

low (Baker, 2001). The item step parameters, indicated by *b,* elucidated which items participants found easier or more difficult to endorse, dependent on the levels of the latent trait they embodied. For example, if at step one (*b*1), the value was negative or low, the probability was more likely that participants would move up a step to option two on the scale. However, if the *b* value for step 4 was higher than step 5, this indicated that participants would need higher levels of the latent construct to endorse the higher scale options (Embretson & Reise, 2000; Gomez, 2008)

Initial findings showed that the sadism item S36 (I enjoy making people suffer), for females 18-25 had to be re-mapped to ensure all categories had a distance of 1. The results showed an irregularly high *a* value (*a*= 7.59), therefore this item was removed from the analysis (Appendix 4.3.2). Item S30 (I enjoy mocking losers to their face), was also removed as it was too similar to item S32 (I get pleasure from mocking people in front of their face). Further inspection of the content of the indicators revealed that Item S34 (I never get tired of pushing people around), did not address the pleasure or power taken from harming others and was excluded, as was item S20 (Trolling behaviour is both cruel and unnecessary (R)), which had a very low *a* (females 18-25 *a*=0.26; males 18-25 *a*=0.21; females over 26 *a*=0.06; males over 26 *a*=0.02). For grandiose narcissism, N15 (If I have to take advantage of somebody to get what I want, so be it), was discarded as it was considered not to align with the study's definition of narcissism, there was also an overlap with Machiavellian items, for example M25 (I will break a promise if it works to my advantage). A further narcissism indicator, N11 (I can usually talk my way out of anything), was removed as the *a* values were consistently low across groups (females 18-25 *a*=0.54; males 18-25 *a*=0.49; females over 26 *a*=0.47; males over 26 *a*=0.49). Item N6 (I resent others who have what I lack), also displayed low values (females 18-25 *a*=0.63; males 18-25 *a*= 0.86; females over 26 *a*=0.53; males over 26 *a*=0.61).

In relation to vulnerable narcissism, the slope values for N22, N35, N40, N43, N44, N52, N54 and N58, throughout the groups were unacceptable, ranging from 0.05 to 0.57, and were discarded. Machiavellianism indicators M1, M4, M11, M12, M14, M19, M20, M21, M39 and M40 also held low slope values for males and females from both groups as did the psychopathy items P2, P30 and P39, which were removed from the analysis. The GPCM of the final models can be seen in Table 16 and 17. The item information curves, which show how well each item measures the constructs at differing levels, as well as the item response curves, which show the likelihood of

participants selecting a certain category on the scale can be found in Appendix 4.3.4 and 4.3.5.

Table 16

Dark Side of Humanity Scale Slope Parameters and Item Locations for Males and Females 18-25

Item	*a* M18	*b1*	*b2*	*b3*	*b4*	*b5*	*a* F18	*b1*	*b2*	*b3*	*b4*	*b5*
P7	1.03	-0.68	0.35	0.81	1.25	3.99	1.07	-0.23	0.89	1.37	2.08	
P9	0.78	-1.21	0.52	-0.20	1.81	2.37	0.77	-0.47	0.64	0.31	1.69	2.32
P10	1.52	-0.51	0.71	0.78	1.66	2.99	2.00	0.18	1.05	1.37	1.57	2.30
P12	1.39	-0.56	0.09	0.64	1.96	0.93	1.48	-0.02	1.27	0.54	2.22	1.53
P13	0.97	-0.33	-0.24	0.46	1.55	1.62	0.91	0.31	0.70	0.59	2.48	1.90
P14	1.15	0.10	0.78	0.43	1.55	2.26	1.30	0.37	1.28	1.03	1.65	1.85
P15	1.25	-0.94	0.08	0.38	1.38	1.71	1.29	-0.76	1.06	0.68	2.07	2.12
P17	2.09	-0.47	-0.07	0.45	1.53	1.91	1.40	0.01	0.60	1.20	2.21	
P18	1.76	-0.21	-0.16	0.32	1.24	1.46	1.24	0.20	0.70	1.18	2.56	
P22	1.12	-1.68	-0.56	0.44	1.35	1.99	1.14	-1.03	0.63	1.05	1.80	2.43
P27	0.78	-0.06	0.38	-0.13	1.68	2.18	0.85	0.32	0.65	0.42	1.97	2.15
M3	0.66	-1.01	-0.22	-0.60	1.81	2.29	0.76	-0.98	0.66	0.07	2.98	1.83
M10	0.69	0.03	0.89	0.79	1.74	5.14	1.18	0.10	1.48	1.10	2.50	2.46
M23	1.16	-1.71	0.05	-0.25	1.69	2.03	1.23	-0.92	0.52	0.42	1.99	
M24	0.72	-1.89	-1.30	0.42	1.52	2.41	0.83	-1.45	-0.11	0.55	1.96	2.86
M26	1.24	-0.60	0.12	0.61	1.69	2.19	1.14	-0.59	0.66	0.93	2.66	2.62
M28	1.38	-0.54	-0.25	0.34	1.45	1.69	1.97	-0.50	0.58	1.32	1.89	
M29	1.86	-0.44	0.09	0.65	1.89	2.33	1.53	0.50	0.85	1.40	2.07	
N1	1.81	-0.03	0.48	1.12	1.67	1.81	2.65	0.27	0.80	1.00	2.28	
N2	3.22	0.07	0.71	0.83	1.55	1.93	4.75	0.22	0.66	1.19	1.65	
N3	2.74	0.04	0.51	0.88	1.38	2.26	3.67	0.31	0.85	1.06	2.09	
N4	0.90	-0.44	0.28	0.62	1.68	1.50	0.90	-0.11	-0.11	0.41	2.31	3.69
N5	1.46	0.14	0.32	0.88	1.48	1.36	1.98	0.43	0.70	1.01	1.72	
N7	1.32	0.20	0.57	0.74	1.63	1.71	2.26	0.42	1.00	1.40	2.55	
N8	1.54	-0.13	0.33	0.92	1.52	2.05	2.41	0.20	0.56	1.10	2.13	
N9	0.84	0.20	1.09	0.63	2.55	1.45	1.24	0.72	1.30	0.86	1.63	
N10	1.35	0.29	0.80	1.13	1.69	1.04	1.54	0.93	1.14	1.04	2.08	
S4	0.91	0.54	0.62	0.61	2.00	2.53	1.18	1.15	0.99	1.20	2.94	2.34
S8	2.16	-0.04	0.63	1.16	1.52	1.98	4.22	0.63	1.25	1.35	1.75	1.97
S10	2.08	0.30	0.90	0.90	1.33	1.94	5.04	0.75	1.17	1.50	1.65	2.35
S12	1.02	0.61	0.99	0.12	1.71	1.14	1.08	0.85	1.27	0.63	2.79	1.20
S13	1.96	0.43	0.88	0.58	1.78	2.22	2.48	0.73	1.32	1.20	2.06	1.59
S15	1.66	0.47	0.45	0.93	2.13	1.05	1.63	0.94	1.33	1.52	2.67	1.40
S24	1.38	1.09	0.78	2.17	2.69	2.69	1.35	1.79	1.92	1.09	1.73	
S32	1.52	0.94	0.69	1.03	1.90	2.23	2.38	0.98	1.58	2.31	1.31	
N23	0.69	-0.82	-0.74	0.12	1.59	1.54	0.73	-0.87	0.05	-1.31	1.92	2.39
N25	1.46	-1.03	-0.05	0.29	1.89	1.96	1.07	-0.79	0.21	0.32	2.16	3.47
N26	1.06	-1.49	-0.29	-0.15	1.38	1.88	1.96	-0.87	0.02	0.16	1.40	1.92
N28	1.32	-0.22	0.36	0.66	1.58	1.77	1.39	0.24	0.73	0.43	1.44	3.33
N29	0.99	0.08	0.36	0.91	2.08	1.69	1.23	0.52	0.99	1.41	2.78	2.24
N31	1.94	-0.53	0.24	0.52	1.81	1.62	1.82	-0.10	0.56	0.60	1.80	
N33	1.50	-0.15	0.80	0.79	1.67	1.91	1.96	0.30	1.01	1.06	1.89	2.58

Note: a = Slope Parameters; *b* = item location; M18 = Males 18-25; F18 = Females 18-25.

Table 17

Dark Side of Humanity Scale Slope Parameters and Item Locations for Males and Females 26 and Over

Item	*a* M26	*b1*	*b2*	*b3*	*b4*	*b5*	*a* F26	*b1*	*b2*	*b3*	*b4*	*b5*
P7	0.77	-0.96	0.80	0.48	1.84	2.56	0.91	0.27	1.17	1.61	3.28	1.54
P9	0.75	-1.04	-0.06	0.28	1.85	1.57	0.92	-0.44	0.68	0.96	2.41	2.75
P10	0.90	-0.06	0.62	1.12	1.72	2.52	1.69	0.65	1.52	2.07	2.53	2.00
P12	1.16	-0.30	0.60	0.80	1.26	2.66	1.09	1.03	1.56	1.51	2.86	
P13	1.26	-0.11	0.14	0.68	1.61	2.81	1.02	1.01	1.39	0.91	2.98	
P14	1.32	0.42	0.68	1.01	2.07		1.21	1.25	1.75	1.51	3.18	1.85
P15	1.09	-0.24	0.28	0.36	1.48	2.44	1.34	0.31	0.91	1.66	2.81	
P17	1.81	-0.25	0.63	0.59	1.83	2.80	1.91	0.42	1.34	1.29	2.57	
P18	2.40	-0.19	0.32	0.61	1.24	2.78	1.56	0.76	0.94	2.02	1.82	
P22	1.11	-0.93	0.06	0.08	1.53	2.49	1.52	-0.16	0.86	1.15	2.48	1.67
P27	0.55	0.25	0.69	0.29	2.06	1.67	0.96	0.60	1.18	1.20	2.59	2.25
M3	0.55	-0.68	0.68	0.09	3.25	1.24	0.47	0.27	1.67	0.58	3.88	
M10	0.82	0.03	0.95	0.96	1.73	4.22	0.64	0.65	2.40	2.62	1.54	3.68
M23	0.79	-1.32	0.72	-0.74	2.36	2.52	1.01	-0.13	1.30	0.92	3.64	
M24	0.99	-1.25	-0.49	0.52	2.11	2.42	0.78	-0.79	0.51	0.41	2.51	3.34
M26	1.62	-0.68	0.40	0.60	1.81	2.58	1.83	-0.42	0.68	1.08	2.22	
M28	1.15	-0.51	0.40	0.78	2.73	1.80	2.98	-0.04	0.65	1.08	2.31	
M29	2.07	-0.15	0.84	1.06	2.10	1.63	1.40	0.53	1.76	1.35	3.22	2.02
N1	1.81	0.15	0.71	0.96	1.50	2.44	2.26	0.31	0.94	1.11	2.47	
N2	3.86	0.22	0.67	1.09	1.77	1.82	2.85	0.62	1.04	1.20	2.28	
N3	2.41	0.42	0.72	0.77	2.10	0.66	5.54	0.54	1.12	1.26	2.01	
N4	1.16	-0.04	0.35	0.28	1.34		1.19	0.10	0.33	-0.05	2.04	1.57
N5	1.58	0.53	0.57	1.13	1.58	2.91	1.50	0.74	0.64	1.55	1.91	
N7	0.96	0.96	0.83	1.01	3.25	3.41	1.69	0.98	1.39	1.41	2.27	2.27
N8	1.85	0.30	0.68	0.38	1.99		2.06	0.48	0.89	1.02	1.78	
N9	1.10	0.80	0.45	0.97	2.38		1.42	0.81	1.29	1.41	2.51	2.26
N10	0.87	1.11	1.35	0.26	2.44		0.90	1.48	2.23	0.85	2.48	2.79
S4	1.02	0.91	1.20	0.63	1.81	2.63	0.79	1.77	1.98	1.67	3.04	
S8	1.66	0.54	1.27	1.14	2.14	2.16	2.49	1.37	1.99	1.89	2.47	
S10	1.59	0.97	1.38	1.03	2.14	2.19	2.91	1.62	1.92	2.37	2.13	
S12	1.26	0.71	1.07	0.32	1.69	2.23	1.53	1.39	1.96	1.04	3.16	1.95
S13	2.83	0.63	1.09	0.99	1.84	2.08	3.47	1.55	2.02	1.64	2.47	
S15	1.86	1.05	1.19	1.23	1.94	1.74	1.17	2.31	1.62	2.22	2.74	
S24	1.68	1.32	1.71	1.66	1.44	2.24	1.25	2.74	2.97	1.11	2.48	
S32	1.00	1.76	2.23	0.33	2.86	2.10	1.32	2.57	3.02	1.72	1.98	
N23	0.49	-0.37	0.38	-0.43	2.71	3.37	0.92	0.09	0.81	-0.81	2.05	2.93
N25	1.01	-0.49	0.50	0.72	2.15	2.65	1.71	0.03	0.40	0.54	2.24	
N26	1.08	-0.17	-0.12	0.18	1.89	1.79	1.34	-0.19	0.56	0.30	1.74	1.94
N28	2.34	0.11	0.77	0.82	1.54	2.58	1.76	0.45	1.22	0.98	1.79	2.24
N29	1.06	0.78	0.75	1.37	2.07	2.99	1.95	0.73	1.24	1.67	1.88	2.39
N31	2.35	0.10	0.50	0.95	1.50		2.52	0.47	0.90	0.74	1.91	
N33	1.82	0.39	0.79	1.64	1.77	2.49	1.65	0.79	1.50	0.95	2.58	2.18

Note: a = Slope Parameters; *b* = item location; M26 = Males 26 and Over; F26 = Females 26 and Over

The two tables serve to highlight a limitation of classical test theories (CTT's), which assume that measurement precision is constant across the spectrum of the latent construct. Yet, it can be seen across groups in the GPCM, that this is not constant across items. Indeed, scales developed solely through CTT's are more likely to have an unequal dissemination of accuracy across the normal range of the construct and do not indicate the degree of the latent trait the items tap into, whereas the GPCM elucidates the range. For example, if the GPCM showed the step values gathered at the higher end of the continuum, the scale would only offer accurate measurement for those who embody high levels of the construct and inaccurate assessment of those with lower levels (Fraley et al., 2000). The scale has items which reflect all levels of the latent constructs being measured. A further strength of this method reveals that for certain items, not all scale options are being utilised, a variance which is elusive in an EFA. The GPCM substantially reduced the item pool and culminated in the remaining items being homogenous across samples, with each item falling within the accepted slope values as well as addressing core facets of the constructs.

Exploratory Graph Analysis

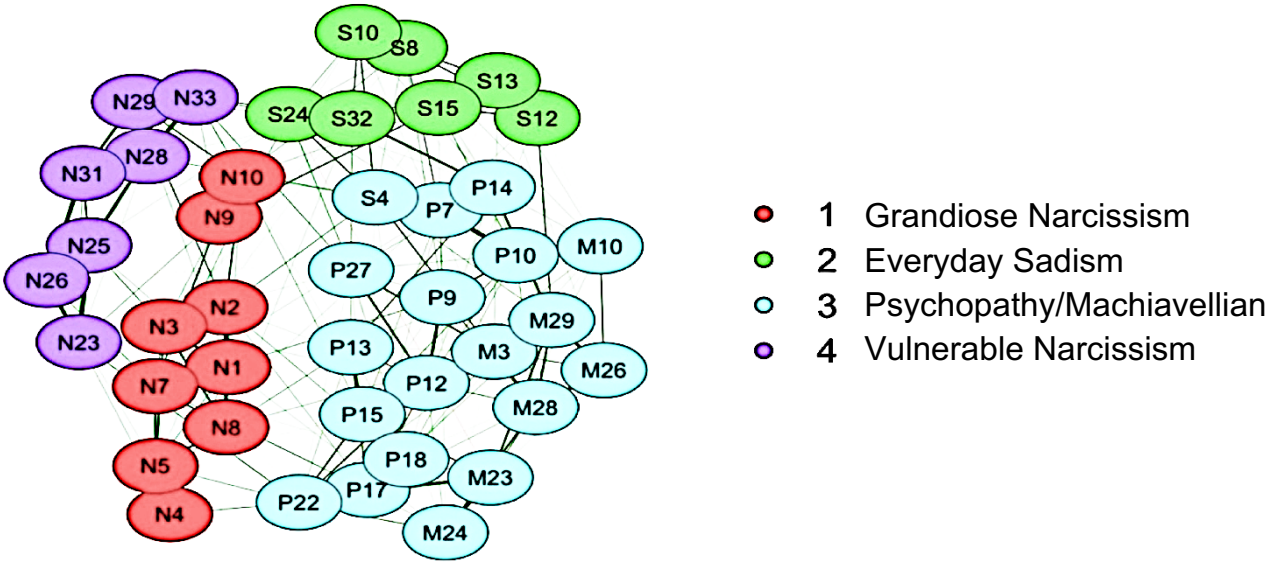
Following the second item reduction, a further EGA was conducted for the Dark Side of Humanity Scale (DSHS).

Males and Females 18-25

The networks showed slight variations across groups, whereby items P17 and P18, in the female 18-25 group, were separate from the remaining psychopathy items, however, the remaining psychopathy as well as Machiavellian indicators formed one cluster. For males 18-25, the everyday sadism item S4 (It give me pleasure to see someone successful get fired), clustered onto the dimension containing the psychopathy and Machiavellian items (Figure 32),

Figure 32

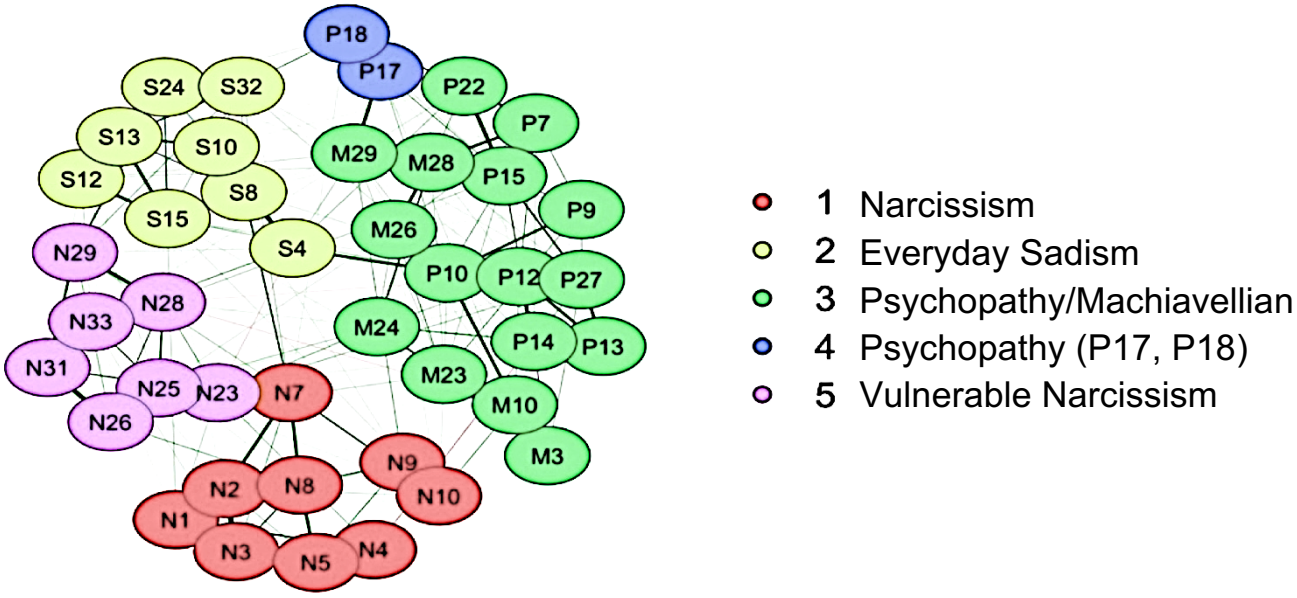
EGA Males 18-25 Final Network



Note: N = Narcissism; P = Psychopathy; M = Machiavellian; S = Everyday Sadism

.

EGA Females 18-25 Final Network



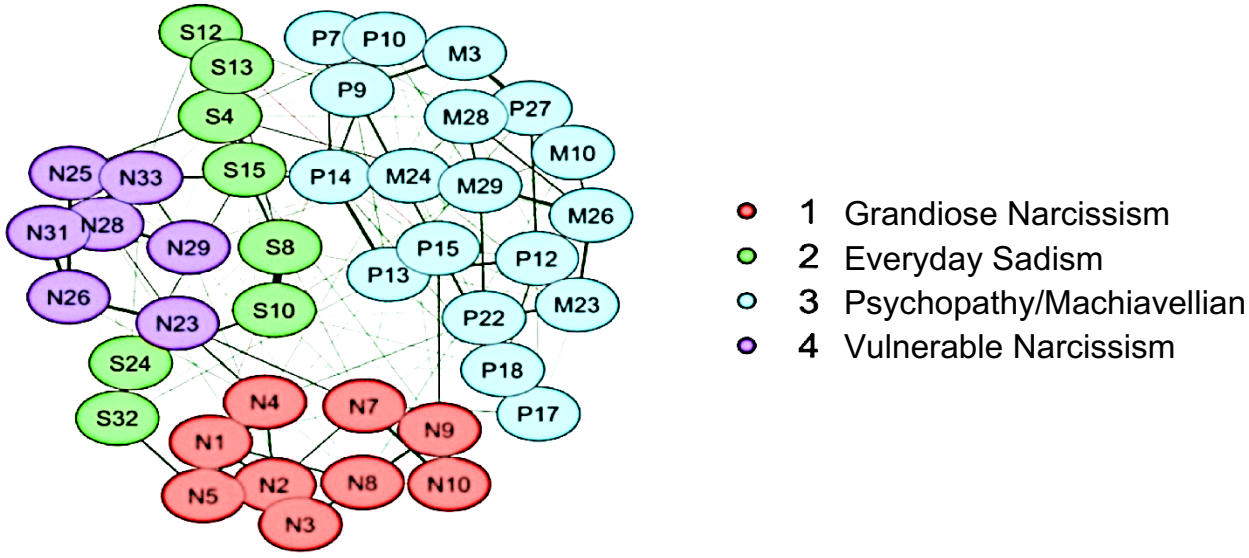
Note: N = Narcissism; P = Psychopathy; M = Machiavellian; S = Everyday Sadism

Males and Females 26 and Over

The females over 26 group showed that Machiavellian items formed their own cluster, with only one indicator loading onto the psychopathy cluster (M29). All items loaded onto their relevant clusters for Males Over 26, with all Machiavellian and psychopathy items formed one cluster (Figure 33)

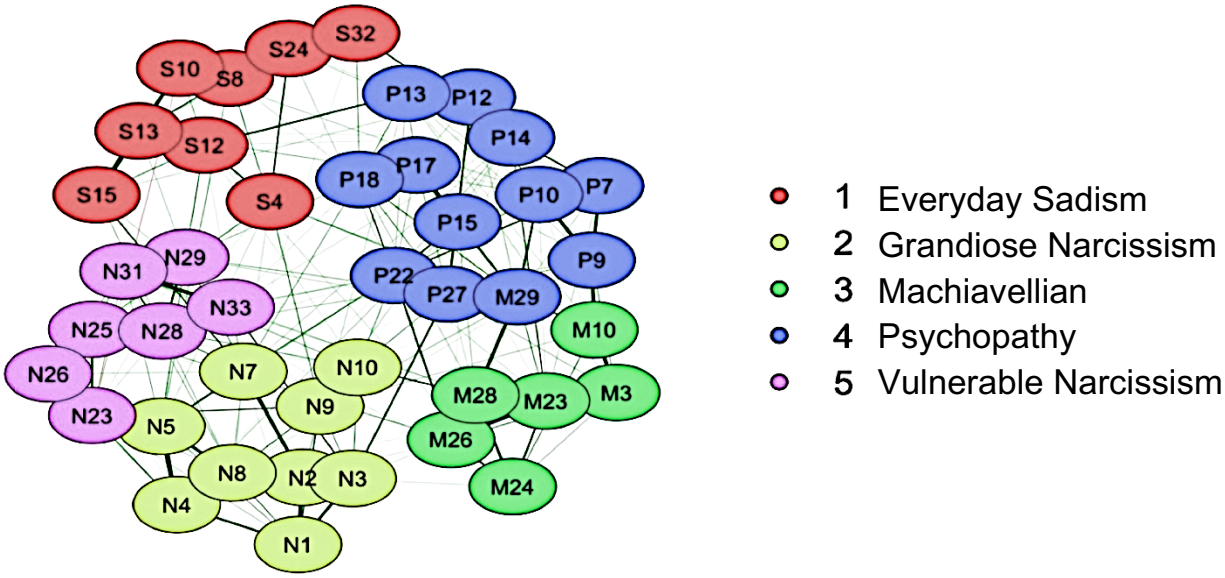
Figure 33

Males 26 and Over Final Network



Note: N = Narcissism; P = Psychopathy; M = Machiavellian; S = Everyday Sadism

Females 26 and Over Final Network



Note: N = Narcissism; P = Psychopathy; M = Machiavellian; S = Everyday Sadism

Network Invariance

To ascertain whether the anomalies found in the networks across groups affected the invariance between males, females and age groups an invariance analysis was conducted. Network structure invariance refers to the similarity of edge weight distributions, whilst network global strength invariance pertains to the overall level of connectivity. M is the maximum difference in edge weights and S is the difference in global strength. A statistically non-significant difference (p >.05), between the networks suggests invariance and therefore the network is replicable (Dodell-Feder et al., 2019). The analysis was conducted in the R package NetworkComparisonTest (van Borkulo et al., 2017). The results showed invariance for; males and females 18-25 (M = 0.15, *p*=0.57); males and females over 26 (M=0.17, *p*=0.42); females 18-25 and females 26 and over (M=0.26, *p*=0.05); females 18-25 and males over 26 (M=0.17, *p* = 0.51); males 18-25 and males over 26 (M=0.11, *p* = 0.80); males 18-25 and females over 26 (M= 0.24, *p* = 0.06). However, global strength invariance was not attained for males and females 18-25 (S=0.15, *p* =0.02), males 18-25 and males over 26 (S= 0.14, *p* = 0.04), males 18-25 and females over 26 (S=0.14, *p* =0.02). Global strength invariance was found for females 18-25 and males over 26 (S= 0.04, *p* = 0.62), females 18-25 and females 26 and over (S= 0.01, *p* = 0.92) and males and females over 26 (S= 0.3, *p*=0.7; Appendix 4.4).

The final data was investigated for outliers using Mahalonobis distance the following were found: (Machiavellian: females 18-25 (2); males 18-25 (4); females over 26 (1); males over 26 (2). Psychopathy; females 18-25 (5); males 18-25 (3); females over 26 (3); males over 26 (2); Grandiose narcissism; females 18-25 (4); males 18-25 (5), females over 26 (7); males over 26 (7); Everyday sadism, females 18-25 (6); males 18-25 (6); females over 26 (17); males over 26 (4); Vulnerable narcissism; (females 18-25 (2); males 18-25 (1); females over 26 (4); males over 26 (2). As previously discussed, the outliers were retained as they were deemed to represent valid values in the data.

Exploratory Graph Confirmatory Factor Analysis

A CFA was then conducted using the EGA model with the WLSMV estimator for each group, which unlike a specified model with a traditional CFA, the model emerges from the EGA. For clarity Figure 34, shows the R code used, where ‘ega.F18DSHS’, is the model used from the female 18-25 EGA analysis.

Figure 34

R Code for EGA Confirmatory Factor Analysis.

```
> cfaF18.net <- CFA(ega.F18DSHS, estimator = "WLSMV", data = F18DSHSN)
```

Traditional CFA scores are generally calculated using a simple structure with regression methods, whereby items only load on one factor. Network scores, however, are computed using a complex structure based on a weighted composite rather than a latent factor (Christensen & Golino, 2020). The models corroborated the EFA graphs whereby for males 18-25, all indicators loaded onto the four clusters of Machiavellianism/psychopathy, grandiose narcissism, vulnerable narcissism and everyday sadism, with a good model fit (χ^2 = 906.60, *df*=813, *p*=.012, CFI= 0.97, RMSEA=0.03). For females 18-25, the two psychopathy indicators P16 and P17 loaded onto their own factor, with all remaining psychopathy and Machiavellian items loading onto one factor, with fit measures for the model (χ^2 = 923.55, *df*=809, *p*=.003, CFI= 0.95, RMSEA=0.03). Females over 26 displayed a five-factor structure, with only one Machiavellian indicator loading with psychopathy, with fit measures of (χ^2 = 932.03, *df*=809, *p*=.002, CFI= 0.93, RMSEA=0.03). Finally, males over 26 showed a four-factor structure, with all Machiavellian and psychopathy items loading onto one factor (χ^2 = 966.96, df=813, *p*=.00, CFI= 0.94, RMSEA=0.03; Appendix 4.5).

Confirmatory Factor Analysis and Measurement Invariance of the DSHS

As many studies use classical test theories, such as regression to analyse data (e.g., Pfattheicher, Keller & Knezevic, 2018), it is critical to ensure, through traditional CFA analysis, that the model provides a good fit. In line with the EGA models, as psychopathy and Machiavellian indicators were subsumed under one cluster for three out of the four groups and congruent with past research, which suggests the two constructs are parallel (McHoskey et al., 1995; Miller, Hyatt et al., 2017; Persson, 2019; Persson, Kajonius & Garcia, 2017b), the CFA model across groups was determined by four factors with the first addressing psychopathy/Machiavellian, the second, grandiose narcissism, the third, everyday sadism and the fourth pertaining to vulnerable narcissism. Thus, a four-factor structure was investigated (Table 18; Appendix 4.6).

Table 18

Confirmatory Factor Analyses for the DSHS using WLSMV and Robust Values

Models	*n*	χ^2	*p*	*df*	CFI	TLI	RMSEA	90% Confidence Lower/Upper
Males 18-25	175	894.71	0.024	813	0.97	0.97	0.02	0.01-0.03
Females 18-25	157	921.41	0.005	813	0.94	0.93	0.03	0.02-0.04
Males over 26	158	953.98	<.001	813	0.93	0.93	0.03	0.02-0.04
Females over 26	177	886.54	0.037	813	0.93	0.93	0.02	0.01-0.03
Whole Sample	667	1558.97	<.001	813	0.91	0.91	0.04	0.03-0.04

Note:: χ^2= Chi Square; df= degrees of freedom; CFI=Comparative fit index: TLI=Tucker-Lewis Index: RMSEA=Root mean square error of approximation. All χ^2 statistics were significant at $p < .001$. DSHS=Dark Side of Humanity Scale.

A good fit was found across samples, without modifications. Although the female groups in the EGA and CFA of the network analyses showed a divergence from the four-factor structure at an item level basis, the fit for the four-factor model with the traditional CFA highlighted the differences between regularised partial correlation analysis for complex networks and latent trait modelling, which seeks to explain the covariation amongst a set of indicators. An analysis was also conducted on the whole sample to ensure that the DSHS could be used with mixed age and sex samples, whereby the CFA showed a good model fit.

Measurement Invariance of the Dark Side of Humanity Scale

The CFA results provided the basis to investigate measurement invariance of all samples for the four-factor model. The whole sample was also included to ensure that when samples are not separated by age or sex, the scale still holds invariance for mixed age and sex participants. Initially, the configural invariance model, with unconstrained factor loadings was calculated which served as the baseline for metric invariance, which tests for equal factor loadings, followed by scalar invariance, which investigates equal factor loadings and item thresholds, whereby intergroup mean scores can meaningfully be compared (Meade et al., 2008; Orri et al., 2018; Table 19; Appendix 4.7).

Table 19

Measurement Invariance for the Four-Factor DSHS using WLSMV and Robust Values

Models	χ^2	*p*	*df*	CFI	TLI	RMSEA	ΔCFI	ΔRMSEA
Four Factor								
M F 18-25								
Configural	1817.40	<.001	1626	0.959	0.957	0.027		
Metric	1826.41	<.001	1664	0.965	0.964	0.024	0.006	-0.003
Scalar	1869.30	.003	1702	0.964	0.964	0.024	-0.001	0.000
MF 26 Over								
Configural	1830.97	<.001	1626	0.932	0.928	0.028		
Metric	1831.86	.002	1664	0.940	0.943	0.025	0.008	-0.003
Scalar	1877.02	.002	1702	0.942	0.942	0.025	0.002	0.000
Whole Sample								
Configural	2204.95	<.001	1626	0.930	0.926	0.033		
Metric	2165.39	<.001	1664	0.939	0.937	0.030	0.009	-0.003
Scalar	2216.00	<.001	1702	0.938	0.937	0.030	-0.001	0.000

Note: df=degrees of freedom; CFI = Confirmatory Fit index; TLI =Tucker-Lewis Index; RMSEA = Root Mean Square of Approximation; DSHS=Dark Side of Humanity Scale.

The four-factor model evidenced scalar invariance across all groups. This model can therefore be used with confidence, whereby multi-group comparisons of factor means and statistically significant differences in group means would not be due to differences in scale properties across ages and sexes (Bialosiewicz et al., 2013).

The Factors of the Dark Side of Humanity Scale

With the four-factor model confirmed through analyses, the semantic context of the factors were considered in further depth. The first factor incorporates both Machiavellian and primary psychopathy indicators which address, strategic planning, callousness, misanthropy, deceit, manipulation, immorality and egocentricity. As discussed in the introduction of this chapter, the characteristics of both constructs pertain to the successful psychopath (Palmen et al., 2018; Poythress and Hall, 2011). Initially, this factor was going to be named ‘Callous Misanthropy’, yet on reflection of the constellation of traits it addresses, it was felt the title did not capture the essence of the construct. Although the successful psychopath has been discussed over decades in the literature (Babiak, 1995; Cooke, Michie, Hart & Clark, 2004; Hall & Benning 2006; Mullins-Sweatt et al., 2010; Persson & Lilienfeld, 2019; Widom, 1977), it has not been specifically incorporated within dark personality measures. Thus, it was decided to label the factor as ‘Successful Psychopathy’, which leaves no ambiguity as to what the factor measures.

The second factor aligns with the grandiose form of narcissistic entitlement whilst the fourth factor pertains to the vulnerable form of entitlement – entitlement rage. The phenotypic structure of narcissistic traits have been shown to comprise the core of entitlement, which is expressed through grandiosity and/or vulnerability (Ackerman, Donnellan & Wright, 2019; Dickinson & Pincus, 2003; Rosenthal & Hooley, 2010). This has been corroborated through network analysis, whereby entitlement acts as the bridge which joins the grandiose and vulnerable facets of narcissism (Dinić, Sokolvksa & Tomašević, 2021). Embodying grandiose entitlement reflects an individual who holds an inflated opinion of their abilities and characteristics, with attitudes alluding to the right to receive special benefits and attention, mediated by a sense of superiority (Freis & Hansen-Brown, 2021). Conversely, entitlement rage reveals individuals who are hypersensitive, anxious with low self-esteem, which mediates emotionally volatile reactions to criticism, expressed through anger and aggression (Donnellan, Ackerman & Wright, 2021). Thus, factor two was labelled Grandiose Entitlement and factor four as Entitlement Rage.

Finally, the third factor pertained to everyday sadism. The indicators imply that both power and pleasure mediate sadistic behaviours (O'Meara et al., 2011; Paulhus & Dutton, 2015). Thus, acts of everyday sadism may provide the individual with positive affect through different motivations, whereby power within interpersonal situations can be sustained through sadistic behaviours from which pleasure is derived. Alternatively, direct, vicarious and/or sadistic fantasies can be realised for no other reason than to obtain pleasure from the suffering of others (Paulhus & Dutton, 2015; Taylor, 2009). The factor was initially going to be labelled 'sadistic pleasure', yet this demeans the cruelty involved, regardless of the motivation. An act is determined as cruel if the perpetrator experiences positive arousal from the victim's pain and can be viewed on a continuum from schadenfreude to direct physical or psychological harm (Nell, 2006). Consequently, factor four was labelled 'Sadistic Cruelty'.

The Dark Side of Humanity Scale

Note: Item numbers in brackets are for ease of reference in relation to the analysis.

Factor One: Successful Psychopathy

1. It's wise to keep track of information that I can use against people (M3).
2. What other people feel doesn't concern me (P7).
3. The only good reason I talk to others is to get information that I can use to my benefit (M10).
4. I can be good at pretending to care about people but must of the time I really don't care (P9).
5. It's sometimes fun to see how far I can push someone before they catch on (P13).
6. Success is based on survival of the fittest, I am not concerned about the losers (P15).
7. I could look people straight in the eye and it means nothing to me to lie or to cheat them (P12).
8. I can simulate emotions like pain and hurt to make others feel sorry for me (P27).
9. I am willing to be unethical if I believe it will help my plans succeed (M23).
10. For me, what's right is whatever I can get away with (P17).
11. I believe that lying is necessary to maintain a competitive advantage over others (M28).
12. If I'm honest all the time it won't lead to the success of my objectives (M24).
13. I am willing to sabotage the efforts of other people if they threaten my own goals (M29).
14. I don't care much if what I do hurts others (P10).
15. In today's world, I feel justified in doing anything I can get away with (P18).
16. Playing by the rules sounds nice but getting what I want is more important (P22).
17. I will break a promise if it works to my advantage (M26).
18. People might describe me as mean and cruel (P14).

Factor Two: Grandiose Entitlement

1. I deserve to receive special treatment (N1).
2. I expect people to bend the rules for me (N2).
3. I tend to expect special favours from others (N3).
4. I deserve to get what I want (N4).
5. I deserve more out of life than other people (N5).
6. I don't think the rules apply to me as much as they apply to others (N7).
7. I expect to be treated better than average (N8).
8. I only associate with people of my calibre (N9).
9. I do not waste my time hanging out with people who are beneath me (N10).

Factor Three: Sadistic Cruelty

1. It gives me pleasure to see someone successful get fired (S4).
2. I enjoy seeing people hurt (S8).
3. I have fantasies which involve hurting other people (S12).
4. Hurting people would be exciting (S13).
5. I post offensive comments on social media forums just so I can take pleasure from the hurt I cause (S24).
6. I would hurt somebody if it meant that I would be in control (S15).
7. I get pleasure from mocking people in front of their friends (S32).
8. I enjoy watching people in pain (S10).

Factor Four: Entitlement Rage

1. It irritates me when people don't notice how good I am (N23).
2. I get into a temper if I don't get the recognition that I deserve (N31).
3. I can get pretty angry when others disagree with me (N25).
4. I hate being criticised so much that I can't control my temper when it happens (N28).
5. It really makes me angry when I don't get what I deserve (N26).
6. I fly into a rage if somebody expects me to do tasks that are really beneath my skill level (N29).
7. I can get really nasty if I don't get what I want (N33).

Descriptive Statistics Dark Side of Humanity Scale

The mean scores for each factor of the DSHS as well as the Cronbach's alpha and McDonald's Omega reliability was investigated with Jamovi (v1.6.1.0). As can be seen, males scored a higher mean score on all factors, with the males 18-25 scoring the highest means across age and sex. Reliability was good for all factors across all groups (Table 20).

Table 20

Mean scores, reliability and inter-factor correlations for the Dark Side of Humanity Scale

Sample	Mean (SE)	Skewness (SE)	Kurtosis (SE)	F1	F2	F3	F4	α	**ω**
Males 18-25									
SP	50.33 (1.33)	0.22 (0.18)	-0.58 (0.37)	-				0.94	0.94
GE	20.56 (0.75)	0.77 (0.18)	-0.24 (0.37)	0.69	-			0.93	0.93
SC	16.55 (0.65)	0.92 (0.18)	-0.08 (0.37)	0.77	0.69	-		0.91	0.91
ER	19.10 (0.56)	0.21 (0.18)	-0.76 (0.37)	0.54	0.69	0.57	-	0.88	0.88
Total Scale								0.97	0.97
Females 18-25									
SP	40.62 (1.17)	0.44 (0.19)	-0.68 (0.38)	-				0.93	0.93
GE	17.51 (0.67)	0.81 (0.19)	-0.46 (0.38)	0.49	-			0.93	0.93
SC	12.54 (0.52)	2.05 (0.19)	4.64 (0.38)	0.62	0.54	-		0.91	0.92
ER	17.13 (0.56)	0.42 (0.19)	-0.74 (0.38)	0.48	0.54	0.51	-	0.88	0.89
Total Scale								0.96	0.96
Males Over 26									
SP	45.68 (1.31)	0.48 (0.19)	-0.28 (0.38)	-				0.93	0.93
GE	18.64 (0.71)	0.76 (0.19)	-0.36 (0.38)	0.60	-			0.92	0.92
SC	13.58 (0.53)	1.34 (0.19)	1.10 (0.38)	0.67	0.57	-		0.87	0.88
ER	16.01 (0.53)	0.57 (0.19)	-0.27 (0.38)	0.47	0.61	0.45	-	0.86	0.87
Total Scale								0.96	0.96
Females Over 26									
SP	33.85 (0.92)	1.22 (0.18)	1.70 (0.36)	-				0.92	0.92
GE	16.03 (0.57)	1.14 (0.18)	0.67 (0.36)	0.61	-			0.91	0.92
SC	9.88 (0.27)	3.20 (0.18)	12.59 (0.36)	0.51	0.50	-		0.82	0.85
ER	14.36 (0.50)	0.82 (0.18)	-0.09 (0.36)	0.57	0.75	0.43	-	0.89	0.89
Total Scale								0.95	0.95
Whole Sample									
SP	42.57 (0.64)	0.62 (0.09)	-0.25 (0.19)	-				0.94	0.94
GE	18.19 (0.35)	0.91 (0.09)	0.02 (0.19)	0.62	-			0.92	0.93
SC	13.13 (0.27)	1.67 (0.09)	2.34 (0.19)	0.71	0.61	-		0.91	0.91
ER	16.65 (0.28)	0.49 (0.09)	-0.60 (0.19)	0.54	0.66	0.53	-	0.89	0.89
Total Scale								0.96	0.97

Note: **ω**= McDonalds; SP = Successful psychopathy; GE=Grandiose entitlement; SC=Sadistic cruelty; ER=Entitlement rage.

5.4 Discussion

The purpose of this study was to develop a scale which measured how the dark personality traits of Machiavellianism, psychopathy, narcissism and everyday sadism manifest in a general population sample across age and sex. It was apparent from the initial Exploratory Graph Analyses (EGA), that there were qualitative differences across all groups, especially for Machiavellianism. Following reduction of the indicators through redundancy analysis and the Generalised Partial Credit Model (GPCM), Machiavellianism and psychopathy were mainly found to load onto one cluster, with grandiose entitlement, entitlement rage and sadistic cruelty forming their own clusters. This model was corroborated through confirmatory factor analysis and measurement invariance. The final four-factor model allows meaningful inter-group comparisons in research (Plouffe, Smith & Saklfoske, 2018).

However, there are a few issues which need to be addressed. The GPCM elucidated that the sixth point, ‘very much like me’, of the scale was not being utilised across some items, except for the 18-25 males, where all options were used for all items. Although collapsing the scale from six to five indicators was considered, it has been argued that collapsing categories upwards (e.g., indicator five into indicator six), or downwards (e.g., indicator five into indicator four), should only be considered if evidence can be provided that both categories tap into the same level of the latent trait (Bond & Fox, 2007), which in this instance, it could not. When participants endorse scale items in answer to a question, they may consider the intrapsychic relatedness of the qualitative description (Klockars & Yamagishi, 2005). Therefore, collapsing category six ‘very much like me’ and five, ‘mainly like me’, would have merged data from categories which had different semantic meanings to participants, thereby, misrepresenting responses (McCullagh, 1985). In light of this, the option to collapse the scale was rejected. This ensured that meaningful information was preserved and participant responses were not altered. It further enabled the maximum amount of useful data to be extracted, which would otherwise have been marginalised.

Consequently, the data provided the opportunity to ensure that the scale items covered the core features of the constructs across varying levels of the latent traits (Monaghan et al., 2019; Revelle, 2010), whilst remaining sex and age invariant. The variation in the GPCM parameters across items suggested this was achieved. The scale items which emerged from the GPCM were submitted to a final EGA, primarily to

determine whether psychopathy and Machiavellianism were subsumed under one factor, however, the older female group deviated from the other groups, whereby Machiavellianism and psychopathy were mainly found on their own dimensions. Moving from a graphical model which focused on the estimation of direct relationships between the indicators through the inverse covariance matrix to a latent factor model (Golino & Epskamp, 2017), and applying the four-factor model to all groups evidenced a good fit and invariance across age and sex. The findings corroborated previous research that Machiavellianism and primary psychopathy are parallel constructs (McHoskey et al., 1995; Miller, Hyatt et al., 2017; Persson, 2019; Persson, Kajonius & Garcia, 2017)

In relation to the constructs themselves, the subsumption of psychopathy and Machiavellianism, adds further support to the concept of the successful psychopath, (Hall & Benning, 2006; Mullins-Sweatt et al., 2010; Widom, 1977). The DSHS indicators for this factor, labelled, 'Successful Psychopathy', are entwined with Cleckley's (1941), stated characteristics of primary psychopathy and those of Machiavellianism (Christie & Geis, 1970; Miller, Hyatt et al., 2017; Mullins-Sweatt et al., 2010). The literature has suggested that the successful psychopath embodies strategic planning (Palmen et al., 2018; Poythress and Hall, 2011), a trait historically attributed to Machiavellianism and contended to be the main facet which separated it from psychopathy (Jones, 2016). Yet, evidence has been provided showing psychopathy to be positively associated with good communication and strategic skills (Babiak, Neumann & Hare, 2010), whereby these characteristics have been instrumental in achieving personal ambitions (Lilienfeld, Waldman et al., 2012). Machiavellianism as a construct has been put forth as not being guided by any theoretical framework, which has been related back to the initial observations made by Christie and Geis (1970), suggesting there were individual differences in Machiavellianism, yet with no clarity as to why (Wilson, Near & Miller, 1996). Researchers have been arguing for the recognition of the successful psychopath in general population samples since Widom (1977). The DSHS provides this.

The narcissism indicators in the DSHS corroborated previous research (Ackerman, Donnellan & Wright. 2019; Dinić, Sokolovksa & Tomašević, 2021), and are in line with the opinion of expert opinions (Rosenthal & Hooley, 2010), that entitlement is the core maladaptive characteristic of narcissism (Dickinson & Pincus, 2003), which bridges the grandiose and vulnerable manifestations (Dinić, Sokolovksa

& Tomašević, 2021). Factor Two, addresses grandiose entitlement whilst factor four taps into entitlement rage (Pincus, Ansell et al., 2009; Wright, Lukowitsky et al., 2010), which represents the vulnerable narcissist (Wright & Edershile, 2018). The DSHS was developed to measure the socially toxic aspects of these specific dark personality traits and entitlement emerged as the core manifestation in the general population sample used. Both facets are a primary determinant of narcissism, which is especially socially malicious and entwined with maladaptive outcomes (Brown, Budzek & Tamborksi, 2009; Maxwell et al., 2011).

Grandiose entitlement indicators in the DSHS align with one who fundamentally believes they embody special privilege, which emerges from a self-image of superiority, which may result in a willingness to exploit others, thereby reflecting a grandiose ego (Ackerman & Donnellan, 2013). Entitlement Rage indicators capture experiences of anger when the individual feels their expectations are not met. Underlying this is a reflection of a fragile ego, entwined with the emotional dysregulation of rage and destructiveness (Bishop & Lane, 2002; Moses & Moses-Hrushovski, 1990; Pincus & Lukowitsky, 2010). Entitlement and superiority are embodied by narcissists across both general population and incarcerated samples, suggesting these traits are the most dangerous aspects of the construct. For example, violent incarcerated offenders scored highest on the NPI subscales of Entitlement and Superiority (Bushman et al., 1999). From all the narcissism indicators included in the item pool, the findings from the analysis evidenced that the darkest aspects of narcissism are not constrained by clinical, forensic or general population origins.

Turning to everyday sadism, this study has shown how this construct manifests in the general population. One noticeable difference across samples was found in the females over 26, as seen in the GPCM, where only one sadism item was endorsed to the fullest extent, which pertained to sadistic fantasies. This was also reflected in the mean score for this factor whereby females over 26 scored the lowest mean. Overall, the indicators address the construct along the continuum from schadenfreude to direct physical or psychological harm (O'Meara et al., 2011). At the most extreme end, rather than physical sadism, those who embody the construct inflict psychological harm on others both in online as well as real world situations. As previously discussed, it appears that power and pleasure are entwined with the construct, which may suggest a shift in the definition, to addresses the cruelty interwoven with everyday sadism. Cruelty emerges from thoughts, yearnings and intentions, which are articulated

intrapsychologically or externally and can be conveyed through action or speech, which results in the real or virtual suffering of others (Taylor, 2009). Thus, the everyday sadist takes intrinsic pleasure, mediated by cruel intentions, from imagining, witnessing or inflicting physical or psychological pain on others (Bulut, 2017; Smith, Powell et al., 2009; van Dijk & Ouwerkerk, 2014).

Sociocultural contexts may motivate sadistic behaviours, whereby some may derive pleasure from the pain of others for intrinsic gratification, others may be compelled to commit these acts through, for example, zealous political or ideological allegiances, by unwavering obedience to authority, or other influences (Kaminer & Stein, 2001). It may be that everyday sadistic acts satisfy the need for power over others, thereby acting as a panacea for powerlessness (Wilson & Seaman, 2007). A converse viewpoint from an evolutionary perspective suggests that power can be maintained through sadistic behaviours, whereby power can confer more sexual opportunities, implying that sadism may have been selected as a reproductive strategy (Dutton, 2012; Nell, 2006).

Turning to sex differences, as previously discussed in Chapters Two and Four, a sex stereotypical narrative has been attributed to the measures (Corry et al., 2008; Jonason & Webster, 2010; Levenson et al., 1995). Yet, as psychometric methods have advanced and statistical software has become readily available, this narrative can be challenged, as seen in Chapter Four. The intersex mean score differences found from the DSHS showed that males in the 18-25 age group scored higher means across all factors, whereas the older female group consistently held the lowest mean scores. Although this in line with past research (e.g., Paulhus, Buckels et al., 2020), it can be questioned as to what comparison of means scores between sexes reveals. Investigating dark personality traits as an ecosystem, through lower-order methods such as network analysis (Briganti & Linkowski, 2020; Dinić, Sokolovska & Tomašević, 2020; Dinić, Wertag, Tomašević & Sokolovska, 2020; Papageorgiou et al., 2019; Truhan et al., 2020), or through item response theory (Carter, Campbell et al., 2015), may provide a more in-depth understanding of the manifestations of the constructs. Indeed, applying semantic meaning to the manifestations of the constructs across and between sexes as well as quantifying them, will facilitate a far more in-depth understanding of how the expression of the traits may change over, across and between sexes and ages.

5.4.1 Conclusion

Many researchers who have developed scales have used classical test theories (CTT), involving the same or similar statistical methods (e.g., Jonason & Webster, 2010; Plouffe, Saklofske & Smith, 2017). These approaches appeared to have become ingrained and to an extent, an unwritten law which researchers historically rarely challenged (Berchtold, 2016). Yet, advances in methods and software enables researchers to investigate data in more psychometrically robust ways, from both top-down and bottom-up approaches. Indeed, network analysis and item response theories are becoming more commonly used (Grigoras et al., 2020; Trahair et al., 2020,). This study has highlighted how they can work symbiotically with CTT, to develop a scale, determined at the item level, which provides a good model fit when analysed through traditional methods.

The indicators of the DSHS factors align with the conceptual frameworks and expert ratings where available, of psychopathy, Machiavellianism, narcissism and everyday sadism. However, in the case of psychopathy and Machiavellianism, the evidence provided confirmation that the two constructs are parallel and are subsumed under one factor (e.g., McHoskey et al., 1995), which represents the successful psychopath (Babiak et al., 2006; Palmen, Derkson & Kolthoff, 2020).

The DSHS is both sex and age invariant and provides an alternative measure to investigate the dark personality traits from a different viewpoint. An advantage of using the Generalised Partial Credit Model during development is that each factor can be used as a stand-alone measure, as the item and latent trait values, when determined through this method, do not fluctuate across samples (Hays et al., 2000; van der Linden, 2016). The DSHS has the potential to produce results which can be meaningfully compared across and between groups, with the preliminary evidence suggesting it is psychometrically robust.

Chapter Six

The Construct Validity and Temporal Reliability of the Dark Side of Humanity Scale

6.1 Introduction

The Dark Side of Humanity Scale (DSHS), has shifted away from the traditional Dark Tetrad (DT), configuration of psychopathy, Machiavellianism, narcissism and everyday sadism (Buckels, Jones & Paulhus, 2013; Paulhus, Buckels et al., 2020). With the symbiotic use of a diverse range of statistical methods as discussed in Chapter Five, the subsumption of primary psychopathy and Machiavellianism (McHoskey et al., 1995; Miller, Hyatt et al., 2017; Persson, 2019; Persson, Kajonius & Garcia, 2017), determined the first factor to represent the successful psychopath (Babiak, Neumann & Hare, 2010; Palmen et al., 2018). Narcissism is addressed by two factors of grandiose entitlement and entitlement rage, which depict the grandiose and vulnerable variants respectively (Pincus, Ansell, et al., 2009), and are considered core components of narcissism (Dickinson & Pincus, 2003). The everyday sadism factor epitomises the intrinsic pleasure, mediated by cruel intentions, from imagining, witnessing or inflicting physical or psychological pain on others (Bulut, 2017; Smith, Powell et al., 2009; van Dijk & Ouwerkerk, 2014). However, further psychometric evidence is required to further confirm the four-factor model as well as the validity of the DSHS against widely available dark personality measures, it's standing in the nomological network and temporal reliability.

However, the first aim of this study is to verify the factor structure of the DSHS through confirmatory factor analysis (CFA), which evaluates the hypothesised structure of a fixed set of variables and is widely used in validation research (Williams, Paulhus & Hare, 2007). Following this, convergent validity will elucidate the extent of the positive and significant relationship with widely used DT measures (Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995), the Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979), the Mach IV (Christie & Geis, 1970), and the Assessment of Sadistic Personalities (Plouffe, Saklofske & Smith, 2017), which will provide an indication that the DSHS is measuring the constructs it was developed to measure (Jonason & Webster, 2010). Further validity will be assessed with measures which are considered to be part of the nomological network surrounding the four dark personality traits, specifically the Big Five (Soto & John,

2017), and a global measure of self-esteem (Rosenberg, 1965). These measures represent 'normal', lower order personality traits and their historical relationships with dark personality measures have provided a good indication of both convergent and discriminant validity (Jonason & Webster, 2010; Seemann et al., 2005). Discriminant validity indicates that the constructs measured by the DSHS are conceptually distinct from other constructs, showing low or negative correlations (Shaffer et al., 2016).

The Big Five personality traits (Soto & John, 2017), are considered fundamental in personality research, as they constitute a shared narrative from which those who score highly on DSHS factors can be described (Jonason, Kaufman et al., 2013). For example, a positive relationship with agreeableness and conscientiousness would indicate personality stability, whereby compassion, respect and trust of others are embodied, as well as the extent of responsibility, organisation and productivity held (Soto & John, 2017; Soto, Kronauer, & Liang, 2016). However, at the centre of dark personality traits, there is an intimation of personality instability, dogmatism and a reluctance to change (Spain et al., 2014). Personality instability suggests the extent to which an individual may be emotionally volatile (Soto & John, 2017; Soto, Kronauer & Liang, 2016).

Whilst dark personality constructs have consistently shown a negative relationship with Agreeableness (Kowalski, Vernon & Schermer, 2019; O'Boyle et al., 2015; Vernon et al., 2008), the remaining DT facets have shown varied associations. For example, extraversion has shown no relationship with Machiavellianism (Jonason & Webster, 2010), whilst narcissism and psychopathy have produced positive associations (Jonason & Webster, 2010; Fernández et al., 2020; O'Boyle et al., 2015; Vernon et al., 2008). Conscientiousness has correlated positively with narcissism (O'Boyle et al., 2015), whilst with psychopathy and everyday sadism there have been negative relationships (Jonason & Webster, 2010; Kowalski et al., 2019; Vernon et al., 2008). Consequently, the connections the DSHS facets have with the Big Five personality domains should highlight their antisocial nature.

One issue which needs addressing before the study begins, is the ongoing debate in the literature as to whether the Honesty/Humility factor from the HEXACO scale should be used alongside the Big Five measures in dark personality research (Bizumic & Monaghan, 2020). On consideration of the responses in the Open Peer Commentary (Bizumic & Monaghan, 2020), such as those of Fleeson (2020), who argues that the Honesty-Humility facet omits important content such as truthfulness,

resulting in incongruity between the label of the trait and the meaning of the trait. Lynam, Crowe, Vize and Miller (2020), in the same commentary, contend that the Honesty-Humility traits do not emerge as separable from Agreeableness indicators. Indeed, on perusal of the facets items, there was considerable overlap with those included in dark personality measures. For example, the Honesty-Humility item, 'I think that I am entitled to more respect than the average person is' (Ashton & Lee, 2005), is similar to the item from the DSHS entitlement factor, 'I expect to be treated better than average', and from the NPI, 'I insist on getting the respect that is due me' (Raskin & Hall, 1979). A psychopathy item from the LSRP, 'Making a lot of money is my most important goal' (Levenson et al., 1995), is comparable to the reverse scored indicator from the Honesty-Humility factor, 'Having a lot of money is not especially important to me'.

Therefore, the decision was made at the outset of this study not to use Honesty-Humility. This has since been supported by Lynam, Miller et al., (2020), who argued that Agreeableness from the Big Five, is more strongly related to the Dark Triad traits than Honesty-Humility. Psychometric evidence from multiple samples using multiple dark personality measures found when using the single construct scales, such as the Mach IV, over brief measures, such as the Dirty Dozen, the BFI-2 (Soto & John, 2017), and Honesty-Humility accounted for a similar amount of variance (Vize, Collison et al., 2020). A further study using Goldberg's (2006), bass-ackwards approach showed that that a separate Honesty-Humility factor did not emerge at any stage in the hierarchy (Crowe, Lynam & Miller, 2018). The evidence strongly suggests that the Honesty-Humility facet does not capture the most fundamental aspects of the Agreeableness domain (Lynam, Miller, et al., 2020), and as such the Agreeableness facet was determined as more pertinent in this study to assess the DSHS.

Investigating an individual's belief in their own worth (Rohmann et al., 2015), is often used to examine the nomological network of dark personality traits (Jonason & Webster, 2010; Malesza et al., 2019; Özsoy et al., 2017). As such, the Rosenberg Self-Esteem scale (RSE: Rosenberg, 1965), which relates to the feeling of being 'good enough' and combines subjective attitudes of self-liking and self-competence (Rosenberg, 1979), will be used. Previous research has found a convergence with self-esteem and the dark triad, suggesting resilience against failures of interpersonal exploitation (Jonason, Koenig & Tost, 2010). Whilst others, finding discriminant validity, have attributed this to low self-esteem, whereby individuals wear a mask of

self-confidence to gain and increase social status, which hides the shame felt at an intrapsychic level, when their behaviours fall outside of societal norms (Volmer et al., 2019).

Variation has been found across studies, with some showing a lack of relationship between DT constructs and self-esteem (Hunter et al., 1982; Jonason & Webster, 2010; Malesza et al., 2019). However, further research has found positive correlations with the grandiose facet of narcissism (Chin et al., 2017; Jonason, Li & Teicher, 2010; Strelan, 2007), and negative relationships with vulnerable narcissism and Machiavellianism (Hunter et al., 1982; Rauthmann, 2012). A further study found no association with entitlement (Rohmann et al., 2015; Strelan, 2007). The relationship between psychopathy and self-esteem appears to be more complex, with findings showing that associations are sex dependent. For males, higher levels of psychopathy have been associated with higher levels of self-esteem, whilst for females, higher levels of psychopathy have revealed lower levels (Visser et al., 2010). Without separating the sample by sex, Chin et al., (2017), found negative associations between psychopathy and self-esteem. To date, no found research has investigated everyday sadism and self-esteem, yet it is proposed that behaviours which seek to dominate and harm others may reflect tactics which maintain and enhance self-esteem levels (Crocker & Park, 2004; Pfattheicher & Schindler, 2015).

Temporal stability will be assessed across two time periods. Following the initial completion of the survey, the same participants will be invited to respond to it once more, sixteen days later (Berchtold, 2016; Watson, 2004). This time lapse was chosen over the two-weeks suggested by Watson (2004), to enable participants to withdraw their data from the study, if they so wished. Sixteen days is a close enough timeframe in which the traits being measured are not expected to change. Shorter time intervals may be impacted by memory effects (Schmidt et al., 2003), and salient life events such as family bereavement, can influence affective traits more so than the Big Five (Vaidya, Gray, Haig, Mroczek & Watson, 2008). Although this time period is generally short enough to ensure, as much as possible, that true changes in the traits are not likely, whereby the impact of contextual factors should be insignificant within sixteen days, (Chmiekewski & Watson, 2009), there is a caveat. The first part of the study was conducted on 6th April 2020, when the United Kingdom was in Covid-19 lockdown, the second part, on the 22nd April 2020, was when an extension to the lockdown was announced. Indeed, during this time period, there were restrictions in place globally.

Thus, any impact from the Covid-19 regulations at an intrapersonal level may be reflected in both data samples and any group differences may be attributed to genuine changes, rather than an unstable scale (Aldridge et al., 2017). If the results show temporal stability as well as convergent and discriminant validity, it will provide confidence that the DSHS is dependable and enables reproducibility, therefore enabling it to be put into wider use in research (Aldridge et al., 2017; Schmidt et al., 2003).

In light of this discussion various predictions can be made. It is considered that negative relationships with the Big Five and the DSHS will be found with Agreeableness and Conscientiousness (Chiorri et al., 2017; Fernández-del-Río et al, 2020; Furnham et al., 2013). It is further hypothesised that Negative Emotionality will show positive correlations with the DSHS (Kowalski et al., 2019; O'Boyle et al., 2015; Vernon et al., 2008). It is also postulated that there will not be any negative associations with self-esteem across the DSHS factors (Hunter et al., 1982; Jonason & Webster, 2010; Malesza et al., 2019; Pfattheicher & Schindler, 2015). Finally, it is predicted that the DSHS will evidence temporal reliability.

6.2 Method

Confirmatory Factor Analysis and Convergent and Discriminant Validity

6.2.1 Participants

An online study was designed to assess the performance of the DSHS against the prevalent Dark Tetrad scales. Invitations to participate were posted on a crowd-data provider, Prolific (Prolific, 2019), to ensure a cross-sectional sample, and mainly equal group sizes. Prolific participants were paid £1.67 for completion of the survey. The invitation to participate also included details relating to the second part of the study of test-retest, informing participants that they may be invited to participate,16 days after completion of this first part. Ethical approval was received from the University of Derby (Appendix 5.1). There were 712 participants, comprising 348 females and 364 males with an age range of 18-70 (M_{age} = 28.87, *SD* = 9.75). Split by groups; Males 18-25 (*n* = 186; M_{age} = 21.45, *SD* = 2.23); Females 18-25 (*n* = 165; M_{age} = 22.12, *SD* = 2.13); Males 26 and Over (*n* = 178; M_{age} = 35.25, *SD* = 9.34); Females 26 and Over (*n* = 183; M_{age} = 36.28, *SD* = 9.21).

6.2.2 Participants Test-Retest

The second part of the online study was designed to assess the temporal reliability of the DSHS. Invitations to participate were sent through the Prolific website, sixteen days after completion of part one of this study, using the participants Prolific ID's, thus ensuring anonymity. Across all groups 413 participants (194 females; 219 males), who completed the first part of the survey accepted the invitation for this second part. The age range varied between 18-64 (M_{age} = 29.42, *SD*=9.86) with no missing data. Mean ages of the groups: Males 18-25 (*n* = 99; M_{age} = 21.33, *SD* = 2.31); Females 18-25 (*n* = 87; M_{age} = 22.15, *SD* = 2.05); Females 26 and Over (*n* = 107; M_{age} = 36.17, *SD* = 9.29); Males 26 and Over (*n* = 120; M_{age} = 35.33, *SD* = 9.12). As there were fewer participants at time two, the Prolific ID codes, which are unique to each participant, and therefore adhere to anonymity, were matched and the data from the two time points were combined into one dataset.

6.2.3 Measures Convergent and Discriminant Validity

Dark Side of Humanity Scale (DSHS)

This scale comprises 42 items over four factors; Successful psychopathy (SP), grandiose entitlement (GE), sadistic cruelty (SC), and entitlement rage (ER; Appendix 5.2.1). It is measured on a six-point Likert type scale from *unlike me* (1) to *very much like me* (6). The scale is recommended to be used as a four-factor model. Reliabilities have been found to be good, with SP α =.92; GE α =.90; SC α =.88 and ER α =.89.

The Mach IV (Christie & Geis, 1970).

The Mach IV comprises 20 items (Appendix 3.2.1) which are measured on a 7-point Likert scale from *strongly disagree* (1) to *strongly agree* (7). Items include 'Anyone who completely trusts anyone else is asking for trouble'. The total scale reliability has been shown to be adequate α=.74 (Abell & Brewer, 2018).

The Levenson Self-Report Psychopathy Scale (LSRP; Levenson et al., 1995)

The LSRP is a 26-item scale (Appendix 3.2.2), which incorporates two factors. Primary psychopathy, measures callous and manipulative interpersonal characteristics and secondary psychopathy, taps into behavioural and antisocial facets. Measured on a four-point Likert scale *disagree strongly* (1) to *agree strongly* (4), items for primary psychopathy include, 'I often admire a really clever scam', and for secondary psychopathy, 'Love is overrated'. Total scale; α=.87; Primary; α=.86 and Secondary; α =.63 (Marion & Sellbom, 2011).

Narcissistic Personality Inventory (NPI: Raskin & Hall, 1979).

The NPI comprises 40 items (Appendix 5.2.2) and is scored on a 5-point Likert scale ranging from *strongly disagree* (1) to *strongly agree* (5). Example items include 'I am an extraordinary person'. The Likert scoring has demonstrated higher internal consistency than the original binary forced choice version, which asks participants to select from a narcissistic or non-narcissistic response (Miller, Gentile, et al., 2018; Raskin & Hall, 1979). Previous research has found that the NPI demonstrates good internal consistency (α =0.95; Miller, Gentile, et al., 2018).

The Assessment of Sadistic Personality (ASP; Plouffe et al., 2017)

The ASP comprises nine items (Appendix 3.2.8), which measure everyday sadism and are rated on a 5-point Likert scale *strongly disagree* (1) to *strongly agree* (5). Scale items consist of statements such as 'Watching people get into fights excites me'. Previous research has evidenced that the ASP demonstrates good internal consistency (α = .86; Kowalski et al., 2019).

Big Five Inventory-2 (BFI-2: Soto & John, 2017).

This measure is a 60-item self-report scale (Appendix 5.2.3), of core personality traits: Negative Emotionality, Extraversion, Open-Mindedness, Agreeableness and Conscientiousness. Participants rate their agreement using a 5-point Likert type scale ranging from *disagree strongly* (1) to *agree strongly* (5). Each subscale is comprised of 12 items and the values obtained by averaging the scores together. Items are short, descriptive phrases, which all begin with 'I am someone who…', followed by item-specific content, for example, 'has a forgiving nature'. Alpha reliability has been shown as good, Extraversion (.90), Agreeableness (.83), Conscientiousness (.88), Negative Emotionality (.90), Open-Mindedness (.78) (Kowalski et al., 2019)

Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1965)

The RSE comprises ten items (Appendix 5.2.4), which measure self-esteem, five of which are positively worded, e.g. 'I feel that I have a number of good qualities', and five which are negatively worded, e.g. 'I wish I could have more respect for myself'. Items are scored on a 4-point Likert type scale, with *strongly disagree* (1) to strongly agree (4). Cronbach alpha reliability for the total scale has been shown to be good with .93 for men and .91 for women (Purton et al., 2018).

Test-Retest Measure

Dark Side of Humanity Scale (DSHS)

Details as above.

6.2.4 Procedure for Convergent and Discriminant Validity

The invitation to participate provided a link to the online survey hosted by Qualtrics (Qualtrics, 2019), where involvement was voluntary and confidential. Inclusion criteria stated that participants had to be aged 18 or over and fluent in English. Prior to commencement of the survey, participants were briefed about the study and its aims, they were then asked for their consent to proceed. Once this was given, participants were presented with demographic questions followed by the questionnaires, for which they were informed they could take their own time completing. Contact information was given for the researcher and supervisor, should participants wish to ask any questions about the nature of the research. They were also informed that they could withdraw their data within two weeks of taking part, with their responses removed from the study by providing the unique identification code they had entered before beginning the survey.

Procedure for Test-Retest

The procedure mirrored the first part of this study. However, as there were fewer participants at time two, than time one, the Prolific ID codes, which are unique to each participant, and therefore adhere to anonymity, were matched and the data from the two time points were combined into one dataset.

6.2.5 Data Analysis

Full details of the statistical methods used are provided in the methodology chapter. All analyses were conducted in R (Version 1.3.1093). In brief, the data from the first sample of this study was assessed by confirmatory factor analysis to verify the four-factor structure (Williams, Paulhus & Hare, 2007). The estimator was set as weighted least means and variance (WLSMV). The model fit was assessed using chi square (χ^2), Confirmatory Fit Index (CFI), Tucker-Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA). A good model fit was determined when CFI and TLI values were more than or equal to .90, in line with Kline's (2015), recommendations. RMSEA values evidenced a good fit if they were .06 or below and an adequate model fit was accepted at .08 or below. Models were rejected if they

displayed a value above 1.0 (Brown, 2015; Hu & Bentler, 1998; Marsh, Wen & Hau, 2004).

The design for both convergent and discriminant validity and test-retest was correlational. Descriptive statistics were investigated for each sample. For convergent and discriminant validity and test-retest validity, robust correlations using the WRS2 package were calculated. This method guards against the effect of outliers (Mair & Wilcox, 2019), and is in line with the robust methods used across studies. Consistent with all previous analyses, the outliers were not removed. Although researchers may view outliers as problematic and remove them to run a 'cleaner' data analysis (Aguinis et al., 2013), it has been argued that in certain research domains the presence of outliers can lead to important theoretical developments and removing them may compromise findings, thereby diminishing the accuracy of the models which portray the constructs under investigation (Aguinis et al., 2013; Neumann, Kosson & Salekin, 2017). As such, the outliers were considered to represent valid values in the data and were therefore retained due to the nature of this study.

6.3 Results

Descriptive Statistics

Descriptive statistics and alpha reliabilities were calculated for total scale and subscale scores, where applicable, shown in Table 21. Outliers were present as assessed by Mahalonobis distance which revealed; males 18-25 (ASP (1); Primary psychopathy (1); SC (3); females 18-25 (SP(2); GE(2); ER(5); Extraversion (1); secondary psychopathy (2); SC(2); Males 26 and Over; ASP(1); SP(1); GE(1); ER(1); SC(2); Females 26 and Over; ASP(1); SP(1); GE(1); SC(40)

Table 21

Descriptive Statistics for the Total Scales and Subfactors convergent/discriminant validity

Scale	F18 Mean (*SE*)	α	M18 Mean (*SE*)	α	F26 Mean (*SE*)	α	M26 Mean (*SE*)	α	Whole Sample Mean (*SE*)	α
DSHS										
Successful Psychopathy	34.88 (1.05)	.93	46.46 (1.22)	.93	31.50 (0.86)	.92	40.56 (1.12)	.92	38.46 (0.58)	.94
Grandiose Entitlement	18.85 (0.67)	.90	21.09 (0.67)	.89	17.12 (0.53)	.89	19.59 (0.65)	.91	19.18 (0.32)	.90
Sadistic Cruelty	10.59 (0.38)	.90	13.41 (0.48)	.89	9.91 (0.27)	,87	11.34 (0.34)	.84	11.34 (0.20)	.88
Entitlement Rage	17.48 (0.59)	.91	18.80 (0.52)	.86	14.60 (0.47)	.89	17.07 (0.54)	.90	16.98 (0.27)	.89
LSRP										
Primary Psychopathy	28.61 (0.53)	.81	33.15 (0.58)	.84	26.10 (0.52)	.86	30.30 (0.63)	.88	29.57 (0.30)	.86
Secondary Psychopathy	21.25 (0.38)	.72	22.15 (0.33)	.64	19.73 (0.33)	.71	21.23 (0.33)	.66	21.09 (0.17)	.69
Mach IV	3.62 (0.54)	.80	4.00 (0.05)	.80	3.42 (0.05)	.80	3.79 (0.06)	.84	3.71 (0.03)	.83
NPI	111.83 (1.73)	.93	118.20 (1.68)	.92	103.39 (1.65)	.93	112.98 (1.78)	.94	111.62 (0.88)	.93
ASP	15.61 (0.47)	.83	19.97 (0.50)	.83	13.85 (0.41)	.85	17.20 (0.47)	.84	16.70 (0.25)	.85
BFI-2										
Extraversion	38.10 (0.70)	.86	36.10 (0.59)	.80	35.49 (0.62)	.82	35.93 (0.59)	.81	36.37 (0.31)	.83
Agreeableness	43.71 (0.57)	.79	40.71 (0.53)	.78	45.75 (0.52)	.79	42.34 (0.55)	.80	43.11 (0.28)	.80
Conscientiousness	39.90 (0.69)	.85	37.87 (0.56)	.80	42.93 (0.63)	.85	39.74 (0.67)	.88	40.11 (0.33)	.85
Negative Emotionality	40.48 (0.75)	.89	33.94 (0.74)	.89	37.46 (0.82)	.92	33.68 (0.72)	.89	36.30 (0.39)	.91
Open Mindedness	45.24 (0.57)	.78	43.78 (0.52)	.75	44.99 (0.60)	.84	44.84 (0.55)	.79	44.70 (0.28)	.79
RSE Self-Esteem	23.43 (0.40)	.88	22.35 (0.42)	.88	23.03 (0.43)	.91	22.62 (0.44)	.89	22.84 (0.21)	.89

Note: DSHS=Dark Side of Humanity Scale; LSRP=Levenson Self-Report Psychopathy Scale; NPI = Narcissistic Personality Inventory; ASP=Assessment of Sadistic Personality; BFI-2=Big Five Inventory-2; RSE= Rosenberg Self-Esteem Scale

The mean scores indicated that for the dark personality scales, males scored higher than females which is consistent with previous research (e.g., Jonason & Webster, 2010). The Big Five personality traits showed that the mean scores for Extraversion and Open Mindedness were higher for females, particularly the 18-25 group, and female groups scored higher than males on Agreeableness, Conscientiousness and Negative Emotionality. Females also showed higher mean scores for self-esteem. All Cronbach alpha scores displayed good reliability across scales and subfactors, except for the LSRP Secondary facet which showed adequate reliability.

Confirmatory Factor Analysis

The robust findings corroborated those in the previous study for the four-factor structure by showing a good fit to the data; males 18-25 (χ^2 =979.55, *df* = 813, *p* =.00, CFI = 0.93, RMSEA = 0.03); females 18-25 (χ^2 =1122, 55,*df* = 813, *p* = 0.00, CFI = 0.96, RMSEA = 0.05); Males 26 and Over (χ^2 = 1217.04, *df* = 813, *p* = 0.00, CFI = 0.95, RMSEA = 0.05); Females 26 and Over (χ^2 =905.72, *df* = 813, *p* =0.01, CFI = 0.93, RMSEA = 0.02). This provided evidence from a new sample, that the DSHS is psychometrically robust (Appendix 5.3).

Convergent and Discriminant Validity

Correlations in bold (Table 22) show the level of association with the DSHS and its corresponding measure. A strong relationship was shown with successful psychopathy, the LSRP primary factor and the Mach IV across all groups. Grandiose Entitlement (GE) mainly displayed a strong association with the NPI, however with entitlement rage, the relationship was moderate, as was the relationship between successful psychopathy (SP) and the secondary facet of the LSRP. There were strong correlations between the ASP and sadistic cruelty (SC), thus across dark personality measures, convergent validity was evidenced (Appendix 5.4).

In relation to the Big Five, discriminant validity was shown across groups for Agreeableness, and Conscientiousness. Extraversion showed a significant positive correlation with GE (*r* =.20), for the younger males, whilst for females 18-25, there was a significant negative association with SP (*r* = -.19). Negative emotionality and entitlement rage (ER), were positively and significantly correlated across groups, however for the younger females, this relationship was also found with SP and SC. Open Mindedness held a significant negative association with SP for all groups except males 18-25, where the relationship was with SC (*r* = -.17).

Table 22

Robust Correlations Among the Dark Side of Humanity Scale, Dark Tetrad Measures, the Big Five and Self-Esteem

Sample	Factor	LSRP 1	LSRP2	NPI	NPI Ent	NPI Exp	Mach IV	ASP	E	A	C	NE	OM	SE
F18-25	SP	**0.71*** **	**0.54*** **	0.34***	0.26***	0.36***	**0.64*** **	0.58***	-0.11	-0.54***	-0.33***	0.16*	-0.15**	0.25***
	GE	0.46***	0.40***	**0.48*** **	**0.41*** **	0.45***	0.43***	0.45***	0.04	-0.34***	-0.24***	0.04	0.02	0.05
	SC	0.57***	0.39***	0.21**	0.18*	0.22**	0.54***	**0.61*** **	-0.19*	-0.47***	-0.26***	0.27***	-0.08	0.21**
	ER	0.42***	0.48***	**0.41*** **	**0.42*** **	0.42***	0.46***	0.42***	-0.03	-0.40***	-0.26***	0.31***	0.07	0.18**
M18-25	SP	**0.81*** **	**0.41*** **	0.44***	0.31***	0.48***	**0.68*** **	0.64***	0.08	-0.63***	-0.18*	-0.01	-0.12	-0.02
	GE	0.58***	0.35***	**0.52*** **	**0.51*** **	0.33***	0.39***	0.58***	0.20**	-0.51***	-0.16*	0.14	0.00	0.00
	SC	0.58***	0.38***	0.28***	0.22***	0.29***	0.47***	**0.66*** **	0.01	-0.44***	-0.18	0.09	-0.17*	0.12
	ER	0.47***	0.44***	**0.41*** **	**0.38*** **	0.28***	0.26***	0.45***	0.08	-0.38***	-0.21**	0.28***	-0.04	0.11
F26+	SP	**0.75*** **	**0.51*** **	0.40***	0.40***	0.45***	**0.70*** **	0.63***	0.05	-0.54***	-0.26***	0.17*	-0.09	0.14
	GE	0.58***	0.36***	**0.53*** **	**0.51*** **	0.45***	0.40***	0.52***	0.09	-0.41***	-0.18*	0.07	-0.02	-0.03
	SC	0.57***	0.47***	0.21**	0.22***	0.30***	0.54***	**0.71*** **	-0.07	-0.47***	-0.22**	0.14	-0.09	0.18**
	ER	0.60***	0.56***	**0.43*** **	**0.50*** **	0.40***	0.55***	0.60***	0.13	-0.48***	-0.28***	0.30***	-0.08	0.25***
M26+	SP	**0.76*** **	**0.45*** **	0.30***	0.30***	0.43***	**0.59*** **	0.59***	-0.01	-0.66***	-0.25**	0.00	-0.18*	0.06
	GE	0.68***	0.38***	**0.55*** **	**0.55*** **	0.50***	0.49***	0.50***	0.12	-0.50***	-0.07	-0.03	-0.09	-0.09
	SC	0.50***	0.39***	0.25***	0.22***	0.39***	0.44***	**0.64*** **	-0.05	-0.45***	-0.30***	0.14	-0.22**	0.15
	ER	0.50***	0.47***	**0.41*** **	**0.51*** **	0.47***	0.48***	0.42***	-0.00	-0.45***	-0.15*	0.20**	-0.04	0.02
Total Sample	SP	**0.79** **	**0.49*** **	0.41**	0.34***	0.47***	**0.69*** **	0.67***	0.00	-0.61***	-0.29***	-0.01	-0.17***	0.07
	GE	0.59***	0.38***	**0.54*** **	**0.51*** **	0.46***	0.45***	0.53***	0.11	-0.46***	-0.19***	0.03	-0.03	-0.03
	SC	0.59***	0.42***	0.28***	0.24***	0.34***	0.54***	**0.68*** **	-0.07	-0.48***	-0.26***	0.07	-0.15**	0.12**
	ER	0.53**	0.51***	**0.44*** **	**0.46*** **	0.42***	0.46***	0.50***	0.05	-0.45***	-0.26***	0.23***	-0.03	0.13***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; F18-25= Females 18-25; M18-25= Males 18-25; F26+= Females 26 and over; M26+=Males 26 and over; LSRP1: Levenson Self-Report Psychopathy Scale Primary; LSRP2; Levenson Self-Report Psychopathy Scale Secondary; NPI=Narcissistic Personality Inventory; NPI Ent = Entitlement; NPI Exp = Exploitativeness; ASP=Assessment of Sadistic Personality; SP=Successful Psychopathy; GE=Grandiose Entitlement; SC=Sadistic Cruelty; ER=Entitlement Rage. E= Extraversion; A = Agreeableness; C=Conscientiousness; NE=Negative Emotionality; OM=Open Mindedness; SE=Self Esteem

Self-esteem varied across groups, with both male groups having no significant relationship with it. The younger females showed a significant positive association with SP, whilst both female groups had the significant and positive relationship with SC and ER.

Test-Retest

Descriptive statistics and alpha reliabilities were calculated for total and subscale scores (Table 23). Outliers were present as assessed by Mahalonobis distance which revealed for Time 1; males 18-25; SC(1); Time 2; males 18-25; SC(1); Time 1; females 18-25; SC(2); SP(1); Time 2; females 18-25; SP(1); SC(2); Time 1; males 26 and Over; GE(1); SC(2); SP(2); Time 2; males 26 and Over; ER(3); Time 1; females 26 and Over; SC(3); Time 2; females 26 and over; SP(1); GE(1); ER(12); SC(3). Cronbach alpha scores displayed good temporal reliability. Robust correlations were calculated to assess temporal reliability (Table 23; Appendix 5.5)

Table 23

Descriptive Statistics for Test-Retest Sample and Test-Retest Correlations

Sample	SP Mean (SE)	a	GE Mean (SE)	a	SC Mean (SE)	a	ER Mean (SE)	a
Time One								
Males 18-25	46.53 (1.65)	.92	19.69 (0.87)	.88	12.61 (0.61)	.87	18.24 (0.74)	.87
Females 18-25	35.94 (1.64)	.95	19.66 (1.01)	.92	11.14 (0.60)	.92	18.85 (0.87)	.92
Males 26 and Over	41.56 (1.40)	.93	19.81 (0.80)	.91	11.53 (0.44)	.85	17.47 (0.69)	.91
Females 26 and Over	31.47 (1.09)	.92	16.51 (0.63)	.89	9.93 (0.35)	.89	14.79 (0.61)	.88
Whole Sample	38.95 (0.77)	.94	18.89 (0.42)	.90	11.29 (0.25)	.89	17.25 (0.37)	.90
Time Two								
Males 18-25	44.37 (1.77)	.94	20.80 (0.92)	.91	13.62 (0.68)	.88	17.81 (0.73)	.87
Females 18-25	36.86 (1.70)	.95	20.43 (0.97)	.93	11.24 (0.57)	.89	19.14 (0.83)	.92
Males 26 and Over	42.34 (1.52)	.95	21.27 (0.82)	.93	12.75 (0.50)	.89	16.82 (0.62)	.91
Females 26 and Over	33.01 (1.37)	.94	17.75 (0.76)	.91	10.13 (0.42)	.84	15.31 (0.68)	.90
Whole Sample	39.26 (0.82)	.95	20.07 (0.44)	.92	11.96 (0.28)	.89	17.15 (0.36)	.91
Test Re-Test	SPT1/SPT2		GE1/GE2		SCT1/SCT2		ERT1/ERT2	
Females 18-25	0.66		0.60		0.67		0.62	
Males 18-25	0.64		0.61		0.54		0.53	
Females Over 26	0.67		0.58		0.44		0.66	
Males Over 26	0.77		0.76		0.62		0.67	
Whole Sample	0.71		0.65		0.59		0.62	

Note: All correlations are significant at $p < .001$; SP=Successful Psychopathy; GE=Grandiose Entitlement; SC=Sadistic Cruelty; ER=Entitlement Rage; SPT1=Successful Psychopathy Time 1; SPT2=Successful Psychopathy Time 2; GE1=Grandiose Entitlement Time 1; GE2=Grandiose Entitlement Time 2;SCT1=Sadistic Cruelty Time 1; SCT2=Sadistic Cruelty Time 2; ERT1=Entitlement Rage Time 1; ERT2=Entitlement Rage Time 2.

The test-retest findings mainly showed high correlations between time one and time two. The medium correlation was found with females 26 and over in relation to sadistic cruelty. This is consistent with previous findings in relation to the DSHS, which have shown that this group scores the lowest intergroup mean score for this factor.

A confounding variable of this study was the global covid-19 lockdown and restriction period from March 2020. As participants were from multiple countries, to investigate whether this had caused a significant difference in responses between time two and time one, robust t-tests were conducted using the R package, 'WRS2' (Mair & Wilcox, 2019). Only one significant difference was found, which was for females 26 and over with entitlement rage $t(12)=3.74$, p =.00, d = 0.64), between time one and time two (Appendix 5.6).

6.4 Discussion

The goal of the present study was to provide further psychometric evidence for the four-factor structure of the DSHS, as well as examine the convergent and discriminant validity and test-retest reliability. The confirmatory factor analysis results corroborated those found in Chapter Five, confirming the structure across age and sex. In line with previous research, convergent validity was evidenced with the factors of the DSHS and their established counterpart measure across all groups (Jones & Paulhus, 2014; Maples et al., 2014), and indicates that the DSHS adequately covers the elements of the personality constructs. Discriminant validity was also present with domains of the Big Five and as predicted, significant negative relationships were found with Conscientiousness and Agreeableness, whilst Negative Emotionality and entitlement rage (ER), were positively and significantly correlated across groups. In relation to self-esteem, there were variations, with no significant relationship found for both male groups, whilst the younger females showed a significant positive association with successful psychopathy (SP), and both female groups had a significant and positive relationship with sadistic cruelty (SC), and entitlement rage (ER). The DSHS also displayed good temporal reliability. The evidence provided suggests that the scale is robust and measures the dark personality constructs across age and sex.

One issue which needs to be addressed as previously mentioned, is the confounding variable of the covid-19 lockdown period. The test-retest time frame

was to rule out true changes in personality over time and conducted within the same contextual timeframe (Watson, 2004). This life alteration of lockdown was experienced by a significant portion of the sample, as globally, the majority of countries were in lockdown or experiencing restrictions. However, it appears that lockdown did not impact the test-retest findings, which provides further confidence in the temporal reliability of the DSHS. Indeed, the only significant difference was found in the female over 26 sample in relation to entitlement rage. This suggests that for this group, context may play a salient role, whereby important life events influence affective traits (Vaidya, Gray, Haig & Watson, 2002). Research has shown that people who embody entitlement were less compliant with Covid regulations, considered the Covid narrative to be disproportionate and were not overly concerned about other people (Zitek & Schlund, 2020).

Convergent and discriminant validity with the Big Five mainly echoed previous studies. All groups showed significant negative associations with Agreeableness and Conscientiousness (e.g., Jonason & Webster, 2010; Fernández-del-Río et al., 2020). Disagreeableness is core to dark personality constructs (Austin et al., 2014), and highlights the antagonistic nature of interpersonal relationships, distrust of others, and the embodiment of callous and belittling attitudes (Soto, Kronauer et al., 2016). The negative relationship with Conscientiousness further emphasises the disdain held to conform to social norms (Soto & John, 2017). The only two constructs which showed significant negative relationships and thus divergence with Open Mindedness were successful psychopathy and sadistic cruelty. This association has previously been found with Machiavellianism and psychopathy (Fernández-del-Río et al., 2020; Kowalski et al., 2019). Open-Mindedness does not have a maladaptive counterpart (Hodson, Hogg & MacInnis, 2009), yet it taps into intrapsychic considerations which have macrosocial implications which are entwined within interpersonal relationships and societal attitudes.

In relation to the remaining Big Five traits, seeing how the relationships varied between age and sex shows the importance of making this distinction, rather than drawing conclusions from a total sample, or by separating groups by sex but not age. Divergent relationships were also found for self-esteem. Specifically, for females, the positive relationships with self-esteem, successful psychopathy, entitlement rage and sadistic cruelty, may point towards an entwinement of the traits. Entitlement can be viewed on a continuum with entitlement rage, where characteristics extend from

showing a willingness to exploit others for their own gain towards severe reactivity to criticism (Zeigler-Hill et al., 2008). Thereby, individuals displaying behaviours of entitlement as well as a lack of empathy, also experience powerful emotions of envy and enact aggressive behaviours due to their emotional instability. Extreme responses, such as a rage against criticism, which may encompass interpersonally sadistic behaviour, enacted without remorse or empathy, can preserve their inflated self-image (Cain et al., 2008).

Limitations

It was initially thought that the impact of Covid-19 may serve as a limitation to this study. However, this was shown to not have any significant affect. In retrospect, it would have been interesting to use the BFI-2 in the re-test study, to ascertain whether the lockdown period significantly affected the basic traits.

6.4.1 Conclusion

The DSHS is a valid and temporally reliable measure to use in dark personality research. This study has further highlighted the importance of separating samples by age and sex to reveal the similarities and differences in the nomological network to facilitate a more nuanced understanding of their associations with the dark facets of the DSHS. Although the confounding variable of covid-19 restrictions and lockdown during the test re-test period, showed a significant difference with entitlement rage in the female 26 and over group, it did not affect the temporal reliability. The DSHS can therefore be applied in a wide variety of studies, including longitudinal research. The DSHS measures the socially toxic aspects of dark personality traits in the general population with findings being a valid representation of the differences between age and sex.

Chapter Seven
General Discussion

7.1 Introduction

This final chapter reflects on the findings of the studies and the implications they have elicited. The main aim of this thesis was to develop a measure which was not only sex and age invariant but also psychometrically robust. The justification for developing a new dark personality scale was provided by the evidence from Chapter Four, which elucidated that the widely available Dark Tetrad (DT), measures were variant across and between age and sex. It further showed that the definitions of the constructs, as determined by the authors of the scales, were not wholly germane to general population samples across the specified groups. This was highlighted through the subsumption of Machiavellianism and psychopathy from the brief measures of the Dirty Dozen (Jonason & Webster, 2010), and to an extent, the Short Dark Triad (Jones & Paulhus, 2014). The findings corroborated previous research, as discussed in Chapter Two, and further intimated that these points contributed to the unstable structures historically found in studies. Consequently, Chapter Five focused on the development of the DSHS using a diverse range of psychometric methods to produce a scale which is sex and age invariant and measures dark personalities from an alternative viewpoint. Chapter Six provided evidence that the DSHS is a valid and temporally reliable measure to use in dark personality research. However, there are numerous considerations to discuss, which range from theoretical implications to psychometric considerations. The discussion will begin with the most relevant for the aim of this thesis:

What unique contribution does a new dark personality measure offer to the field of dark tetrad research?

7.2 The unique contribution of the DSHS to the field of dark personality research.

It may be questioned as to why a new scale is needed or is indeed necessary, when there are numerous measures across constructs which can already be used in dark personality research. The evidence provided in Chapter Four emphasised that scales which had been developed before more sophisticated methods of psychometric scrutiny were readily available were unstable in their structures. It further highlighted that the definitions of the constructs as determined by the authors

of the measures were not wholly germane across age and sex, which may have inadvertently perpetuated sex-stereotypical associations.

To emphasise the unique contribution the DSHS offers to the field and for comparison, as this thesis draws to a conclusion, a new scale which seeks to measure the Dark Tetrad has emerged – The Short Dark Tetrad (SD4; Paulhus, Buckels et al., 2020). The SD4 is an extension of the Short Dark Triad (SD3; Jones & Paulhus, 2014), which was investigated in Chapter Four and was found to be variant over sex and age, from both a top-down and bottom-up perspective. Criticisms by previous researchers can also be found in Chapter Two. Although some indicators have changed and some re-worded, the conceptualisations of the constructs have not. Machiavellianism is defined as one who is manipulative, callous and strategic, for psychopathy, the emphasis is on the secondary facet of antisocial behaviours and criminality whilst narcissism is represented by the grandiose variant. The everyday sadism facet mainly addresses vicarious and psychological pleasure taken from the suffering of others. Although some may consider the SD4 and DSHS to be analogous because they seek to measure the Dark Tetrad constructs, there are marked differences. These differences highlight a unique contribution of the DSHS.

The authors of the SD4 chose to restrict their samples to younger age groups, considering age a confounding variable, following their initial item pool reduction study with an MTurk sample (mean age = 35.2). Ensuing exploratory and confirmatory factor analyses were conducted with samples with mean ages ranging between 19.8 and 20.7. This would have no impact on the psychopathy facet as evidence has supported the decline of secondary psychopathy traits and behaviours with age in both institutionalised and general population samples (Gill & Crino, 2012; Harpur & Hare, 1994). Indeed, as the authors contended that student samples were the most usual source of data in personality research (Paulhus, Buckels et al., 2020), this implies it is not necessarily intended for wider population use.

One main aim of the DSHS was that it is invariant over age. There was substantial evidence for the exclusion of secondary psychopathy in general population sample measures (Boduszek, Dhingra, et al., 2016; Boduszek & Debowska, 2016; Cooke & Logan, 2015; Dinić, Wertag et al., 2020; Skeem & Cooke, 2010; Sellbom & Drislane, 2020), as discussed in Chapters Two and Four. Primary psychopathic traits have been shown to remain stable over the lifespan (Gill & Crino,

2012; Sturek et al., 2008), and the affective-interpersonal traits are put forth as, on the whole, constituting psychopathy (Poythress & Hall, 2011). Thus, the successful psychopathy factor which emerged from the data across age and sex (Chapter Five), is an entwinement of Cleckley's (1941), stated characteristics of primary psychopathy and those of Machiavellianism, as determined by experts (Christie & Geis, 1970; Miller, Hyatt et al., 2017; Mullins-Sweatt et al., 2010). It therefore moves away from the synthesis of the personality and behavioural theories determined for institutionalised males (Hare, 1980), and further significantly diverges from the facets represented in the SD4, providing a unique way of measuring these traits in the general population.

The narcissism factor of the SD4 does not address the vulnerable facet, but mainly pertains to the adaptive, grandiose characteristics found in the NPI (Rosenthal & Hooley, 2010). Although there are controversies and disagreements in the literature about the defining or fundamental features of the construct (Cain et al., 2008; Pincus & Lukowitsky, 2010), advancements have been made which suggest that the phenotypic structure of narcissism is attributed to a core of entitlement which manifests in grandiosity and/or vulnerability (Ackerman, Donnellan & Wright, 2019; Dinić, Sokolovksa & Tomašević, 2021). It has been argued that failing to distinguish between grandiose and vulnerable narcissism has led to the literature being disjointed and inconsistent (Miller, Lynam et al., 2017).

Consequently, there are no theoretical implications which arise from the narcissism constructs in the DSHS, as the grandiose entitlement and entitlement rage factors address the concerns put forth by the literature. Indeed, the DSHS is, to date, the first short DT measure which addresses the two forms of entitlement, providing a unique contribution. It is also worth noting, that these factors again, emerged from the data, as the only items which were homogenous across the specified groups, following analyses through the multiple methods used. The factors can be used as stand-alone measures, which was a benefit of using the Generalised Partial Credit Model (Hays et al., 2000; van der Linden, 2016).

In relation to everyday sadism, the authors of the SD4 (Paulhus, Buckels et al., 2020), argued that items which addressed direct sadism, such as 'I like to hurt people', interrelated with psychopathic aggression and would be likely to cross load over the two factors. Thus, they sought to minimise the indicators which tapped into direct, physical sadism. There are two points to make here. Firstly, in the sadistic

cruelty factor of the DSHS, the item, ‘I enjoy seeing people hurt’, could be considered to align with the callous nature of an individual who embodies psychopathy. Conversely, the item ‘People might describe me as mean and cruel ‘, from the successful psychopathy facet, may be considered more appropriate for the sadistic cruelty factor. In the development of the DSHS, there was no minimisation of characteristics from any of the included constructs. When the homogenous indicators across age and sex were incorporated into the final EGA network model in Chapter Five, the indicators mainly loaded onto their parent clusters across groups, with no pre-determination as to where they should load. Although there is a characteristic overlap between successful psychopathic features and those of the everyday sadist, each has its unique, nuanced motivations. The everyday sadist takes pleasure from the suffering of others, whilst those high in psychopathic traits may seek to harm others to accomplish goals (Carton & Egan, 2017; Hughes & Samuels, 2020). As with narcissism, there are no theoretical implications arising from the sadistic cruelty facet of the DSHS.

The purpose of comparison between the DSHS and the SD4 was not to suggest the SD4 is invalid or lacks robustness, indeed, there have not yet been any studies which have investigated its psychometric properties. The main point was to elucidate the differences and highlight how the DSHS offers an alternative and unique way of investigating dark personality traits, in a short measure of the constructs, within the general population. This provides justification as to why the DSHS is considered needed and/or necessary.

7.3 Theoretical Implications

The main theoretical deviation from existing measures, was in relation to Machiavellianism and psychopathy. As was discussed in Chapters Two and Four, and mentioned above, there was substantial evidence for the exclusion of secondary psychopathy from the DSHS (Boduszek, Dhingra, et al., 2016; Boduszek & Debowska, 2016; Cooke & Logan, 2015; Dinić, Wertag et al., 2020; Skeem & Cooke, 2010; Sellbom & Drislane, 2020). People across ages in the general population, who embody primary traits, display sufficient behavioural control and an ability to plan ahead. Potential risk taking is reflected upon and the consequences considered in advance of their actions. Thus, they operate successfully in society, abstaining from significant antisocial behaviour (Poythress & Hall, 2011).

Machiavellianism shares many overlapping characteristics with primary psychopathy as evidenced across numerous studies (Kavish et al., 2019; Maneiro et al., 2019; McHoskey et al., 1998; Miller, Hyatt et al., 2016; Persson. 2019; Persson, Kajonius & Garcia, 2017b; Siddiqi et al., 2020), which has led to the construct being considered redundant (Dinić, Wertag et al., 2020). Indeed, Machiavellianism may be considered a victim of the jangle fallacy (Gonzalez, MacKinnon & Muniz, 2020), whereby two constructs have been labelled differently but are empirically identical or identical enough to be subsumed under one construct (McHoskey et al., 1998; Miller, Hyatt et al., 2016; Persson. 2019). Previous research had shown that Machiavellianism and primary psychopathy formed a dark dyad in analysis (Maneiro et al.,2019; Rogoza & Cieciuch, 2018; Siddiqi et al., 2020), this was also mainly found across Chapters Four and Five. When synthesised, primary psychopathy and Machiavellianism were considered to represent the successful psychopath (Babiak et al., 2006; Gao & Raine, 2010; Mullins-Sweatt et al., 2010), in general population samples.

The theory supporting the successful psychopath, suggests that low levels of the behavioural inhibition system mediates the hypo-responsive fight-flight freeze response (Gray, 1987; Lykken, 1995). This stimulates self-control due to high levels of executive functioning, which is converse to the impulsiveness found in the traditional psychopathic profile (Babiak, Neumann & Hare, 2010; Babiak, Hare & McLaren, 2006; Board & Fritzon, 2005; Cleckley, 1941; Dutton, 2012; Gao & Raine, 2010; Lilienfeld, Waldman et al., 2012; Lykken, 1995). The integration of primary psychopathic and Machiavellian traits elucidate the characteristics of the successful psychopath, whereby they are manipulative and deceitful, assess risks through strategic planning and are callous, fearless and emotionally shallow. They lie and feel no remorse or guilt, and although they lack empathy, they can feign it, whilst embodying disdainful and apathetic attitudes to others. Outwardly, they are controlled, goal oriented, intelligent and charismatic (Babiak, Hare & McLaren, 2006; Babiak, Neumann & Hare, 2010; Board & Fritzon, 2005; Dutton, 2012; Fehr et al, 1992; Gao & Raine, 2010; Lilienfeld, Waldman et al., 2012; McHoskey, 1995; Palmen et al., 2018). These are addressed by the indicators in the successful psychopathy factor of the DSHS. This factor offers a unique contribution in being the first short DT measure to assess the successful psychopath.

7.4 Sex Stereotypical Associations

An underlying thread which has been interwoven through this thesis is how females have been portrayed in DT research. Researchers may consider intergroup mean score differences, which, in DT personality research, dependably shows that males achieve higher means. Yet, unless there is certainty that the measure used with the sample being compared is invariant, these results cannot necessarily be relied upon. However, these findings have often been used to maintain a sex stereotypical narrative.

Turning back to historical literature which addressed female stereotypes, Friedan (1963), argued that the influence of Freud and those who followed him had been provided with the momentum to preserve the perception that women are weak, in psychological research. She contended that 'instead of destroying the old prejudices that restricted women's lives, social science in America merely gave them new authority' (Friedan, 1963, p. 117). Women were depicted by psychologists as reliant and submissive, motivated by their desire to find a husband and have children (Weisstein, 1968/1971). This implied an underlying societal inevitability that women were devoid of power and status due to their inherent nature (Eagly, Eaton et al., 2012). Although these arguments are historical and many fields in psychology have addressed these issues, they have still been, to an extent, emphasised in dark personality research (Corry et al., 2015; Jonason & Webster, 2010). Indeed, there has been a noticeable absence within dark tetrad research, to dispel these assumptions, instead they are often supported and reinforced (Eagly, 1987; Hyde, 2005; Prentice & Carranza, 2002).

Yet, the stereotypical associations are also applied to men in dark personality research, where they are consistently determined as agentic (e.g., Jonason & Fletcher, 2018). However, sex stereotypes provide a distorted view of reality. It may therefore be questioned as to whether the implicit beliefs of some researchers in Western countries are conflating stereotypical associations mediated by the portrayal of sex differences from the measures, thus assuming sex role stereotypes (Baez et al., 2017). Assumptions such as these have salient societal implications whereby the impact not only affects those who appraise others based on stereotypical beliefs but also on those who are exposed to these discriminating suppositions (Ellemers, 2018).

The findings shown in Chapter Four and the development of the DSHS, have shown that males and females transcend their stereotypical associations, thereby refuting previous assertations (e.g., Corry et al., 2008; Tschanz et al.,1998). Moving forwards, it is hoped that the presumption that the male is the norm by which to measure females (Crawford & Marecek, 1989; Tiefer, 1987), will begin to dissipate and inter and intrasex distinctions will be acknowledged and enacted upon. However, this can only be achieved if it is accepted that males and females across ages, manifest idiosyncratic behaviours, which although go against the grain of homogeneity, is recognised as an objective actuality (Honey, 2017). The DSHS offers a unique contribution to the literature such that each item is sex and age invariant, thus, dispelling sex stereotypical associations which may inadvertently be present in other widely available measures.

7.5 Psychometric Considerations

A further unique contribution of this thesis to the literature are the methods used in the development of the DSHS. The decision to move away from traditional methods of exploratory factor analysis and other forms of classical test theories (CTT), was mediated by the discussions which put forth that the topology of the constructs under investigation can be analysed in a way that other statistical methods cannot provide (Hevey, 2018). The bottom-up and top-down analysis of the widely used scales, as discussed in Chapter Four, with samples split across and between groups, emphasised the limitations in previous scale development, whereby only CTT's of mainly exploratory and confirmatory factor analysis have been used, without consideration of group differences at an item level (e.g., Jonason & Webster, 2010), although this latter method may not have been readily available. The rigid CTT approach to scale development appeared to have become an unwritten law which researchers rarely challenged (Berchtold, 2016), until recently.

As an example, the Assessment of Sadistic Personality (ASP: Plouffe, Saklofske & Smith, 2017), was initially developed as a nine-item scale, using principal axis factoring. Yet in a recent revision (Plouffe, Kowalksi et al., 2020), using the graded response model and differential item functioning analysis, showed that one item should be discarded from the scale. This points to a salient consideration in relation to the widely used measures, such as the Mach IV (Christie & Geis, 1970) and the NPI (Raskin & Hall, 1979). Although criticisms have been consistently directed at them as discussed in Chapter Two, these measures should be situated

within their historical time frames, when technology and freely available statistical software were either not available or not easily accessible. The statistics software ‘R’, for example, was released in its first version to the public in 2000 (Peng, 2020). Although it is unclear when it became more widely used in psychological research, it now offers a diverse range of packages that can be used for multiple statistical methods and models, thereby addressing many of the limitations found in licenced packages such as SPSS. However, a limitation of R for some, may be the syntax used.

Thus, although the widely used measures may be criticised for their structural instability, sex and age invariance, they were developed with the tools that were available at the time. It is only with the benefit of hindsight and improved methods of psychometric scrutiny, that the limitations of these measures can be discussed. These discussions inform the future and it is only by taking on board criticisms of previous measures, that more psychometrically robust scales, mediated by more powerful tools, will emerge.

After extensive reading of the literature, the ethos of network analysis was considered an appropriate method to use to investigate the widely available DT measures and contribute to the development of the DSHS. Personality traits are viewed as systems which encompass multiple components that interact with each other (Hevey, 2018). Thus, exploratory graph analysis (EGA), elucidated the number of dimensions in the data by estimating the relationships between the observed variables rather than evaluating them as a function of a latent construct (Golino & Epskamp, 2017). By further incorporating the Generalised Partial Credit Model, into the development process, which determined the degree in which the scale items for each construct captures the latent trait levels in participants (Muraki & Muraki, 2016), the homogeneity of items across age and sex could be further elucidated. The final model was then assessed through confirmatory factor analysis, which following invariance testing showed a four-factor structure was invariant for all the groups. The methods used during the development of the DSHS are all discussed in depth in Chapter Three and shown in Chapters Four and Five.

To date, scales have not been developed with the symbiotic process used for the DSHS. It is therefore considered that this method of scale development provides a unique contribution to the literature.

7.6 Future Directions for Dark Tetrad Research

Research is now increasingly using network analysis to consider the dark personality traits as systems of complex structures, rather than latent constructs (Briganti & Linkowski, 2019; Christensen, Golino & Silvia, 2020; Trahair et al., 2020). Various methods of item response theories are also becoming more popular (Dinić, Bulut Allred et al., 2020; Garcia et al., 2018; Grigorias et al., 2020). Thus, there are a multitude of opportunities for expanding knowledge into how dark personalities across varying groups manifest (Nariman et al., 2020). The DSHS offers an alternative measure of the DT constructs. This will provide a divergent perspective from existing short measures, on how the attitudes and behaviours, which are under the umbrella of successful psychopathy, grandiose narcissism, sadistic cruelty and entitlement rage, interact with each other and in turn with external correlates. This will take research beyond mean score comparisons and factor models, to a more specific and nuanced approach, whereby the inter and intragroup differences across and between sex and age can be robustly investigated.

7.7 Limitations

The popular and freely available measures used in dark personality research were incorporated within Chapter Four, of this thesis. Although some studies utilise The Psychopathic Personality Inventory Revised (PPI-R) and the Self-Report Psychopathy Scale (SRP-4), there are charges for their use, which limits their accessibility. Thus, the decision was taken to use the freely available measures which are mostly used in academic research due to budgetary limitations. This decision also aligns with the usage of the DSHS, which will be freely available to researchers, once published.

The studies contained within this thesis solely relied on self-report data. It has been contended that this method is a valid and reliable way to investigate dark personality constructs (Jones and Paulhus, 2014). Although there may be a concern about how honestly participants report subjective assessments of themselves, it has been found that erroneous responding is not a problem in psychopathy or narcissism research (Ray, Hall et al., 2013; Sleep et al., 2017), as long as there are no incentives to skew responses (Kelsey et al., 2015).

Discussions have suggested that although measurement invariance analysis enables researchers to investigate whether a scale holds a similar structure across groups, which implies there is a safeguard against measurement bias, cultural

variations should be considered (Collison, South et al., 2020). Although the studies in this thesis have provided evidence that the DSHS is invariant across age and sex, cultural variations were not incorporated.

On reflection, it is seen as a limitation that the Big Five measure (Soto & John, 2017), was not included in the retest study (Chapter Six). Research across two time-points at the start of the coronavirus outbreak, has shown that there were small, decreased changes in Negative Emotionality amongst individuals who were not in quarantine or isolation. The Big Five constructs are considered to remain temporally stable and resilient to stressful, normative life events (Sutin et al., 2020). However, the global pandemic and the restrictions on personal freedoms may not be considered a normative life event. Their inclusion in the study would have provided further evidence as to the stability of the basic traits and any impact on their relationship with the DT constructs during significant life changes.

7.8 Conclusion

To many people, dark personalities are a fascinating concept, which are amplified by the media, through films and television programmes, which have thrived on characters who embody various dark personality traits (Jonason, Webster, Schmitt et al., 2012). The news brings into our homes, stories of people who cause psychological or physical harm to others (Hare, 1999). Yet these narratives which inform their watchers or readers about murderers or sadistic abusers align with the extreme end of the dark personality continuum. However, most people will have come across those who are manipulative and deceitful, who hold a sense of entitlement and possibly, individuals who take pleasure from causing or seeing others in pain. People may find it hard to conceive that there are those who operate within society, who enact these behaviours, strategically planning their moves in a callous and detached manner, without feeling empathy or remorse. Yet, these are the more common characteristics of individuals in society who embody dark personality traits. Highlighting how these traits manifest homogenously across age and sex, within the general population, was the main aim of this thesis and this has been achieved.

The contributions provided by these studies have served to draw attention to the importance of measuring samples by sex and age. The salience of this was highlighted through the criticisms of existing scales and the psychometric evidence which showed poor model fits, heterogenous factor structures across and between

groups, as well as a lack of invariance. Machiavellianism and psychopathy were also shown through to be parallel constructs and therefore subsumed under one construct to form the successful psychopath. Narcissism in the DSHS is represented by the core of the construct, which is a shift away from the adaptive characteristics prevalent in previous scales, as assessed by grandiose entitlement and entitlement rage. The sadistic cruelty factor elucidates how the everyday sadist takes intrinsic pleasure, mediated by cruel intentions, from imagining, witnessing or inflicting physical or psychological pain on others (Bulut, 2017; Smith, Powell et al., 2009; van Dijk & Ouwerkerk, 2014).

The DSHS retained some original items from the widely available scales used to measure the DT traits, that were explored in Chapter Four. The indicators which were not shown to be sex and age invariant in the Mokken Scale Analysis, were discarded. The resulting measure has preserved the most useful and relevant Indicators drawn from other DT scales, as well as those written specifically for everyday sadism, in a way that has allowed existing conceptualisations to be updated (e.g., through the use of questions around online trolling), that are both sex and age invariant. This has resulted in a revised measure, which can confidently be used to assess the DT traits effectively across age and sex.

The boundaries of sex stereotypical associations have been breached with the development of the DSHS. It has been suggested that research considers both intra and intersex and age similarities and differences. Although there may be idiosyncrasies, investigating participants from an item and/or network level as well as through classical test theories, will provide a far greater depth of knowledge in how dark personality traits manifest, instead of reducing males and females to mean score differences. Although age has been considered a confounding variable (Paulhus, Buckels et al., 2020), it is questioned as whether it is justifiable to exclude participants from scale development studies when they are outside of the student age range. The embodiment of dark personality traits may decrease to an extent with age (Gill & Crino, 2012), yet these traits are still present, as shown by the studies in this thesis. To marginalise a group to ensure the measure aligns with student populations because they are the most commonly used samples in psychology research, may be thought to undermine homogeneity. This is one main aim the development of the DSHS sought to avoid – and it succeeded.

To summarise, this thesis is considered to have made the following unique contributions to the field of dark personality research:

The DSHS is both sex and age invariant.

The retention of sex and age invariant indicators from the widely used measures, items from other DT measures and those written for everyday sadism enabled existing conceptualisations to be updated.

The DSHS is the first short DT measure to date, which contains two separate narcissism factors addressing the grandiose and vulnerable facets as they manifest invariantly in the general population.

The DSHS was developed with a unique method, using exploratory graph analysis, the generalised partial credit model and confirmatory factor analysis symbiotically.

The DSHS is the first DT measure to incorporate a factor which assesses the successful psychopath.

The DSHS dispels stereotypical assumptions which may have been inadvertently present in other widely available measures.

It is anticipated that the DSHS, which measures the Dark Tetrad, in an alternative manner, will provide the opportunity for research to move in diverse directions, with the knowledge that each factor is specific to the traits and behaviours they address, whilst also being sex and age invariant. Ultimately, it is hoped that the contributions made by these studies and the methods used, will add a unique contribution to dark personality research.

7.9 Final Reflections

Looking back on my PhD journey, it is difficult to disentwine my academic development from the personal changes it has brought. My desire to have an extensive understanding and knowledge of the dark personality constructs mediated my journey into the psychological domain and the focus became refined over time. At the onset of my PhD, I was armed with a basic knowledge of statistics and SPSS, which rapidly grew into an arsenal of statistical methods conducted in R. As there were no modules available which taught even the basics of R, I taught myself. For someone to whom statistics was an alien concept, to being the person, whereby a new dataset to investigate is akin to getting all your birthday and Christmas presents in one go, has been a remarkable change. The statistics went beyond numbers or plots, to telling a story. Who could ever have imagined that your heart would beat faster when waiting for confirmatory factor analysis results to appear on the screen? Or that when they did, you would cover your eyes and peak through, just in case the results were awful? Well, that has been me over this time. Needless to say, I now have an equivalent passion with statistics as I do with the dark personality traits.

The research has been fascinating, especially when you read a sentence or paragraph in a paper which puts a slightly different slant on a concept or theory. Instead of thinking inside a box or doing something in a particular way to conform with what has gone before, you begin to step outside that box, initially with trepidation. Yet, this seemed to be the moment when I realised that anything is possible, that if you are brave and truly believe in what you are doing, it can be achieved. My way of thinking changed, questions which I may never have considered before, enabled me to critically investigate research. This then extended into my life in general. I found myself listening more intently to conversations, having a more in-depth and nuanced understanding of what was being said and responding in a way I would never have done before my PhD.

I would honestly say that I am a different person today to the one I was when I began my PhD. I have developed both academically and personally and one would not have happened without the other. Although there are still self-doubts, I now have a level of confidence in both my knowledge and myself, which has only been made possible by this journey. Ultimately, for me, the memories of this journey will not be about reaching the top of the mountain but the long and sometimes arduous climb to get there.

References

References marked with an asterisk indicate studies included in the literature review.

Abell, L., & Brewer, G. (2018). Machiavellianism and Schadenfreude in Women's Friendships. *Psychological Reports*, *121*(5), 909-919.

Abrams, M. D., Ómarsdóttir, A. Ó., Björnsdóttir, M. D., Einarsdóttir, S., Martin, C., Carr, A., Brown, S.D., & Rector, C. (2013). Measurement invariance of the career indecision profile: United States and Iceland. *Journal of Career Assessment*, *21*(3), 469-482.

Ackerman, R. A., Corretti, C. A., & Carson, K. J. (2018). Psychometric properties of the narcissistic personality inventory. In *Handbook of Trait Narcissism* (pp. 125-131). Springer, Cham.

Ackerman, R. A., & Donnellan, M. B. (2013). Evaluating self-report measures of narcissistic entitlement. *Journal of Psychopathology and Behavioral Assessment*, *35*(4), 460-474.

Ackerman, R. A., Donnellan, M. B., & Robins, R. W. (2012). An item response theory analysis of the Narcissistic Personality Inventory. *Journal of Personality Assessment*, *94*(2), 141-155.

Ackerman, R. A., Donnellan, M. B., & Wright, A. G. (2019). Current conceptualizations of narcissism. *Current Opinion in Psychiatry*, *32*(1), 32-37.

*Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., & Kashy, D. A. (2011). What does the narcissistic personality inventory really measure? *Assessment*, *18*(1), 67-87.

*Ahmed, S. M. S., & Stewart, R. A. (1981). Factor analysis of the Machiavellian scale. *Social Behavior and Personality: An International Journal, 9*(1), 113-115.

Adedoyin, O. O., & Mokobi, T. (2013). Using IRT psychometric analysis in examining the quality of junior certificate mathematics multiple choice examination test items. *International Journal of Asian Social Science*, *3*(4), 992-1011.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270-301.

Aldridge, V. K., Dovey, T. M., & Wade, A. (2017). Assessing test-retest reliability of psychological measures: Persistent methodological problems. *European Psychologist, 22*(4), 207–218. https://doi.org/10.1027/1016-9040/a000298

Allen, E., & Seaman, C. (2007). *Likert scales and data analyses. Quality Progress, 40(7), 64–65.*

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders (3rd ed.)*. Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (3rd ed., revised)*. Washington, DC: Author

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders (4th ed., Text Revision*). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Publisher.

American Psychological Association. (2015). Definitions related to sexual orientation and gender diversity in APA documents. *APA Dictionary of Psychology*. https://www.apa.org/pi/lgbt/resources/sexuality-definitions.pdf

*Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism. *Journal of Research in Personality, 40*(4), 440-450.

An, X., & Yung, Y. F. (2014). Item response theory: What it is and how you can use the IRT procedure to apply it. *SAS Institute Inc. SAS364-2014*, *10*(4).

*Anestis, J. C., Caron, K. M., & Carbonell, J. L. (2011). Examining the impact of gender on the factor structure of the Psychopathic Personality Inventory—Revised. *Assessment, 18*(3), 340-349.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469.

Arnett, J. J., & Taber, S. (1994). Adolescence terminable and interminable: When does adolescence end? *Journal of Youth and Adolescence*, *23*(5), 517-537.

*Arseneault, R., & Catano, V. (2019). An extension of the dark triad and five-factor model to three Asian societies. *Asian Journal of Social Psychology*, *22*(4), 358-368.

Ashton, M. C., & Lee, K. (2005). Honesty-humility, the Big Five, and the five-factor model. *Journal of Personality, 73*(5), 1321-1354.

*Atari, M., & Chegeni, R. (2016). Assessment of dark personalities in Iran: Psychometric evaluation of the Farsi translation of the Short Dark Triad (SD3-F). *Personality and Individual Differences, 102,* 111-117.

Atlas, G. D., & Them, M. A. (2008). Narcissism and sensitivity to criticism: A preliminary investigation. *Current Psychology*, *27*(1), 62.

Austin, E. J., Saklofske, D. H., Smith, M., & Tohver, G. (2014). Associations of the managing the emotions of others (MEOS) scale with personality, the Dark Triad and trait EI. *Personality and Individual Differences*, *65*, 8-13.

Babiak, P. (1995). When psychopaths go to work: A case study of an industrial psychopath. *Applied Psychology*, *44*(2), 171-188.

Babiak, P., Hare, R. D., & McLaren, T. (2006). *Snakes in suits: When psychopaths go to work*. New York, NY: Regan Books.

Babiak, P., Neumann, C. S., & Hare, R. D. (2010). Corporate psychopathy: Talking the walk. *Behavioral Sciences & the Law*, *28*(2), 174-193.

Baez, S., Flichtentrei, D., Prats, M., Mastandueno, R., García, A. M., Cetkovich, M., & Ibáñez, A. (2017). Men, women… who cares? A population-based study on sex differences and gender roles in empathy and moral cognition. *PloS one*, *12*(6), e0179336.

Baker, F. B. (2001). *The basics of item response theory*. https://eric.ed.gov/?id=ED458219

*Barelds, D. P., & Dijkstra, P. (2010). Narcissistic personality inventory: structure of the adapted Dutch version. *Scandinavian Journal of Psychology, 51*(2), 132-138

Barlett, C. P., & Barlett, N. D. (2015). The young and the restless: Examining the relationships between age, emerging adulthood variables, and the Dark Triad. *Personality and Individual Differences, 86*, 20-24.

Baumeister, R. F., Bushman, B. J., & Campbell, W. K. (2000). Self-esteem, narcissism, and aggression: Does violence result from low self-esteem or from threatened egotism? *Current Directions in Psychological Science, 9*(1), 26-29.

Baumeister, R. F., & Campbell, W. K. (1999). The intrinsic appeal of evil: Sadism, sensational thrills, and threatened egotism. *Personality and Social Psychology Review*, *3*(3), 210-221.

*Beller, J., & Bosse, S. (2017). Machiavellianism has a dimensional latent structure: Results from taxometric analyses. *Personality and Individual Differences*, *113*, 57-62.

Bendig, A. W. (1953). The reliability of self-ratings as a function of the amount of verbal anchoring and of the number of categories on the scale. *Journal of Applied Psychology, 37,* 38–41. http://dx.doi.org/10.1037/h0 057911

*Benning, S. D., Patrick, C. J., Hicks, B. M., Blonigen, D. M., & Krueger, R. F. (2003). Factor structure of the psychopathic personality inventory: validity and implications for clinical assessment. *Psychological Assessment, 15*(3), 340.

Berchtold, A. (2016). Test–retest: agreement or reliability? *Methodological Innovations*, *9*, 2059799116672875.

Bertl, B., Pietschnig, J., Tran, U. S., Stieger, S., & Voracek, M. (2017). More or less than the sum of its parts? Mapping the Dark Triad of personality onto a single Dark Core. *Personality and Individual Differences*, *114*, 140-144.

Bialosiewicz, S., Murphy, K., & Berry, T. (2013). Do our measures measure up? The critical role of measurement invariance. *Demonstration Session at American Evaluation Association*.

Bishop, J., & Lane, R. C. (2002). The dynamics and dangers of entitlement. *Psychoanalytic Psychology*, *19*(4), 739.

Bizumic, B., & Monaghan, C. (2020). The HEXACO Model: Clinical Extensions and Universality. *European Journal of Personality.*34, 511-590.

Blinkhorn, V., Lyons, M., & Almond, L. (2015). The ultimate femme fatale? Narcissism predicts serious and aggressive sexually coercive behaviour in females. *Personality and Individual Differences, 87*, 219-223.

Board, B. J., & Fritzon, K. (2005). Disordered personalities at work. *Psychology, Crime & Law*, *11*(1), 17-32.

Boduszek, D., & Debowska, A. (2016). Critical evaluation of psychopathy measurement (PCL-R and SRP-III/SF) and recommendations for future research. *Journal of Criminal Justice*, *44*, 1-12.

Boduszek, D., Debowska, A., Dhingra, K., & DeLisi, M. (2016). Introduction and validation of Psychopathic Personality Traits Scale (PPTS) in a large prison sample. *Journal of Criminal Justice*, *46*, 9-17.

Boduszek, D., Dhingra, K., Hyland, P., & Debowska, A. (2016). A bifactorial solution to the Psychopathy Checklist: Screening Version in a sample of civil psychiatric patients. *Criminal Behaviour and Mental Health*, *26*(3), 174-185.

Boldero, J. M., Bell, R. C., & Davies, R. C. (2015). The structure of the narcissistic personality inventory with binary and rating scale items. *Journal of personality assessment*, *97*(6), 626-637.

Boomsma, A. (1985). Nonconvergence, Improper Solutions, and Starting Values in LISREL Maximum Likelihood Estimation. *Psychometrika, 50*, 229-242.

Bond, T. G., & Fox, C. M. (2007). Fundamental measurement in the human sciences. *Chicago, IL: Institute for Objective Measurement*.

Book, A., Visser, B. A., & Volk, A. A. (2015). Unpacking “evil”: Claiming the core of the Dark Triad. *Personality and Individual Differences*, *73*, 29-38.

Borgers, N., Sikkel, D., & Hox, J. (2004). Response effects in surveys on children and adolescents: The effect of number of response options, negative wording, and neutral midpoint. *Quality and Quantity, 38*(1), 17–33.

Bornstein, M. H., Jager, J., & Putnick, D. L. (2013). Sampling in developmental science: Situations, shortcomings, solutions, and standards. *Developmental Review*, *33*(4), 357–370.

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, *64*, 1089–1108. https://doi.org/10.1002/jclp.20503

Borsboom, D., & Cramer, A. O. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91-121.

Brailovskaia, J., Rohmann, E., Bierhoff, H. W., & Margraf, J. (2020). The anxious addictive narcissist: The relationship between grandiose and vulnerable narcissism, anxiety symptoms and Facebook Addiction. *Plos One*, *15*(11), e0241632.

*Braun, S., Kempenaers, C., Linkowski, P., & Loas, G. (2016). French adaptation of the narcissistic personality inventory in a Belgian French-speaking sample. *Frontiers in Psychology, 7*, 1980.

* Briganti, G., & Linkowski, P. (2020). Exploring network structure and central items of the Narcissistic Personality Inventory. *International Journal of Methods in Psychiatric Research*, *29*(1), e1810.

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

Brinkley, C. A., Diamond, P. M., Magaletta, P. R., & Heigel, C. P. (2008). Cross-validation of Levenson's Psychopathy Scale in a sample of federal female inmates. *Assessment, 15*(4), 464-482.

Brown, T. A. (2015*). Confirmatory factor analysis for applied research (*2nd ed*)*. New York, NY: Guilford.

*Brown, J., Akers, M. D., & Giacomino, D. E. (2013). Narcissism and Accounting Majors. *American Journal of Business Education, 6*(3), 375-384.

Brown, R. P., Budzek, K., & Tamborski, M. (2009). On the meaning and measure of narcissism. *Personality and Social Psychology Bulletin*, *35*(7), 951-964.

Brown, E. C., & Guy, R. F. (1983). The effects of sex and Machiavellianism on self-disclosure patterns. *Social Behavior and Personality: An International Journal, 11*(1), 93-96.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. *Handbook of Structural Equation Modeling,* 361-379.

Brown, M. F., Stanton, K., & Watson, D. (2020). Replicable Factor Structure and Correlates of an Alternate Version of the Narcissistic Personality Inventory. *Journal of Psychopathology and Behavioral Assessment*, *42*(1), 69-85.

Buckels, E. E., Jones, D. N., & Paulhus, D. L. (2013). Behavioral confirmation of everyday sadism. *Psychological Science, 24*(11), 2201-2209.

Buckels, E. E., Trapnell, P. D., Andjelovic, T., & Paulhus, D. L. (2019). Internet trolling and everyday sadism: Parallel effects on pain perception and moral judgment. *Journal of Personality*, *87*(2), 328-340.

Bulut, T. (2017). The concept of sadism in the current empirical literature. *UDK: 343*.

Bushman, B. J., Baumeister, R., Phillips, C., & Gilligan, J. (1999). Narcissism and self-esteem among violent offenders in a prison population. *Unpublished manuscript*.

Buss, D. M., & Duntley, J. D. (2008). Adaptations for exploitation. *Group Dynamics: Theory, Research, and Practice, 12*(1), 53.

Buss, A. R., & Royce, J. R. (1975). Detecting cross-cultural commonalties and differences: Intergroup factor analysis. *Psychological Bulletin*, *82,* 128–136.

Cain, N. M., Pincus, A. L., & Ansell, E. B. (2008). Narcissism at the crossroads: Phenotypic description of pathological narcissism across clinical theory, social/personality psychology, and psychiatric diagnosis. *Clinical Psychology Review*, *28*(4), 638-656.

Campbell, W. K., & Miller, J. D. (2011). *The handbook of narcissism and narcissistic personality disorder: Theoretical approaches, empirical findings, and treatments*. John Wiley & Sons.

*Carter, G. L., Campbell, A. C., Muncer, S., & Carter, K. A. (2015). A Mokken analysis of the Dark Triad 'Dirty Dozen': sex and age differences in scale structures, and issues with individual items. *Personality and Individual Differences, 83*, 185-191.

Carter, G. L., & Douglass, M. D. (2018). The ageing Narcissus: Just a myth? Narcissism moderates the age-loneliness relationship in older age. *Frontiers in Psychology, 9*, 1254.

Carton, H., & Egan, V. (2017). The dark triad and intimate partner violence. *Personality and Individual Differences*, *105*, 84-88.

CASP. (2017). *Critical Appraisal Skills Programme Cohort Study Checklist*. http://docs.wixstatic.com/ugd/dded87_5ad0ece77a3f4fc9bcd3665a7d1fa91f.pdf

Center for Countering Digital Hate. (2020). *Don't Feed the Trolls*. *A Practical Guide to Dealing With Hate on Social Media.* https://www.counterhate.com/dont-feed-the-trolls

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1-29.

Chang, L. (1994). A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity. *Applied Psychological Measurement, 18*(3), 205–215.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464-504.

Chen, J., & Chen, Z. (2008). Extended bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*, 759–771. https://doi.org/10.1093/biomet/asn034

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255.

Chiarotto, A., Bishop, A., Foster, N. E., Duncan, K., Afolabi, E., Ostelo, R. W., & Paap, M. C. (2018). Item response theory evaluation of the biomedical scale of the Pain Attitudes and Beliefs Scale. *PloS one*, *13*(9).

Chin, K., Atkinson, B. E., Raheb, H., Harris, E., & Vernon, P. A. (2017). The dark side of romantic jealousy. *Personality and Individual Differences*, *115*, 23-29.

*Chiorri, C., Garofalo, C., & Velotti, P. (2019). Does the dark triad manifest similarly in men and women? Measurement invariance of the dirty dozen across sex. *Current Psychology*, *38*(3), 659-675.

Chmielewski, M., & Watson, D. (2009). What is being assessed and why it matters: the impact of transient error on trait research. *Journal of Personality and Social Psychology*, *97*(1), 186.

Chou, C.P., & Bentler, P. M. (1995). Estimation and tests in structural equation modeling. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and application (pp. 37-55). Thousand Oaks, CA: Sage. 567

Christensen, A. P., Cotter, K. N., & Silvia, P. J. (2019). Reopening openness to experience: A network analysis of four openness to experience inventories. *Journal of Personality Assessment*, *101*(6), 574-588.

Christensen, A. P., & Golino, H. (2020). Estimating factors with psychometric networks: A Monte Carlo simulation comparing community detection algorithms. *PsyArXiv*. https://doi.org/10.31234/osf.io/hz89e

Christensen, A. P., Golino, H., & Silvia, P. J. (2020). A psychometric network perspective on the validity and validation of personality trait questionnaires. *European Journal of Personality*, https://doi.org/10.1002/per.2265

Christensen, A. P., Gross, G. M., Golino, H. F., Silvia, P. J., & Kwapil, T. R. (2019). Exploratory graph analysis of the multidimensional schizotypy scale. *Schizophrenia Research*, *206*, 43-51.

*Christian, E., & Sellbom, M. (2016). Development and validation of an expanded version of the three-factor Levenson Self-Report Psychopathy Scale. *Journal of Personality Assessment*, *98*(2), 155-168.

Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. NY, USA: Academic Press.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309.

*Clarke, I. E., Karlov, L., & Neale, N. J. (2015). The many faces of narcissism: Narcissism factors and their predictive utility. *Personality and Individual Differences, 81*, 90-95.

Cleckley, H. (1941). *The mask of sanity*. St. Louis, MO: Mosby.

Cloninger, C. R. (1978). The antisocial personality. *Hospital Practice*, *13*(8), 97-106.

Coid, J., & Ullrich, S. (2010). Antisocial personality disorder is on a continuum with psychopathy. *Comprehensive Psychiatry*, *51*(4), 426-433.

Collison, K. L., South, S., Vize, C. E., Miller, J. D., & Lynam, D. R. (2020). Exploring Gender Differences in Machiavellianism Using a Measurement Invariance Approach. *Journal of Personality Assessment*, 1-9.

Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology press.

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13-29.

Cooke, D. J., & Logan, C. (2015). Capturing clinical complexity: Towards a personality-oriented measure of psychopathy. *Journal of Criminal Justice, 43*(4), 262–273. https://doi.org/10.1016/j.jcrimjus.2015.04.004

Cooke, D. J., Michie, C., Hart, S. D., & Clark, D. A. (2004). Reconstructing psychopathy: Clarifying the significance of antisocial and socially deviant behavior in the diagnosis of psychopathic personality disorder. *Journal of Personality Disorders*, *18*(4), 337-357.

*Corral, S., & Calvete, E. (2000). Machiavellianism: Dimensionality of the Mach IV and its relation to self-monitoring in a Spanish sample. *The Spanish Journal of Psychology, 3,* 3-13.

*Corry, N., Merritt, R. D., Mrug, S., & Pamp, B. (2008). The factor structure of the Narcissistic Personality Inventory. *Journal of Personality Assessment*, *90*(6), 593-600.

Côté, J., & Bynner, J. M. (2008). Changes in the transition to adulthood in the UK and Canada: The role of structure and agency in emerging adulthood. *Journal of Youth Studies*, *11*(3), 251-268.

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., Kendler, K.S., & Borsboom, D. (2012). Measurable like temperature or mereological like flocking? On the nature of personality traits. *European Journal of Personality*, *26*, 451–459. https: //doi.org/10.1002/per.1879

Crawford, A. V., Green, S. B., Levy, R., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement*, *70*(6), 885-901.

Crawford, M., & Marecek, J. (1989). Psychology reconstructs the female 1968–1988. *Psychology of Women Quarterly*, *13*(2), 147-165.

Crede, M., & Harms, P. (2019). Questionable research practices when using confirmatory factor analysis. *Journal of Managerial Psychology*.

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*, *130*(3), 392.

Crowe, M. L., Lynam, D. R., & Miller, J. D. (2018). Uncovering the structure of agreeableness from self-report measures. *Journal of Personality*, *86*(5), 771-787.

Cunliffe, T., & Gacono, C. B. (2005). A Rorschach investigation of incarcerated female offenders with antisocial personality disorder. *International Journal of Offender Therapy and Comparative Criminology*, *49*(5), 530-546.

*Czibor, A., Szabo, Z. P., Jones, D. N., Zsido, A. N., Paal, T., Szijjarto, L., Carre, J.R., & Bereczkei, T. (2017). Male and female face of Machiavellianism: Opportunism or anxiety?. *Personality and Individual Differences, 117*, 221-229.

Dahling, J. J., Whitaker, B. G., & Levy, P. E. (2009). The development and validation of a new Machiavellianism scale. *Journal of Management*, *35*(2), 219-257.

Davies, J., & Hand, N. (2003, March). Sadistic interest: An exploratory study. In *Division of Forensic Psychology Annual Conference, Cambridge, England*.

DeShong, H. L., Helle, A. C., Lengel, G. J., Meyer, N., & Mullins-Sweatt, S. N. (2017). Facets of the dark triad: Utilizing the five-factor model to describe Machiavellianism. *Personality and Individual Differences*, *105*, 218-223.

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Thousand Oaks, CA: SAGE Publications.

de Vogel, V., Stam, J., Bouman, Y. H., Ter Horst, P., & Lancel, M. (2016). Violent women: A multicentre study into gender differences in forensic psychiatric patients. *The Journal of Forensic Psychiatry & Psychology*, *27*(2), 145-168.

Dickinson, K. A., & Pincus, A. L. (2003). Interpersonal analysis of grandiose and vulnerable narcissism. *Journal of Personality Disorders, 17*(3), 188-207.

Diefenbach, M.A., Weinstein, N.D., & O'Reilly, J. (1993). Scales for assessing perceptions of health hazard susceptibility. *Health Education Research, 8 2*, 181-92.

Dinić, B. M., Bulut Allred, T., Petrović, B., & Wertag, A. (2020). A test of three sadism measures: Short Sadistic Impulse Scale, varieties of sadistic tendencies, and assessment of sadistic personality. *Journal of Individual Differences.* Advance online publication.

*Dinić, B. M., Petrović, B., & Jonason, P. K. (2018). Serbian adaptations of the Dark Triad Dirty Dozen (DTDD) and Short Dark Triad (SD3). *Personality and Individual Differences, 134*, 321-328.

Dinić, B.M., Sokolovska, V. & Tomašević, A. (2021). The narcissism network and centrality of narcissism features. *Curr Psychol*. https://doi.org/10.1007/s12144-020-01250-w

Dinić, B. M., & Vujić, A. (2019). The Pathological Narcissism Inventory: Measurement invariance across Serbian and USA samples and further validation. *European Journal of Psychological Assessment*. Online first. https://doi.org/10.1027/1015- 5759/a000537

*Dinić, B. M., Wertag, A., Tomašević, A., & Sokolovska, V. (2020). Centrality and redundancy of the Dark Tetrad traits. *Personality and Individual Differences*, *155*, 109621.

Dodell-Feder, D., Saxena, A., Rutter, L., & Germine, L. (2019). The network structure of schizotypal personality traits in a population-based sample. *Schizophrenia Research*, *208*, 258-267.

Dong, Y., & Dumas, D. (2020). Are personality measures valid for different populations? A systematic review of measurement invariance across cultures, gender, and age. *Personality and Individual Differences*, *160*, 109956.

Donnellan, MB, Ackerman, RA, Wright, AGC. (2021). Narcissism in contemporary personality psychology. In O. P. John and R. W. Robins (Eds.), *Handbook of Personality: Theory and Research (4th Edition)*. New York: Guilford.

*Dotterer, H. L., Waller, R., Neumann, C. S., Shaw, D. S., Forbes, E. E., Hariri, A. R., & Hyde, L. W. (2017). Examining the factor structure of the self-report of psychopathy short-form across four young adult samples. *Assessment, 24*(8), 1062-1079.

*Douglas, H., Bore, M., & Munro, D. (2012). Distinguishing the dark triad: Evidence from the five-factor model and the Hogan development survey. *Psychology, 3*(3), 237-242.

Dutton, K. (2012). *The wisdom of psychopaths: Lessons in life from saints, spies and serial killers*. Random House. USA.

Eagly, A. H. (1987). *John M. MacEachran memorial lecture series; 1985. Sex differences in social behavior: A social-role interpretation.* Lawrence Erlbaum Associates, Inc.

Eagly, A. H., Eaton, A., Rose, S. M., Riger, S., & McHugh, M. C. (2012). Feminism and Psychology: Analysis of a Half-Century of Research on Women and Gender. American Psychologist. *Advance online publication*. doi: 10.1037/a0027260

Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, *111*(1), 3.

Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research*, *16*(1), 5.

*Eichenbaum, A. E., Marcus, D. K., & French, B. F. (2019). Item Response Theory analysis of the Psychopathic Personality Inventory–Revised. *Assessment*, *26*(6), 1046-1058.

Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology*, *69*, 275-298.

Embretson, S. E., & Reise, S. P. (2000). Multivariate Applications Books Series. *Item response theory for psychologists. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers*.

*Emmons, R. A. (1984). Factor analysis and construct validity of the narcissistic personality inventory. *Journal of Personality Assessment, 48*(3), 291-300.

Emmons, R. A. (1987). Narcissism: Theory and measurement. *Journal of Personality and Social Psychology, 52*(1), 11.

Emons, W. H., Sijtsma, K., & Pedersen, S. S. (2012). Dimensionality of the Hospital Anxiety and Depression Scale (HADS) in cardiac patients: comparison of Mokken scale analysis and factor analysis. *Assessment, 19*(3), 337-353.

Epskamp, S., Borsboom, D., & Fried, E. I. (2017). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195-212.

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*, 617–634. https://doi.org/10.1037/met0000167

Fehr, B., Samson, D., & Paulhus, D. L. (1992). *The construct of Machiavellianism: Twenty years later.* In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in Personality Assessment, Vol. 9* (p. 77–116). Lawrence Erlbaum Associates, Inc.

Fernández-del-Río, E., Ramos-Villagrasa, P. J., & Barrada, J. R. (2020). Bad guys perform better? The incremental predictive validity of the Dark Tetrad over Big Five and Honesty-Humility. *Personality and Individual Differences, 154*, 109700.

Fiester, S. J., & Gay, M. (1991). Sadistic personality disorder: A review of data and recommendations for DSM-IV. *Journal of Personality Disorders*, *5*(4), 376-385.

Finn, J. A., Ben-Porath, Y. S., & Tellegen, A. (2015). Dichotomous versus polytomous response options in psychopathology assessment: Method or meaningful variance? *Psychological Assessment, 27,* 184 –193. http://dx.doi.org/10.1037/pas0000044

Finstad, K. (2010). Response interpolation and scale sensitivity: Evidence against 5-point scales. *Journal of Usability Studies*, *5*(3), 104-110.

Fleeson, W. (2020). Live Simply So That Others May Simply Live: Trying to Get Clarity on the Meaning of Honesty/Humility. A commentary to appear with: Bizumic, B., & Monaghan, C. (2020). The HEXACO Model: Clinical Extensions and Universality. *European Journal of Personality.*34, 511-590.

Forouzan, E., & Cooke, D. J. (2005). Figuring out la femme fatale: Conceptual and assessment issues concerning psychopathy in females. *Behavioral Sciences & the Law*, *23*(6), 765-778.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *3–5*, 75–174. https://doi.org/10.1037/met0000255

Foulkes, L. (2019). Sadism: Review of an elusive construct. *Personality and Individual Differences*, *151*, 109500.

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology*, *78*(2), 350.

Freis, S. D., & Hansen-Brown, A. A. (2021). Justifications of entitlement in grandiose and vulnerable narcissism: The roles of injustice and superiority. *Personality and Individual Differences*, *168*, 110345.

Freud, S. (1931). *Libidinal types. Collected papers (Vol. 5).* London: Hogarth Press.

Freud, S. (1950). Libidinal types. In J. Strachey (Ed. and Trans.), The standard edition of the complete psychological works of Sigmund Freud (Vol. 21, pp. 217– 220). London: Hogarth Press. (Original work published 1931).

Friedan, B. (1963). The feminine mystique. New York, NY: Norton

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432-441.

Furlong, A. (2009). *Handbook of Youth and Young Adulthood: New Perspectives and Agendas*. Routledge.

Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The Dark Triad of personality: A 10 year review. *Social and Personality Psychology Compass*, *7*(3), 199-216.

Furr, M. (2011). *Scale construction and psychometrics for social and personality psychology*. SAGE Publications Ltd.

*Gamache, D., Savard, C., & Maheux-Caron, V. (2018). French adaptation of the Short Dark Triad: Psychometric properties and a head-to-head comparison with the Dirty Dozen. *Personality and Individual Differences, 122,* 164-170.

Gao, Y., & Raine, A. (2010). Successful and unsuccessful psychopaths: A neurobiological model. *Behavioral Sciences & the Law*, *28*(2), 194-210.

* Garcia, D., Persson, B. N., Al Nima, A., Brulin, J. G., Rapp-Ricciardi, M., & Kajonius, P. J. (2018). IRT analyses of the Swedish Dark Triad Dirty Dozen. *Heliyon*, *4*(3), e00569.

Garland, R. (1991). The mid-point on a rating scale: Is it desirable. *Marketing Bulletin, 2,* 66–70.

*Garofalo, C., Noteborn, M. G., Sellbom, M., & Bogaerts, S. (2019). Factor structure and construct validity of the Levenson Self-Report Psychopathy Scale (LSRP): A replication and extension in Dutch nonclinical participants. *Journal of Personality Assessment*, *101*(5), 481-492.

Gill, D. J., & Crino, R. D. (2012). The relationship between psychopathy and age in a non-clinical community convenience sample. *Psychiatry, Psychology and Law*, *19*(4), 547-557.

Gladden, P. R., Figueredo, A. J., & Jacobs, W. J. (2009). Life history strategy, psychopathic attitudes, personality, and general intelligence. *Personality and Individual Differences*, *46*(3), 270-275.

Glenn, A. L., & Sellbom, M. (2015). Theoretical and empirical concerns regarding the dark triad as a construct. *Journal of Personality Disorders*, *29*(3), 360-377.

Gluck, M., Heesacker, M., & Choi, H. D. (2020). How much of the dark triad is accounted for by sexism? *Personality and Individual Differences*, *154*, 109728.

Goldberg, L. R. (2006). Doing it all bass-ackwards: The development of hierarchical factor structures from the top down. *Journal of Research in Personality*, *40*(4), 347-358.

Golino, H., & Christensen, A. P. (2020). EGAnet: Exploratory Graph Analysis – A framework for estimating the number of dimensions in multivariate data using network psychometrics. Retrieved from https://cran.r-project.org/package=EGAnet

Golino, H. F., & Demetriou, A. (2017). Estimating the dimensionality of intelligence like data using Exploratory Graph Analysis. *Intelligence*, *62*, 54-70.

Golino, H., & Epskamp, S. (2017). Exploratory Graph Analysis: A new approach for estimating the number of dimensions in psychological research. *PloS ONE*, *12*, e0174035. https://doi.org/10.1371/journal.pone.0174035

Golino, H., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Sadana, R., … Martinez-Molina, A. (2020). Investigating the performance of Exploratory Graph Analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychological Methods*, *25*, 292–320. https://doi.org/10.1037/met0000255

Gomez, R. (2008). Parent ratings of the ADHD items of the disruptive behavior rating scale: Analyses of their IRT properties based on the generalized partial credit model. *Personality and Individual Differences*, *45*(2), 181-186.

Gonzalez, J. M., & Greitemeyer, T. (2018). The relationship between everyday sadism, violent video game play, and fascination with weapons. *Personality and Individual Differences*, *124*, 51-53.

Gonzalez, O., MacKinnon, D. P., & Muniz, F. B. (2020). Extrinsic Convergent Validity Evidence to Prevent Jingle and Jangle Fallacies. *Multivariate Behavioral Research*, 1-17.

*Gordts, S., Uzieblo, K., Neumann, C., Van den Bussche, E., & Rossi, G. (2017). Validity of the Self-Report Psychopathy Scales (SRP-III full and short versions) in a community sample. *Assessment, 24*(3), 308-325.

Gore, W. L., & Widiger, T. A. (2016). Fluctuation between grandiose and vulnerable narcissism. *Personality Disorders: Theory, Research, and Treatment*, *7*(4), 363.

Gosling, S. D., Sandy, C. J., John, O. P., & Potter, J. (2010). Wired but not WEIRD: The promise of the Internet in reaching more diverse samples. *Behavioral and Brain Sciences*, *33*(2-3), 94-95.

Götz, F. M., Bleidorn, W., & Rentfrow, P. J. (2020). Age differences in Machiavellianism across the life span: Evidence from a large-scale cross-sectional study. *Journal of Personality*, *88*(5), 978-992.

Gov.UK (2018). Data protection. https://www.gov.uk/data-protection

Gray, J. A. (1987). *The psychology of fear and stress* (Vol. 5). CUP Archive.

Green, A., MacLean, R., & Charles, K. (2020). Unmasking gender differences in narcissism within intimate partner violence. *Personality and Individual Differences*, *167*, 110247.

Green, S. B., Redell, N., Thompson, M. S., & Levy, R. (2016). Accuracy of revised and traditional parallel analyses for assessing dimensionality with binary data. *Educational and Psychological Measurement*, *76*(1), 5-21.

*Grigoraș, M., Butucescu, A., Miulescu, A., Opariuc-Dan, C., & Iliescu, D. (2020). The Measurement Invariance of the Short Dark Triad. Implications for High-and-Low-Stakes Contexts. *Journal of Individual Differences*, *41,* 207-218.

Grijalva, E., Newman, D. A., Tay, L., Donnellan, M. B., Harms, P. D., Robins, R. W., & Yan, T. (2015). Gender differences in narcissism: A meta-analytic review. *Psychological Bulletin*, *141*(2), 261.

*Gummelt, H. D., Anestis, J. C., & Carbonell, J. L. (2012). Examining the Levenson Self Report Psychopathy Scale using a graded response model. *Personality and Individual Differences, 53*(8), 1002-1006.

Gysi, D. M., Voigt, A., de Miranda Fragoso, T., Almaas, E., & Nowick, K. (2018). wTO: an R package for computing weighted topological overlap and a consensus network with integrated visualization tool. *BMC Bioinformatics*, *19*(1), 392.

Halamová, J., Kanovský, M., Gilbert, P., Troop, N. A., Zuroff, D. C., Petrocchi, N., ... & Matos, M. (2019). Multiple Group IRT Measurement Invariance Analysis of the Forms of Self-Criticising/Attacking and Self-Reassuring Scale in Thirteen International Samples. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *37*(4), 411-444.

Hall, J. R., & Benning, S. D. (2006). The “successful” psychopath. *Handbook of Psychopathy*, 459-478.

Hallquist, M., Wright, A. C. G., & Molenaar, P. C. M. (2019). Problems with centrality measures in psychopathology symptom networks: Why network psychometrics cannot escape psychometric theory. *Multivariate Behavioral Research*. https://doi.org/10. 1080/00273171.2019.1640103

Hambleton, R. K., & Jones, R. W. (1993). An NCME instructional module on: Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, *12*(3), 38-47.

Hanel, P. H., & Vione, K. C. (2016). Do student samples provide an accurate estimate of the general public?. *PloS one*, *11*(12).

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences, 1*(2), 111-119.

Hare, R. D. (1999). *Without conscience: The disturbing world of the psychopaths among us.* The Guilford Press.

Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The revised psychopathy checklist: reliability and factor

structure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *2*(3), 338.

Hare, R. D., Hart, S. D., & Harpur, T. J. (1991). Psychopathy and the DSM-IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology*, *100*(3), 391.

Hare, R. D., & Neumann, C. S. (2006). The PCL-R assessment of psychopathy: Development, structural properties, and new directions. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 58–88). New York: Guilford.

Harpur, T. J., & Hare, R. D. (1994). Assessment of psychopathy as a function of age. *Journal of Abnormal Psychology*, *103*(4), 604.

Harpur, T. J., Hakstian, A. R., & Hare, R. D. (1988). Factor structure of the Psychopathy Checklist. *Journal of Consulting and Clinical Psychology*, *56*(5), 741.

Hart, W., Richardson, K., Tortoriello, G., & Tullett, A. (2017). Strategically out of control: A self-presentational conceptualization of narcissism and low self-control. *Personality and Individual Differences, 114*, 103-107.

*Hauck-Filho, N., & Teixeira, M. A. P. (2014). Revisiting the psychometric properties of the Levenson self-report psychopathy scale. *Journal of Personality Assessment, 96*(4), 459-464.

Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical Care*, *38*(9 Suppl), II28.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). Most people are not WEIRD. *Nature*, *466*(7302), 29-29.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2–3), 61-83.

Hevey, D. (2018). Network analysis: a brief overview and tutorial. *Health Psychology and Behavioral Medicine*, *6*(1), 301-328.

Higgins, J. P. T., & Green, S. (2014). *Cochrane handbook for systematic reviews of interventions Version 5.1. 0* [updated March 2011]. 2011. The Cochrane Collaboration.

Hilbert, S., Küchenhoff, H., Sarubin, N., Nakagawa, T. T., & Bühner, M. (2016). The influence of the response format in a personality questionnaire: An analysis of a dichotomous, a Likert-type, and a visual analogue scale. *TPM-Testing, Psychometrics, Methodology in Applied Psychology, 23,* 3–24.

Hirschfeld, G., & Von Brachel, R. (2014). Improving Multiple-Group confirmatory factor analysis in R–A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research, and Evaluation*, *19*(1), 7.

Hodson, G., Hogg, S. M., & MacInnis, C. C. (2009). The role of “dark personalities”(narcissism, Machiavellianism, psychopathy), Big Five personality factors, and ideology in explaining prejudice. *Journal of Research in Personality*, *43*(4), 686-690.

Hoertel, N., Peyre, H., Lavaud, P., Blanco, C., Guerin-Langlois, C., René, M., Schuster, J.P., Lemogne, C., Delorme, R., & Limosin, F. (2018). Examining sex differences in DSM-IV-TR narcissistic personality disorder symptom expression using Item Response Theory (IRT). *Psychiatry Research*, *260*, 500-507.

Honey, P. L. (2017). *The element of surprise: Women of the Dark Triad.* In M. L. Fisher (Ed.), *Oxford library of psychology. The Oxford handbook of women and competition* (p. 147–165). Oxford University Press.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Ageing Research*, *18*(3), 117-144.

Hoyle, R. H., & Duvall, J. L. (2004). Determining the number of factors in exploratory and confirmatory factor analysis. *Handbook of Quantitative Methodology for the Social Sciences*, 301-315.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424.

Hughes, S., & Samuels, H. (2020). Dark desires: The Dark Tetrad and relationship control. *Personality and Individual Differences*, 110548.

*Hunter, J. E., Gerbing, D. W., & Boster, F. J. (1982). Machiavellian beliefs and personality: Construct invalidity of the Machiavellianism dimension. *Journal of Personality and Social Psychology*, *43*(6), 1293.

Hussein, M. H. (2010). Evaluation of an Arabic Version of Children's Self-Report Social Skills Scale (CS 4) Based on Item Response Theory. *Online Submission*. fhttps://www.semanticscholar.org/paper/Social-Skills-Scale-1-Evaluation-of-an-Arabic-of-'-Hussein/99692610c93768cb729f8df17dea9d451bb06bbb

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581.

Irvine, S. H. (1969). Contributions of ability and attainment testing in Africa to a general theory of intellect. *Journal of Biosocial Science, 1,* 91–102.

Ishikawa, S. S., Raine, A., Lencz, T., Bihrle, S., & Lacasse, L. (2001). Autonomic stress reactivity and executive functions in successful and unsuccessful criminal psychopaths from the community. *Journal of Abnormal Psychology*, *110*(3), 423.

*Jackson, L. A., Ervin, K. S., & Hodge, C. N. (1992). Narcissism and body image. *Journal of Research in Personality*, *26*(4), 357-370.

Jackson, R. L., Rogers, R., Neumann, C. S., & Lambert, P. L. (2002). Psychopathy in female offenders: An investigation of its underlying dimensions. *Criminal Justice and Behavior*, *29*(6), 692-704.

James, S., Kavanagh, P. S., Jonason, P. K., Chonody, J. M., & Scrutton, H. E. (2014). The Dark Triad, schadenfreude, and sensational interests: Dark personalities, dark emotions, and dark behaviors. *Personality and Individual Differences, 68,* 211–216. https://doi.org/10.1016/j.paid.2014.04.020

Jauk, E., Weigle, E., Lehmann, K., Benedek, M., & Neubauer, A. C. (2017). The relationship between grandiose and vulnerable (hypersensitive) narcissism. *Frontiers in Psychology*, *8*, 1600.

Johns, R. (2005). One size doesn't fit all: Selecting response scales for attitude items. *Journal of Elections, Public Opinion & Parties, 15*, 237-264. doi: 10.1080/13689880500178849

Jonason, P.K., & Davis, M.D. (2018). A gender role view of the Dark Triad traits. *Personality and Individual Differences, 125*, 102-105.

Jonason, P. K., & Fletcher, S. A. (2018). Agentic and communal behavioral biases in the Dark Triad traits. *Personality and Individual Differences*, *130*, 76-82.

Jonason, P.K, Foster, J.D., Kavanagh, P.S., Gouveia, V.V. & Birkás, B. (2018). Basic Values and the Dark Triad Traits. *Journal of Individual Differences,* 39(4), 220-228.

Jonason, P. K., Kaufman, S. B., Webster, G. D., & Geher, G. (2013). What lies beneath the Dark Triad Dirty Dozen: Varied relations with the Big Five. *Individual Differences Research, 11*(2).

Jonason, P. K., Koenig, B. L., & Tost, J. (2010). Living a fast life: The Dark Triad and life history theory. *Human Nature, 21*(4), 428–442.

Jonason, P. K., & Krause, L. (2013). The emotional deficits associated with the Dark Triad traits: Cognitive empathy, affective empathy, and alexithymia. *Personality and Individual Differences, 55(*5), 532-537.

Jonason, P. K., Li, N. P., & Teicher, E. A. (2010). Who is James Bond? The Dark Triad as an agentic social style. *Individual Differences Research*, *8*(2), 111.

Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment, 22*(2), 420.

Jonason, P. K., Webster, G. D., Schmitt, D. P., Li, N. P., & Crysel, L. (2012). The antihero in popular culture: Life history theory and the dark triad personality traits. *Review of General Psychology*, *16*(2), 192-199

Jones, D. N. (2016). *The nature of Machiavellianism: Distinct patterns of misbehavior.* In V. Zeigler-Hill & D. K. Marcus (Eds.), *The dark side of personality: Science and practice in social, personality, and clinical psychology* (p. 87–107). American Psychological Association. https://doi.org/10.1037/14854-005

Jones, D. N., & Figueredo, A. J. (2013). The core of darkness: Uncovering the heart of the Dark Triad. *European Journal of Personality*, *27*(6), 521-531.

Jones, D.N., & Paulhus, D.L. (2009). *Machiavellianism.* In M.R. Leary & R.H. Hoyle (Eds.) *Individual differences in social behaviour* (p.93-108). New York: Guilford.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment*, *21*(1), 28-41.

*Kajonius, P. J., Persson, B. N., Rosenberg, P., & Garcia, D. (2016). The (mis) measurement of the Dark Triad Dirty Dozen: exploitation at the core of the scale. *PeerJ*, *4*, e1748.

Kaminer, D., & Stein, D. J. (2001). Sadistic personality disorder in perpetrators of human rights abuses: A South African case study. *Journal of Personality Disorders*, *15*(6), 475-486.

Karpman, B. (1948). The myth of the psychopathic personality. *American Journal of Psychiatry, 104*(9), 523-534.

Kavish, N., Jones, M. A., Rock, R. C., Johnson, A. K., & Anderson, J. L. (2019). On the Overlap between Psychopathic Traits and Machiavellianism in a Forensic

Population. *Journal of Psychopathology and Behavioral Assessment*, *41*(2), 198-207

Kayrouz, R., Dear, B. F., Karin, E., & Titov, N. (2016). Facebook as an effective recruitment strategy for mental health research of hard to reach populations. *Internet Interventions*, *4*(Part 1), 1–10.

*Kawamoto, T., Shimotsukasa, T., & Oshio, A. (2020). Cross-sectional age differences in the Dark Triad traits in two Japanese samples. *Psychology and Aging, 35*(1), 91–96. https://doi.org/10.1037/pag0000399

Kealy, D., Laverdière, O., & Pincus, A. L. (2020). Pathological Narcissism and Symptoms of Major Depressive Disorder Among Psychiatric Outpatients: The Mediating Role of Impaired Emotional Processing. *The Journal of Nervous and Mental Disease*, *208*(2), 161-164.

Kean, J., & Reilly, J. (2014). Item response theory. *Handbook for Clinical Research: Design, Statistics and Implementation*, 195-198.

Keith, T. Z., Caemmerer, J. M., & Reynolds, M. R. (2016). Comparison of methods for factor extraction for cognitive test-like data: Which overfactor, which underfactor?. *Intelligence*, *54*, 37-54.

Kelsey, K. R., Rogers, R., & Robinson, E. V. (2015). Self-report measures of psychopathy: What is their role in forensic assessments? *Journal of Psychopathology and Behavioral Assessment*, *37*(3), 380-391.

Klein Tuente, S., Vogel, V. D., & Stam, J. (2014). Exploring the criminal behavior of women with psychopathy: Results from a multicenter study into psychopathy and violent offending in female forensic psychiatric patients. *International Journal of Forensic Mental Health*, *13*(4), 311-322.

*Klimstra, T. A., Jeronimus, B. F., Sijtsema, J. J., & Denissen, J. J. (2020). The unfolding dark side: Age trends in dark personality features. *Journal of Research in Personality*, *85*, 103915.

Kline, R. B. (2015). *Principles and practice of structural equation modelling (4th ed.).* New York, NY: Guilford Press.

Klockars, A. J., & Yamagishi, M. (1988). The influence of labels and positions in rating scales. *Journal of Educational Measurement, 25*(2), 85-96.

Kowalski, C. M., Di Pierro, R., Plouffe, R. A., Rogoza, R., & Saklofske, D. H. (2019). Enthusiastic acts of evil: The assessment of sadistic personality in polish and Italian populations. *Journal of Personality Assessment,* 1-11.

Kowalski, C. M., Vernon, P. A., & Schermer, J. A. (2019). The Dark Triad and facets of personality. *Current Psychology*, 1-12.

Krizan, Z., & Herlache, A. D. (2018). The narcissism spectrum model: A synthetic view of narcissistic personality. *Personality and Social Psychology Review*, *22*(1), 3-31.

Krosnick, J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Hanemann, W., Kopp, R. J., Mitchell, R.C., Presser, S., Rudd, P.A., Smith, V.K., Moody, W.R., Green, M.C. & Conaway, M. (2002). The impact of 'no opinion' response options on data quality: Non-attitude reduction or an invitation to satisfice? *Public Opinion Quarterly, 66,* 371-403.

*Kubarych, T. S., Deary, I. J., & Austin, E. J. (2004). The Narcissistic Personality Inventory: Factor structure in a non-clinical sample. *Personality and Individual Differences, 36*(4), 857-872.

Kulas, J. T., & Stachowski, A. A. (2013). Respondent rationale for neither agreeing nor disagreeing: Person and item contributors to middle cate- gory endorsement intent on Likert personality indicators. *Journal of Research in Personality, 47, 254–262. http://dx.doi.org/10.1016/j.jrp .2013.01.014*

Kushari, U., Sidhu, S., Tahlani, K., & Chaterjee, S. (2017). A study of attitude towards women in relation to narcissism, ethics and personality. *Indian Journal of Community Psychology, 13 (2),* 286-307.

Lancaster, S. L., Melka, S. E., & Rodriguez, B. F. (2009). A factor analytic comparison of five models of PTSD symptoms. *Journal of Anxiety Disorders*, *23*(2), 269-274.

Langer, M. M., Hill, C. D., Thissen, D., Burwinkle, T. M., Varni, J. W., & DeWalt, D. A. (2008). Item response theory detected differential item functioning between healthy and ill children in quality-of-life measures. *Journal of Clinical Epidemiology*, *61*(3), 268-276.

Lauritzen, S. L. (1996). *Graphical models.* Oxford, UK: Clarendon Press.

Lehmann, A., & Ittel, A. (2012). Aggressive behavior and measurement of psychopathy in female inmates of German prisons—A preliminary study. *International Journal of Law and Psychiatry*, *35*(3), 190-197.

*Lester, W. S., Salekin, R. T., & Sellbom, M. (2013). The SRP-II as a rich source of data on the psychopathic personality. *Psychological Assessment, 25*(1), 32.

Leung, S.O. (2011). A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point Likert scales. *Journal of Social Service Research, 37*(4), 412-421.

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology, 68*(1), 151.

Lewis, B. R. (1993). Service quality measurement. *Marketing Intelligence & Planning*, *11*(4), 4-12.

Ligtvoet, R., van der Ark, L. A., Bergsma, W. P., & Sijtsma, K. (2011). Polytomous latent scales for the investigation of the ordering of items. *Psychometrika, 76*(2), 200–216.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives de Psychologie, 140,* 5–53.

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal population. *Journal of Personality Assessment*, *66*(3), 488-524.

Lilienfeld, S. O., Waldman, I. D., Landfield, K., Watts, A. L., Rubenzer, S., & Faschingbauer, T. R. (2012). Fearless dominance and the US presidency: Implications of psychopathic personality traits for successful and unsuccessful political leadership. *Journal of Personality and Social Psychology*, *103*(3), 489.

Lilienfeld, S. O., Watts, A. L., & Smith, S. F. (2015). Successful psychopathy: A scientific status report. *Current Directions in Psychological Science, 24*(4), 298-303.

Logan, C., & Blackburn, R. (2009). Mental disorder in violent women in secure settings: Potential relevance to risk for future violence. *International Journal of Law and Psychiatry*, *32*(1), 31-38.

Lui, L. F., Sassenrath, C., & Pfattheicher, S. (2020). When is your pain my gain? The use of perspective taking by everyday sadists. *Personality and Individual Differences*, *167*, 110213.

Lykken, D. T. (1995). *The Antisocial Personalities*. New Jersey: Psychology Press.

Lynam, D.R., Crowe, M.L., Vize, C., & Miller, J. (2020). Little Evidence That Honesty-Humility Lives Outside of FFM Agreeableness. A commentary to appear with:

Bizumic, B., & Monaghan, C. (2020). The HEXACO Model: Clinical Extensions and Universality. *European Journal of Personality*.34, 511-590.

Lynam, D, R., Miller, J., Vize, C., & Crowe, M. L. (2020). Agreeableness in the HEXACO. A commentary to appear with: Ashton, M. C., & Lee, K. (in press). Objections to the HEXACO model of personality structure–and why those objections fail. *European Journal of Personality*. https://doi.org/10.1002/per.2242.

Lynam, D. R., & Widiger, T. A. (2001). Using the five-factor model to represent the DSM-IV personality disorders: An expert consensus approach. *Journal of Abnormal Psychology*, *110*(3), 401.

*Lynam, D. R., Whiteside, S., & Jones, S. (1999). Self-reported psychopathy: A validation study. *Journal of Personality Assessment*, *73*(1), 110-132.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84.

*Mahmut, M. K., Menictas, C., Stevenson, R. J., & Homewood, J. (2011). Validating the factor structure of the Self-Report Psychopathy Scale in a community sample. *Psychological Assessment, 23*(3), 670.

Mair, P., & Wilcox, R. (2019). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, 1-25.

Malesza, M., Ostaszewski, P., Büchner, S., & Kaczmarek, M. C. (2019). The adaptation of the Short Dark Triad personality measure–psychometric properties of a German sample. *Current Psychology*, *38*(3), 855-864.

* Maneiro, L., López-Romero, L., Gómez-Fraguela, J. A., Cutrín, O., & Romero, E. (2019). Pursuing the Dark Triad: Psychometric properties of the Spanish version of the Dirty Dozen. *Journal of Individual Differences*, *40*(1), 36.

*Marion, B. E., & Sellbom, M. (2011). An examination of gender-moderated test bias on the Levenson Self-Report Psychopathy Scale. *Journal of Personality Assessment, 93*(3), 235-243.

Marsh, H. W., & Hau, K. T. (1999). Confirmatory Factor Analysis: Strategies for Small Sample Sizes. *Statistical Strategies for Small Sample Research, 1,* 251-284

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and

dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural equation Modeling*, *11*(3), 320-341.

Marsh, H. W., Nagengast, B., & Morin, A. J. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, *49*(6), 1194.

Marsh, H. W., Wen, Z., & Hau, K. T. (2004). Structural equation models of latent interactions: evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, *9*(3), 275.

Maples, J. L., Lamkin, J., & Miller, J. D. (2014). A test of two brief measures of the dark triad: The dirty dozen and short dark triad. *Psychological assessment*, *26*(1), 326.

Marion, B. E., & Sellbom, M. (2011). An examination of gender-moderated test bias on the Levenson Self-Report Psychopathy Scale. *Journal of Personality Assessment*, *93*(3), 235-243.

Massar, K., Winters, C. L., Lenz, S., & Jonason, P. K. (2017). Green-eyed snakes: The associations between psychopathy, jealousy, and jealousy induction. *Personality and Individual Differences, 115*, 164-168.

Maxwell, K., Donnellan, M. B., Hopwood, C. J., & Ackerman, R. A. (2011). The two faces of Narcissus? An empirical comparison of the Narcissistic Personality Inventory and the Pathological Narcissism Inventory. *Personality and Individual Differences*, *50*(5), 577-582.

Mayer, J., Savard, C., Brassard, A., Lussier, Y., & Sabourin, S. (2019). Subclinical Psychopathic Traits and Romantic Attachment in Treatment-Seeking Couples. *Journal of Marital and Family Therapy*, *46*(1), 165-178.

McCord, W., & McCord, J. (1964). *The psychopath: An essay of the criminal mind.* New York, NY: Van Nostrand Reinhold.

McCrae, R. R. (2015). A more nuanced view of reliability: Specificity in the trait hierarchy. *Personality and Social Psychology Review*, *19*, 97–112. https://doi.org/10. 1177/1088868314541857

McCullagh, P. (1985). On the asymptotic distribution of Pearson's statistic in linear exponential-family models. *International Statistical Review/Revue Internationale de Statistique*, 61-67.

McDonald, R. P. (2014). *Factor analysis and related methods*. Psychology Press.

*McHoskey, J. W., Worzel, W., & Szyarto, C. (1998). Machiavellianism and psychopathy. *Journal of Personality and Social Psychology*, *74*(1), 192.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, *93*(3), 568.

Menold, N. (2020). How Do Reverse-keyed Items in Inventories Affect Measurement Quality and Information Processing? *Field Methods*, *32*(2), 140-158.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525–543.

Miller, J. D., Gentile, B., Carter, N. T., Crowe, M., Hoffman, B. J., & Campbell, W. K. (2018). A comparison of the nomological networks associated with forced-choice and likert formats of the narcissistic personality inventory. *Journal of Personality Assessment*, *100*(3), 259-267.

*Miller, J. D., Hyatt, C. S., Maples-Keller, J. L., Carter, N. T., & Lynam, D. R. (2017). Psychopathy and Machiavellianism: A distinction without a difference?. *Journal of Personality*, *85*(4), 439-453.

Miller, J. D., Lynam, D. R., Hyatt, C. S., & Campbell, W. K. (2017). Controversies in narcissism. *Annual Review of Clinical Psychology*, *13*, 291-315

Miller, J. D., Maples, J., & Campbell, W. K. (2011). Comparing the construct validity of scales derived from the Narcissistic Personality Inventory: A reply to. *Journal of Research in Personality*, *45*(5), 401-407.

Miller, B. K., Nicols, K., & Konopaske, R. (2019). Measurement invariance tests of revisions to archaically worded items in the Mach IV scale. *PloS one*, *14*(10), e0223504.

Miller, J. D., Price, J., & Campbell, W. K. (2012). Is the Narcissistic Personality Inventory still relevant? A test of independent grandiosity and entitlement scales in the assessment of narcissism. *Assessment, 19*(1), 8-13.

Miller, J. D., Sleep, C. E., Crowe, M. L., & Lynam, D. R. (2020). Psychopathic boldness: Narcissism, self-esteem, or something in between? *Personality and Individual Differences*, *155*, 109761.

*Min, H., Pavisic, I., Howald, N., Highhouse, S., & Zickar, M. J. (2019). A systematic comparison of three sadism measures and their ability to explain workplace mistreatment over and above the dark triad. *Journal of Research in Personality*, *82*, 103862.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine, 6*(7), e1000097.

*Monaghan, C., Bizumic, B., & Sellbom, M. (2016). The role of Machiavellian views and tactics in psychopathology. *Personality and Individual Differences*, *94*, 72-81.

Morf, C. C., Weir, C., & Davidov, M. (2000). Narcissism and intrinsic motivation: The role of goal congruence. *Journal of Experimental Social Psychology, 36*(4), 424-438.

Moses, R., & Moses-Hrushovski, R. (1990). Reflections on the sense of entitlement. *The Psychoanalytic Study of the Child*, *45*(1), 61-78.

Moss, J. (2005). Race effects on the employee assessing political leadership: a review of Christie and Geis'(1970) Mach IV measure of Machiavellianism. *Journal of Leadership & Organizational Studies, 11*(2), 26-33.

Mudrack, P. E. (1989). Age-related differences in Machiavellianism in an adult sample. *Psychological Reports*, *64*(3_suppl), 1047-1050.

Mullins-Sweatt, S. N., Glover, N. G., Derefinko, K. J., Miller, J. D., & Widiger, T. A. (2010). The search for the successful psychopath. *Journal of Research in Personality*, *44*(4), 554-558.

Muraki, E., & Muraki, M. (2016). Generalized partial credit model. In *Handbook of Item Response Theory, Volume One* (pp. 155-166). Chapman and Hall/CRC.

Myers, W. C., Burket, R. C., & Husted, D. S. (2006). Sadistic personality disorder and comorbid mental illness in adolescent psychiatric inpatients. *Journal of the American Academy of Psychiatry and the Law Online*, *34*(1), 61-71.

Nariman, H. S., Hadarics, M., Kende, A., Lášticová, B., Poslon, X. D., Popper, M., ... & Minescu, A. (2020). Anti-Roma bias (stereotypes, prejudice, behavioral tendencies): A network approach toward attitude strength. *Frontiers in psychology*, *11*.

*Neal, T. M., & Sellbom, M. (2012). Examining the factor structure of the Hare self-report psychopathy scale. *Journal of Personality Assessment, 94*(3), 244-253.

Nell, V. (2006). Cruelty's rewards: The gratifications of perpetrators and spectators. *Behavioral and Brain Sciences*, *29*(3), 211-224.

Neumann, C. S., Kosson, D. S., & Salekin, R. T. (2017). Exploratory and confirmatory factor analysis of the psychopathy construct: Methodological and conceptual issues. In *The Psychopath* (pp. 79-104). Routledge.

*Neumann, C. S., Schmitt, D. S., Carter, R., Embley, I., & Hare, R. D. (2012). Psychopathic traits in females and males across the globe. *Behavioral Sciences & The Law, 30*(5), 557-574.

Neumann, C. S., Vitacco, M. J., Hare, R. D., & Wupperman, P. (2005). Reconstruing the “reconstruction” of psychopathy: A comment on Cooke, Michie, Hart, and Clark. *Journal of Personality Disorders*, *19*(6), 624-640.

Nowlis, S. M., Kahn, B. E., & Dhar, R. (2002). Coping with ambivalence: The effect of removing a neutral option on consumer attitude and preference judgments. *Journal of Consumer Research, 29*, 319-334. doi: 10.1086/344431

O'Boyle, E. H., Forsyth, D. R., Banks, G. C., Story, P. A., & White, C. D. (2015). A meta-analytic test of redundancy and relative importance of the dark triad and five-factor model of personality. *Journal of Personality, 83*(6), 644-664.

O'Connor, E. M., & Simms, C. M. (1990). Self-revelation as manipulation: The effects of sex and Machiavellianism on self-disclosure. *Social Behavior and Personality: An International Journal, 18(*1), 95-99.

Ogunsakin, I. B., & Shogbesan, Y. O. (2018). Item Response Theory (IRT): A modern statistical theory for solving measurement problem in 21st century. *International Journal of Scientific Research in Education*, *11*, 627-635.

Olderbak, S., & Figueredo, A. J. (2009). Predicting romantic relationship satisfaction from life history strategy. *Personality and Individual Differences*, *46*(5-6), 604-610.

O'Meara, A., Davies, J., & Hammond, S. (2011). The psychometric properties and utility of the Short Sadistic Impulse Scale (SSIS). *Psychological Assessment*, *23*(2), 523.

Orri, M., Rouquette, A., Pingault, J. B., Barry, C., Herba, C., Côté, S. M., & Berthoz, S. (2018). Longitudinal and sex measurement invariance of the affective neuroscience personality scales. *Assessment*, *25*(5), 653-666.

Özsoy, E., Rauthmann, J. F., Jonason, P. K., & Ardıç, K. (2017). Reliability and validity of the Turkish versions of dark triad dirty dozen (DTDD-T), short dark

triad (SD3-T), and single item narcissism scale (SINS-T). *Personality and Individual Differences*, *117*, 11-14.

Palan, S., & Schitter, C. (2018). Prolific.ac – A subject pool for online experiments *Journal of Behavioural and Experimental Finance 17*(3), 22-27.

Palmen, D., Derksen, J., & Kolthoff, E. (2018). House of cards: psychopathy in politics. *Public Integrity*, *20*(5), 427-443.

Palmen, D. G., Derksen, J. J., & Kolthoff, E. (2020). High self-control may support 'success' in psychopathic leadership: Self-control versus impulsivity in psychopathic leadership. *Aggression and Violent Behavior*, *50*, 101338.

Pan, W., Zhang, Q., Teo, T. S., & Lim, V. K. (2018). The dark triad and knowledge hiding. *International Journal of Information Management, 42*, 36-48.

Panitz, E. (1989). Psychometric investigation of the Mach IV scale measuring Machiavellianism. *Psychological Reports, 64*(3), 963-968.

Papageorgiou, K. A., Benini, E., Bilello, D., Gianniou, F. M., Clough, P. J., & Costantini, G. (2019). Bridging the gap: A network approach to Dark Triad, Mental Toughness, the Big Five, and perceived stress. *Journal of Personality*, *87*(6), 1250-1263.

Paulhus, D.L. (2014). Toward a taxonomy of dark personalities. *Current Directions in Psychological Science, 23*, 421–426. https://doi.org/10.1177/0963721414547737

Paulhus, D. L., Buckels, E. E., Trapnell, P. D., & Jones, D. N. (2020). Screening for Dark Personalities: The Short Dark Tetrad (SD4). *European Journal of Psychological Assessment*. Advance Online Publication.

Paulhus, D.L., & Dutton, D.G. (2015). Everyday Sadism. In V. Zeigler-Hill & D. K. Marcus (Eds.), *The dark side of personality* (pp.109-120). Washington, D.C.: American Psychological Association Press.

Paulhus, D.L., & Jones, D.N., (2015). Measuring dark personalities via questionnaire. In G. J. Boyle, D. H. Saklofske & G. Matthews (Eds.), *Measures of Personality and Social Psychological Constructs* (pp.562-594*)*. San Diego, CA: Academic Press.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*(6), 556-563.

Peer, E., Samat, S., Brandimarte, L., Acquisti, A., 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. J. Exp. Soc. Psychol. 70, 153– 163.

Peng. R.D. (2020). *R Programming for Data Science. https://bookdown.org/rdpeng/rprogdatascience/*

Pérez, M. E., & Pericchi, L. R. (2014). Changing statistical significance with the amount of information: The adaptive α significance level. *Statistics & probability letters*, *85*, 20-24.

Persson, B. N. (2019). Searching for Machiavelli but finding psychopathy and narcissism. *Personality Disorders: Theory, Research, and Treatment, 10*(3), 235–245.

*Persson, B. N., Kajonius, P. J., & Garcia, D. (2017a). Testing construct independence in the short dark triad using item response theory. *Personality and Individual Differences*, *117*, 74-80.

*Persson, B. N., Kajonius, P. J., & Garcia, D. (2017b). Revisiting the structure of the Short Dark Triad. *Assessment*, 1073191117701192.

Persson, B. N., & Lilienfeld, S. O. (2019). Social status as one key indicator of successful psychopathy: An initial empirical investigation. *Personality and Individual Differences*, *141*, 209-217.

Peterson, S. J., Walumbwa, F. O., Byron, K., & Myrowitz, J. (2009). CEO positive psychological traits, transformational leadership, and firm performance in high-technology start-up and established firms. *Journal of Management, 35(*2), 348-368.

Pfattheicher, S., Keller, J., & Knezevic, G. (2019). Destroying things for pleasure: On the relation of sadism and vandalism. *Personality and Individual Differences*, *140*, 52-56.

Pfattheicher, S., & Schindler, S. (2015). Understanding the dark side of costly punishment: The impact of individual differences in everyday sadism and existential threat. *European Journal of Personality*, *29*(4), 498-505.

Pincus, A. L., Ansell, E. B., Pimentel, C. A., Cain, N. M., Wright, A. G., & Levy, K. N. (2009). Initial construction and validation of the Pathological Narcissism Inventory. *Psychological Assessment*, *21*(3), 365.

Pincus, A. L., & Lukowitsky, M. R. (2010). Pathological narcissism and narcissistic personality disorder. *Annual Review of Clinical Psychology*, *6*, 421-446.

*Plouffe, R.A., Kowalski, C.M., Tremblay, P.F., Saklofske, D., Rogoza, R., Di Pierro, R., Chahine, S. (2021). Examination of the Assessment of Sadistic Personality Using Item Response Theory and Differential Item Functioning. *European Journal of Psychological Assessment, https://doi.org/10.1027/1015-5759/a000634*

Plouffe, R. A., Saklofske, D. H., & Smith, M. M. (2017). The assessment of sadistic personality: Preliminary psychometric evidence for a new measure. *Personality and Individual Differences, 104*, 166-171.

*Plouffe, R. A., Smith, M. M., & Saklofske, D. H. (2018). A psychometric investigation of the Assessment of Sadistic Personality. *Personality and Individual Differences*. DOI: 10.1016/j.paid.2018.01.002

Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*, *10*, 191–218. https://doi.org/10.7155/jgaa.00185

Poythress, N. G., & Hall, J. R. (2011). Psychopathy and impulsivity reconsidered. *Aggression and Violent Behavior*, *16*(2), 120-134.

Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, *26*(4), 269-281.

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica, 104,* 1–15. http://dx.doi .org/10.1016/S0001-6918(99)00050-5

Prolific (2019). Quickly find research participants you can trust. *Prolific.* https://www.prolific.co/

Purton, T., Officer, C., Bullivant, B., Mitchison, D., Griffiths, S., Murray, S. B., & Mond, J. (2018). Body dissatisfaction, narcissism and self-esteem in young men and women: A moderated mediation analysis. *Personality and Individual Differences*, *131*, 99-104.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71-90.

Qualtrics (2019). Online Survey Software. https://www.qualtrics.com/uk/

Ramo, D. E., Rodriguez, T. M. S., Chavez, K., Sommer, M. J., & Prochaska, J. J. (2014). Facebook recruitment of young adult smokers for a cessation trial: Methods, metrics, and lessons learned. *Internet Interventions*, *1*(2), 58–64.

Ramsey, S. R., Thompson, K. L., McKenzie, M., & Rosenbaum, A. (2016). Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior*, *58*, 354-360.

Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports, 45,* 590.

Raskin, R., & Hall, C. S. (1981). The Narcissistic Personality Inventory: Alternative form reliability and further evidence of construct validity. *Journal of Personality Assessment*, *45*(2), 159-162.

*Raskin, R., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology, 54*(5), 890.

Rauthmann, J. F. (2012). Towards multifaceted Machiavellianism: Content, factorial, and construct validity of a German Machiavellianism Scale*. Personality and Individual Differences, 52(*3), 345-351.

*Rauthmann, J. F. (2013). Investigating the MACH–IV with item response theory and proposing the trimmed MACH. *Journal of Personality Assessment*, *95*(4), 388-397.

*Ray, J. J. (1983). Defective validity of the Machiavellianism scale. *Journal of Social Psychology, 119*(2), 291-292.

Ray, J. V., Hall, J., Rivera-Hudson, N., Poythress, N. G., Lilienfeld, S. O., & Morano, M. (2013). The relation between self-reported psychopathic traits and distorted response styles: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment*, *4*(1), 1.

Raykov, T. (2012). Scale Construction and Development Using Structural Equation Modeling. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 472-492). New York, NY: Guilford Press.

Reifman, A., Arnett, J. J., & Colwell, M. J. (2007). Emerging adulthood: Theory, assessment and application. *Journal of Youth Development, 2*(1), 37-48.

Revelle, W. (2010). The “New Psychometrics” = Item Response Theory. https://pdfs.semanticscholar.org/068c/fdbea4ee95c0a4f8c982c54bc2e4f89f9663.pdf?_ga=2.18482687.772994346.1579706751-1209984339.1578439396

Robins, L. N. (1966). *Deviant children grown up: A sociological and psychiatric study of sociopathic personality.* Baltimore: Williams and Wilkins.

*Rogoza, R., & Cieciuch, J. (2018). Dark Triad traits and their structure: An empirical approach. *Current Psychology*, 1-16.

*Rogoza, R., Żemojtel-Piotrowska, M., Jonason, P. K., Piotrowski, J., Campbell, K. W., Gebauer, J. E., ... & Włodarczyk, A. (2020). Structure of Dark Triad Dirty Dozen Across Eight World Regions. *Assessment*, 1073191120922611.

Rohmann, E., Neumann, E., Herner, M. J., & Bierhoff, H. (2015). Grandiose and vulnerable narcissism: Self-construal, attachment, and love in romantic relationships. *European Psychologist, 17*, 279–290.

Rosenberg, M. (1965). Rosenberg self-esteem scale (RSE). Acceptance and commitment therapy. *Measures Package, 61*(52), 18.

Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.

Rosenthal, S.A. (2006). Narcissism and Leadership. A Review and Research Agenda. Working Papers. Center for Public Leadership.47-57. http://hdl.handle.net/1721.1/55948

Rosenthal, S. A., & Hooley, J. M. (2010). Narcissism assessment in social–personality research: Does the association between narcissism and psychological health result from a confound with self-esteem? *Journal of Research in Personality, 44*(4), 453-465.

Rosseel, Y. (2020). The lavaan tutorial. *Department of Data Analysis: Ghent University*.

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, *48*(1), 165-179.

Rudnev, M., Lytkina, E., Davidov, E., Schmidt, P., & Zick, A. (2018). Testing measurement invariance for a second-order factor: a cross-national test of the Alienation Scale. *Methods, data, analyses: A journal for quantitative methods and survey methodology (mda)*, *12*(1), 47-76.

Ruscio, J., & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment*, *24*(2), 282.

Russell, C. J., & Bobko, P. (1992). Moderated regression analysis and Likert scales: Too coarse for comfort. *Journal of Applied Psychology, 77,* 336–342. http://dx.doi.org/10.1037/0021-9010.77.3.336

Rutherford, M. J., Cacciola, J. S., Alterman, A. I., & McKay, J. R. (1996). Reliability and validity of the Revised Psychopathy Checklist in women methadone patients. *Assessment*, *3*(2), 145-156.

*Salekin, R. T., Chen, D. R., Sellbom, M., Lester, W. S., & MacDougall, E. (2014). Examining the factor structure and convergent and discriminant validity of the Levenson Self-Report Psychopathy Scale: Is the two-factor model the best fitting model? *Personality Disorders: Theory, Research, and Treatment, 5*(3), 289.

Salekin, R. T., Rogers, R., Ustad, K. L., & Sewell, K. W. (1998). Psychopathy and recidivism among female inmates. *Law and Human Behavior*, *22*(1), 109-128.

Salkind, N. J. (2010). *Encyclopedia of research design* (Vols. 1-0). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412961288

Samuel, D. B., & Widiger, T. A. (2004). Clinicians' personality descriptions of prototypic personality disorders. *Journal of Personality Disorders*, *18*(3: Special issue), 286-308.

Sapnas, K. G., & Zeller, R. A. (2002). Minimizing sample size when using exploratory factor analysis for measurement. *Journal of Nursing Measurement*, *10*(2), 135-154.

Sass, D. A., & Schmitt, T. A. (2010). A comparative investigation of rotation criteria within exploratory factor analysis. *Multivariate Behavioral Research*, *45*(1), 73-103.

Shchebetenko, S., Kalugin, A. Y., Mishkevich, A. M., Soto, C. J., & John, O. P. (2020). Measurement Invariance and Sex and Age Differences of the Big Five Inventory–2: Evidence from the Russian Version. *Assessment*, *27*(3), 472-486.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23-74.

Schimmack, U. (2010). What multi-method data tell us about construct validity. European Journal of Personality: *Published for the European Association of Personality Psychology, 2*4(3), 241-257.

Schmidt, F. L., Le, H., & Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual-differences constructs. *Psychological Methods*, *8*(2), 206.

Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, *31*, 43–53. https://doi.org/10.1016/j.newideapsych. 2011.02.007

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, *99*(6), 323-338.

Schumpe, B. M., & Lafrenière, M. A. K. (2016). Malicious joy: Sadism moderates the relationship between schadenfreude and the severity of others' misfortune. *Personality and Individual Differences, 94*, 32-37.

Sellbom, M., & Drislane, L. E. (2020). The classification of psychopathy. *Aggression and Violent Behavior*, 101473.

Shen, L., Condit, C. M., & Wright, L. (2009). The psychometric property and validation of a fatalism scale. *Psychology and Health, 24*(5), 597-613.

Shenkin, S. D., Watson, R., Laidlaw, K., Starr, J. M., & Deary, I. J. (2014). The attitudes to ageing questionnaire: Mokken Scaling Analysis. *PloS one, 9*(6), e99100.

*Shou, Y., Sellbom, M., & Han, J. (2017). Evaluating the construct validity of the Levenson self-report psychopathy scale in China. *Assessment, 24*(8), 1008-1023.

Seemann, E. A., Buboltz, W. C., Thomas, A., Soper, B., & Wilkinson, L. (2005). Normal Personality Variables and Their Relationship to Psychological Reactance. *Individual Differences Research, 3*(2).

Shaffer, J. A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation: A guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods*, *19*(1), 80-110.

Sherman, E. D., Miller, J. D., Few, L. R., Campbell, W. K., Widiger, T. A., Crego, C., & Lynam, D. R. (2015). Development of a Short Form of the Five-Factor Narcissism Inventory: The FFNI-SF. *Psychological Assessment*, *27*(3), 1110.

*Siddiqi, N., Shahnawaz, M. G., & Nasir, S. (2020). Reexamining construct validity of the Short Dark Triad (SD3) scale. *Current Issues in Personality Psychology*, *8*(1), 18-30.

Sijtsma, K., & Hemker, B. T. (1998). Nonparametric polytomous IRT models for invariant item ordering, with results for parametric models. *Psychometrika*, *63*(2), 183-200.

Sijtsma, K., Meijer, R. R., & van der Ark, L. A. (2011). Mokken scale analysis as time goes by: An update for scaling practitioners. *Personality and Individual Differences*, *50*(1), 31-37.

Sijtsma, K., & Molenaar, I. W. (2002). *Introduction to nonparametric item response theory* (Vol. 5). Sage.

Sijtsma, K., & van der Ark, L. A. (2017). A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. *British Journal of Mathematical and Statistical Psychology*, *70*(1), 137-158.

Simms, L. J., Zelazny, K., Williams, T. F., & Bernstein, L. (2019). Does the number of response options matter? Psychometric perspectives using personality questionnaire data. *Psychological Assessment*, *31*(4), 557.

Skeem, J. L., & Cooke, D. J. (2010). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment*, *22*(2), 433.

Sleep, C. E., Sellbom, M., Campbell, W. K., & Miller, J. D. (2017). Narcissism and response validity: Do individuals with narcissistic features underreport psychopathology?. *Psychological Assessment*, *29*(8), 1059.

Smith, R. J. (1985). The concept and measurement of social psychopathy. *Journal of Research in Personality, 19,* 219-231.

Smith, R. H., Powell, C. A., Combs, D. J., & Schurtz, D. R. (2009). Exploring the when and why of schadenfreude. *Social and Personality Psychology Compass*, *3*(4), 530-546.

*Somma, A., Fossati, A., Patrick, C., Maffei, C., & Borroni, S. (2014). The three-factor structure of the Levenson Self-Report Psychopathy Scale: Fool's gold or true gold? A study in a sample of Italian adult non-clinical participants. *Personality and Mental Health, 8*(4), 337-347.

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology, 113*(1), 117.

Soto, C. J., Kronauer, A., & Liang, J. K. (2016). *Five-factor model of personality*. In S. K. Whitbourne (Ed.), Encyclopedia of adulthood and aging (Vol. 2, pp. 506-510). Hoboken, NJ: Wiley

Spain, S. M., Harms, P., & LeBreton, J. M. (2014). The dark side of personality at work. *Journal of Organizational Behavior, 35*(S1), S41-S60.

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience & Biobehavioral Reviews*, *24*(4), 417-463.

Steenkamp, J. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross- national consumer research. *Journal of Consumer Research, 25,* 78–90.

Stevens, G. W., Deuling, J. K., & Armenakis, A. A. (2012). Successful psychopaths: Are they unethical decision-makers and why? *Journal of Business Ethics*, *105*(2), 139-149.

Straat, J. H., van der Ark, L. A., & Sijtsma, K. (2014). Minimum sample size requirements for Mokken scale analysis. *Educational and Psychological Measurement*, *74*(5), 809-822.

Strachan, K., Williamson, S., & Hare, R. D. (1990). Psychopathy and female offenders. *Unpublished study. Department of Psychology, University of British Columbia, Canada*.

Strand, S., & Belfrage, H. (2005). Gender differences in psychopathy in a Swedish offender sample. *Behavioral Sciences & the Law*, *23*(6), 837-850.

Stewart, J., Batty, A. O., & Bovee, N. (2012). Comparing multidimensional and continuum models of vocabulary acquisition: An empirical examination of the vocabulary knowledge scale. *TESOL Quarterly*, *46*(4), 695-721.

Strelan, P. (2007). Who forgives others, themselves, and situations? The roles of narcissism, guilt, self-esteem, and agreeableness. *Personality and Individual Differences*, *42*(2), 259-269.

Sturek, J. C., Loper, A. B., & Warren, J. I. (2008). Psychopathy in female inmates: The SCID-II Personality Questionnaire and the PCL-R. *Psychological Services, 5*(4), 309.

Suhr, D. D. (2006) Principal Component Analysis vs. Exploratory Factor Analysis. Retrieved from https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/203-30.pdf

Sutin, A. R., Luchetti, M., Aschwanden, D., Lee, J. H., Sesker, A. A., Strickhouser, J. E., Stephan, Y., & Terracciano, A. (2020). Change in five-factor model personality traits during the acute phase of the coronavirus pandemic. *PloS one*, *15*(8), e0237056.

Szabó, E., & Jones, D. N. (2019). Gender differences moderate Machiavellianism and impulsivity: Implications for Dark Triad research. *Personality and Individual Differences, 141*, 160-165.

Tarka, P. (2011). Measurement, Reliability and Scales Construction in a View of Classical True-Score Theory. *Argumenta Oeconomica*, *2*(27) 65-99.

Tavris, C. (1993). The Mismeasure of Woman. *Feminism & Psychology, 3*(2), 149-168.

Taylor, K. (2009). *Cruelty: Human evil and the human brain.* Oxford University Press.

The British Psychological Society, (2018). *Code Of Ethics And Conduct*. [online] Leicester: The British Psychological Society. Available at: https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct

Thomas, K. M., Wright, A. G., Lukowitsky, M. R., Donnellan, M. B., & Hopwood, C. J. (2012). Evidence for the criterion validity and clinical utility of the Pathological Narcissism Inventory. *Assessment*, *19*(2), 135-145.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tiefer, L. (1987). Social Construction and the Study of Human Sexuality, In P. Shaver and C. Hendrick (Eds), *Sex and Gender* (pp. 70-94). Newbury Park. CA: USA. Sage.

*Trahair, C., Baran, L., Flakus, M., Kowalski, C. M., & Rogoza, R. (2020). The structure of the Dark Triad traits: A network analysis. *Personality and Individual Differences*, *167*, 110265.

*Tran, U. S., Bertl, B., Kossmeier, M., Pietschnig, J., Stieger, S., & Voracek, M. (2018). “I'll teach you differences”: Taxometric analysis of the Dark Triad, trait

sadism, and the Dark Core of personality. *Personality and Individual Differences, 126*, 19-24.

Trémolière, B., & Djeriouat, H. (2016). The sadistic trait predicts minimization of intention and causal responsibility in moral judgment. *Cognition*, *146*, 158-171.

*Truhan, T. E., Wilson, P., Mõttus, R., & Papageorgiou, K. A. (2020). The many faces of dark personalities: An examination of the Dark Triad structure using psychometric network analysis. *Personality and Individual Differences*, 110502.

Trzesniewski, K. H., & Donnellan, M. B. (2010). Rethinking “Generation Me”: A study of cohort effects from 1976-2006. *Perspectives on Psychological Science, 5*(1), 58-75.

*Tschanz, B. T., Morf, C. C., & Turner, C. W. (1998). Gender differences in the structure of narcissism: A multi-sample analysis of the Narcissistic Personality Inventory. *Sex Roles*, *38*(9-10), 863-870.

*Uzieblo, K., Verschuere, B., Van den Bussche, E., & Crombez, G. (2010). The validity of the Psychopathic Personality Inventory-Revised in a community sample. *Assessment, 17*(3), 334-346.

Vaidya, J. G., Gray, E. K., Haig, J. R., Mroczek, D. K., & Watson, D. (2008). Differential stability and individual growth trajectories of big five and affective traits during young adulthood. *Journal of Personality*, *76*(2), 267-304.

Vaidya, J. G., Gray, E. K., Haig, J., & Watson, D. (2002). On the temporal stability of personality: Evidence for differential stability and the role of life experiences. *Journal of Personality and Social Psychology*, *83*(6), 1469.

Van Borkulo, C. D., Boschloo, L., Kossakowski, J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2017). Comparing network structures on three aspects: A permutation test. *Manuscript submitted for publication*, *10*.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4-70.

Van der Ark, L. A. (2012). New developments in Mokken scale analysis in R. *Journal of Statistical Software*, *48*(5), 1-27.

Van der Linden, W. J. (Ed.). (2016). *Handbook of Item Response Theory, Three Volume Set*. CRC Press. FL, USA.

van Dijk, W. W., & Ouwerkerk, J. W. (Eds.). (2014). *Schadenfreude: Understanding pleasure at the misfortune of others*. Cambridge University Press.

Van Schuur, W. H. (2003). Mokken scale analysis: Between the Guttman scale and parametric item response theory. *Political Analysis, 11*(2), 139-163.

*Vaughan, R., Madigan, D. J., Carter, G. L., & Nicholls, A. R. (2019). The Dark Triad in male and female athletes and non-athletes: Group differences and psychometric properties of the Short Dark Triad (SD3). *Psychology of Sport and Exercise*, *43*, 64-72.

Vazire, S., Naumann, L. P., Rentfrow, P. J., & Gosling, S. D. (2008). Portrait of a narcissist: Manifestations of narcissism in physical appearance. *Journal of Research in Personality, 42*(6), 1439-1447.

Vedel, A., & Thomsen, D. K. (2017). The Dark Triad across academic majors. *Personality and Individual Differences*, *116*, 86-91.

Vernon, P. A., Martin, R. A., Schermer, J. A., & Mackie, A. (2008). A behavioral genetic investigation of humor styles and their correlations with the Big-5 personality dimensions. *Personality and Individual Differences, 44*, 1116–1125.

Visser, B. A., Pozzebon, J. A., Bogaert, A. F., & Ashton, M. C. (2010). Psychopathy, sexual behavior, and esteem: It's different for girls. *Personality and Individual Differences*, *48*(7), 833-838.

Vitale, J. E., Smith, S. S., Brinkley, C. A., & Newman, J. P. (2002). The reliability and validity of the Psychopathy Checklist–Revised in a sample of female offenders. *Criminal Justice and Behavior*, *29*(2), 202-231.

Vize, C. E., Collison, K. L., Miller, J. D., & Lynam, D. R. (2020). The "core" of the dark triad: A test of competing hypotheses. *Personality Disorders: Theory, Research, and Treatment*, *11*(2), 91.

*Vize, C. E., Lynam, D. R., Collison, K. L., & Miller, J. D. (2018). Differences Among Dark Triad Components: A Meta-analytic Investigation. *Personality Disorders: Theory, Research, and Treatment, 9*(2), 101-111.

Volmer, J., Koch, I. K., & Wolff, C. (2019). Illuminating the 'dark core': Mapping global versus specific sources of variance across multiple measures of the dark triad. *Personality and Individual Differences*, *145*, 97-102.

Walters, G. D., Brinkley, C. A., Magaletta, P. R., & Diamond, P. M. (2008). Taxometric analysis of the Levenson self-report psychopathy scale. *Journal of Personality Assessment*, *90*(5), 491-498.

Wang, J., & Wang, X. (2012). *Structural Equation Modeling: Applications Using Mplus*. Hoboken, NJ: Wiley, Higher Education Press.

Ward, J. H. (1963). Hierarchical clustering to optimise an objective function. *Journal of the American Statistical Association*, *58*, 238–244.

Warren, J. I., & South, S. C. (2006). Comparing the constructs of antisocial personality disorder and psychopathy in a sample of incarcerated women. *Behavioral Sciences & the Law*, *24*(1), 1-20.

Warren, J. I., South, S. C., Burnette, M. L., Rogers, A., Friend, R., Bale, R., & Van Patten, I. (2005). Understanding the risk factors for violence and criminality in women: The concurrent validity of the PCL-R and HCR-20. *International Journal of Law and Psychiatry*, *28*(3), 269-289.

Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality*, *38*(4), 319-350.

Watson, R., Deary, I. J., & Shipley, B. (2008). A hierarchy of distress: Mokken scaling of the GHQ-30. *Psychological Medicine, 38*(4), 575-579.

Watson, R., Wang, W., & Thompson, D. R. (2014). Violations of local stochastic independence exaggerate scalability in Mokken scaling analysis of the Chinese Mandarin SF-36. *Health and Quality of Life Outcomes, 12*(1), 149.

*Watts, A. L., Waldman, I. D., Smith, S. F., Poore, H. E., & Lilienfeld, S. O. (2017). The nature and correlates of the dark triad: The answers depend on the questions. *Journal of Abnormal Psychology, 126*(7), 951

Weaver, K. F., Morales, V. C., Dunn, S. L., Godde, K., & Weaver, P. F. (2017). *An introduction to statistical analysis in research: with applications in the biological and life sciences*. John Wiley & Sons.

*Webster, G. D., & Jonason, P. K. (2013). Putting the “IRT” in “Dirty”: Item Response Theory analyses of the Dark Triad Dirty Dozen—An efficient measure of narcissism, psychopathy, and Machiavellianism. *Personality and Individual Differences, 54*(2), 302-306.

Weiser, E. B. (2015). # Me: Narcissism and its facets as predictors of selfie-posting frequency. *Personality and Individual Differences*, *86*, 477-481.

Weisstein, N. (1968). Kinder, Kirche, Kuche as scientific law: Psychology constructs the female. Boston, MA: New England Press.

Weisstein, N. (1971). Psychology constructs the female; or, the fantasy life of the male psychologist (with some attention to the fantasies of his friends, the male biologist and the male anthropologist). Journal of Social Education, 35, 362–373. doi:10.1177/0959353593032005

Weizmann-Henelius, G., Viemerö, V., & Eronen, M. (2004). Psychopathy in violent female offenders in Finland. *Psychopathology*, *37*(5), 213-221.

Wetzel, E., Grijalva, E., Robins, R. W., & Roberts, B. W. (2019). You're still so vain: Changes in narcissism from young adulthood to middle age. *Journal of Personality and Social Psychology*.

Widiger, T. A., Trull, T. J., Clarkin, J. F., Sanderson, C., & Costa, P. T., Jr. (2002). A description of the DSM-IV personality disorders with the five-factor model of personality. In P. T. Costa, Jr. & T. A. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (p. 89–99). American Psychological Association.

Widom, C. S. (1977). A methodology for studying noninstitutionalized psychopaths. *Journal of Consulting and Clinical Psychology*, *45*(4), 674.

*Williams, M. L., Hazleton, V., & Renshaw, S. (1975). The measurement of Machiavellianism: A factor analytic and correlational study of Mach IV and Mach V. *Communications Monographs, 42*(2), 151-159.

*Williams, K. M., Nathanson, C., & Paulhus, D. L. (2003, August). Structure and validity of the self-report psychopathy scale-III in normal populations. *In 111th annual convention of the American Psychological Association.*

Williams, K. M., Paulhus, D. L., & Hare, R. D. (2007). Capturing the four-factor structure of psychopathy in college students via self-report. *Journal of Personality Assessment, 88*(2), 205-219.

Wilson, C., & Seaman, D. (2007). *The serial killers: A study in the psychology of violence*. Random House. London, UK.

*Wilson, M. S., & Sibley, C. G. (2011). 'Narcissism creep?': Evidence for Age-Related Differences in Narcissism in the New Zealand General Population. *New Zealand Journal of Psychology, 40*(3).

Wilson, D. S., Near, D., & Miller, R. R. (1996). Machiavellianism: A synthesis of the evolutionary and psychological literatures. *Psychological Bulletin*, *119*(2), 285.

Wilson, M. S., & Sibley, C. G. (2011). 'Narcissism creep?': Evidence for Age-Related Differences in Narcissism in the New Zealand General Population. *New Zealand Journal of Psychology*, *40*(3).

Wink, P. (1991). Two faces of narcissism. *Journal of Personality and Social Psychology*, *61*(4), 590.

Woods, C. M. (2002). Factor analysis of scales composed of binary items: illustration with the Maudsley Obsessional Compulsive Inventory. *Journal of Psychopathology and Behavioral Assessment*, *24*(4), 215-223.

Wothke, W. (1993). Nonpositive definite matrices in structural modelling. In K.A. Bollen & J.S Long (Eds.), *Testing structural equation models* (pp. 256-93). Newbury Park, CA: Sage.

Wright, A. G., & Edershile, E. A. (2018). Issues resolved and unresolved in pathological narcissism. *Current Opinion in Psychology*, *21*, 74-79.

Wright, A. G., Lukowitsky, M. R., Pincus, A. L., & Conroy, D. E. (2010). The higher order factor structure and gender invariance of the Pathological Narcissism Inventory. *Assessment*, *17*(4), 467-483.

Wu, A. D., Zhen, L., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research, and Evaluation*, *12*(1), 3.

Xu, H., & Tracey, T. J. (2017). Use of multi-group confirmatory factor analysis in examining measurement invariance in counseling psychology research. *The European Journal of Counselling Psychology, 6*(1), 75-82.

Yang, F. M. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry*, *26*(3), 171.

Zanon, C., Hutz, C. S., Yoo, H. H., & Hambleton, R. K. (2016). An application of item response theory to psychological test development. *Psicologia: Reflexão e Crítica*, *29*(1), 1-10.

Zeigler-Hill, V., Clark, C. B., & Pickard, J. D. (2008). Narcissistic subtypes and contingent self-esteem: Do all narcissists base their self-esteem on the same domains? *Journal of Personality*, *76*(4), 753-774.

Zitek, E. M., & Schlund, R. J. (2020). Psychological entitlement predicts noncompliance with the health guidelines of the COVID-19 pandemic. *Personality and Individual Differences*, 110491.

The Dark Side of Humanity Scale: A Reconstruction of the Dark Tetrad Constructs as they Manifest in the General Population

Appendices

Appendix 1

An Interrogation of the Structure, Sex and Age Invariance of the Dark Tetrad Scales: A Systematic Review

1.1 Study Quality Appraisal Form (CASP, 2017)

Criteria to assess the studies

Study Aim

Did the study address a clearly focused issue?

Sample

Was the sample recruited in an acceptable way?

Was the sample representative of a defined population?

Was everybody included who should have been?

Selection bias?

Measurement

Was the outcome accurately measured to minimise bias?

Measurement or classification bias:

Did they use subjective or objective measurements?

Do the measures truly reflect what you want them to? (Have they been validated)

Confounding Factors

Have the authors identified all important confounding factors?

Have they taken account of the confounding factors in the design and/or analysis?

Study Results

What are the results of this study?

How precise are the results?

Do you believe the results?

Can it be due to bias, chance or confounding?

Are the design and methods of this study sufficiently flawed to make the results unreliable?

Do the results of this study fit with other available evidence?

Study Quality Appraisal Form (CASP, 2017

Key = C, comprehensive; A, acceptable; I, incomplete.	Aim	Ethics	Sample	Data collection	Analysis	Bias	Reliability	Fit w/o evidence
Psychopathy								
Anestis, Caron & Carbonell (2011)	C	A	C	C	C	C	C	C
Benning, Patrick, Hicks, Blonigen & Krueger (2003)	C	C	A	C	A	C	C	C
Christian & Sellbom (2016)	C	A	C	C	A	C	C	A
Dotterer, Waller, Neumann, Shaw, Forbes, Hariri & Hyde (2017)	C	C	C	C	C	C	C	C
Douglas, Bore & Munro (2012)	C	C	C	C	C	C	C	C
Eichenbaum, Marcus & French (2018)	C	C	C	C	C	C	C	C
Garofalo, Noteborn, Sellbom & Bogaerts (2018)	C	C	C	C	C	C	C	C
Gordts, Uzieblo, Neumann, Van den Bussche & Rossi (2017)	C	C	C	C	C	A	C	C
Gummelt, Anestis & Carbonell (2012)	C	C	C	C	C	C	C	C
Hauck-Filho & Teixeira (2014)	C	C	C	C	C	C	C	C
Lester, Salekin & Sellbom (2013)	C	C	C	C	C	C	C	C
Lynam, Whiteside & Jones (1999)	C	A	C	C	A	C	C	A
Mahumut, Menictas, Stevenson & Homewood (2011)	C	A	C	C	C	C	C	C
Marion & Sellbom (2011)	C	C	C	C	C	C	C	C
Neal & Sellbom (2012)	C	C	C	C	C	C	C	C

Key = C, comprehensive; A, acceptable; I, incomplete.	Aim	Ethics	Sample	Data collection	Analysis	Bias	Reliability	Fit w/o evidence
Neumann, Schmitt, Carter, Embley & Hare (2012)	C	C	C	C	C	A	C	C
Salekin, Chen, Sellbom, Lester & MacDougall (2014)	C	C	C	C	C	C	C	C
Shou, Sellbom & Han (2017)	C	C	C	C	C	C	C	C
Somma, Fossati, Patrick, Maffei & Borroni (2014)	C	C	C	C	A	C	C	C
Uzieblo, Verschuere, Van den Bussche & Crombez (2010)	C	C	C	C	A	C	C	C
Williams, Nathanson & Paulhus (2003)	C	A	C	C	C	A	C	C
Narcissism								
Ackerman, Witt, Donnellan, Trzesniewski, Robins & Kashy (2011)	C	C	C	C	C	C	C	C
Ames, Rose & Anderson (2006)	C	C	C	C	A	C	C	A
Barelds & Dijkstra (2010)	C	A	C	C	C	C	C	C
Braun, Kempenaers, Linkowski & Loas (2016)	C	C	C	C	A	C	C	C
Briganti & Linkowski, 2020	C	I	A	C	C	C	C	C
Brown, Akers & Giacomino (2013)	C	I	C	C	A	C	A	A
Clarke, Karlov & Neale (2015)	C	C	C	C	A	C	C	C
Corry, Merritt, Mrug & Pamp (2008)	C	C	C	C	C	C	C	C
Emmons (1984)	C	A	C	C	A	C	C	C
Jackson, Ervin & Hodge (1992)	C	A	A	C	C	C	C	C
Kubarych, Deary & Austin (2004)	C	C	C	C	C	C	C	C

Key = C, comprehensive; A, acceptable; I, incomplete.	Aim	Ethics	Sample	Data Collection	Analysis	Bias	Reliability	Fit with other evidence
Raskin & Terry (1988)	C	C	C	C	C	C	C	C
Tschanz, Morf & Turner (1998)	C	C	C	C	C	C	C	C
Wilson & Sibley (2011)	C	C	C	C	C	C	C	C
Machiavellianism								
Ahmed & Stewart (1981)	C	I	C	C	C	C	C	A
Corral & Calvete (2000)	C	C	C	C	C	A	C	C
Czibor et al., (2017)	C	C	C	C	C	C	C	C
Hunter, Gerbing & Boster (1982)	C	C	C	C	C	C	C	C
Monaghan, Bizumic & Sellbom (2016)	C	C	C	C	C	C	C	C
Rauthmann (2013)	C	C	C	C	C	C	C	C
Ray (1983)	C	I	A	C	A	C	A	C
Williams, Hazelton & Renshaw (1975)	C	C	C	C	C	C	C	C
Psychopathy and Machiavellianism								
McHoskey, Worzel & Szyarto (1988)	C	C	C	C	C	C	C	C
Miller, Hyatt, Maples-Keller, Carter & Lynam (2017)	C	C	C	C	C	C	C	C
Vize, Lynam, Collison & Miller (2018).	C	C	C	C	C	C	C	C
Dark Triad								
Arseneault & Catano (2019)	C	C	C	C	C	C	C	C

Key = C, comprehensive; A, acceptable; I, incomplete.	Aim	Ethics	Sample	Data Collection	Analysis	Bias	Reliability	Fit w/o evidence
Atari & Chegeni (2016)	C	C	C	C	C	C	C	C
Carter, Campbell, Muncer & Carter (2015)	C	C	C	C	C	C	C	C
Chiorri, Garofalo & Velotti, (2017)	C	C	C	C	C	C	C	C
Dinić, Petrović & Jonason (2018)	C	C	C	C	C	C	C	C
Dinić, Wertag, Tomaševič & Soklovska (2020)	C	C	C	C	C	C	C	C
Gamache, Savard & Maheux-Caron (2018)	C	C	C	C	C	C	C	C
Garcia, MacDonald & Rapp-Ricciardi (2017)	C	C	C	C	C	C	C	C
Grigoras, Butucescu, Miulescu, Opariuc-Dan & Iliescu (2020)	C	I	C	C	C	C	C	C
Kajonius, Persson, Rosenberg & Garcia (2016)	C	C	C	C	C	C	C	C
Kawamoto, Shimotsukasa & Oshio (2020)	C	C	C	C	C	C	C	C
Klimstra, Jeronimus, Sitsema & Denissne, (2020)	C	I	C	C	C	C	C	C
Maneiro, López-Romero, Gómez-Fraguela, Cutrin & Romero (2018)	C	C	C	C	C	C	C	C
Persson, Kajonius & Garcia (2017a)	C	C	C	C	C	C	C	C
Persson, Kajonius & Garcia (2017b)	C	C	C	C	C	C		
Rogoza & Cieciuch (2018)	C	C	C	C	C	C	C	C
Rogoza, Zemoitel-Piotrowska, Jonason et al., (2020)	C	C	C	C	C	C	C	C
Siddiqi, Shahnawaz & Nasir (2020)	C	I	C	C	A	C	A	C
Trahair, Baran, Flakus, Kowalski & Rogoza (2020)	C	I	C	C	C	C	C	C

Key = C, comprehensive; A, acceptable; I, incomplete.	Aim	Ethics	Sample	Data Collection	Analysis	Bias	Reliability	Fit w/o evidence
Tran, Bertl, Kossmeier, Pietschnig, Stieger & Voracek (2018)	C	C	C	C	C	C	C	C
Truhan, Wilson, Möttus & Papageorgious (2020)	C	I	C	C	C	C	C	C
Vaughan, Madigan, Carter & Nicholls (2019)	C	C	C	C	C	C	C	C
Watts, Waldman, Smith, Poore & Lilienfeld (2017)	C	C	C	C	C	C	C	C
Webster & Jonason (2013)	C	C	C	C	C	C	C	C
Everyday Sadism								
Dinić, Allred, Petrović & Wertage, 2020	C	C	C	C	C	C	C	C
Min, Pavisic, Howald, Highhouse & Zickar (2019)	C	I	C	C	C	C	C	C
Plouffe, Kowalski, Tremblay, Saklofrske, Rogoza, Di Pierro & Chaine (2020)	C	I	C	C	C	C	C	C
Plouffe, Smith & Saklofske (2018)	C	C	C	C	C	C	C	C

Note: C= Comprehensive; A = Acceptable; I = Incomplete.

Appendix 2

R Code for the methods analysed with R in Chapters Four, Five and Six

Confirmatory Factor Analysis

```
Models$DD <- 'Machiavellian =~ DDM1+DDM2+DDM3+DDM4
Psychopathy =~ DDP1+DDP2+DDP3+DDP4
Narcissism =~ DDN1+DDN2+DDN3+DDN4'
```

Measurement Invariance

Configural Invariance

```
configuralASP <- cfa(ASPmodel, data=ASPMFOv26,estimator = "WLSMV", group="Sex")
summary (configuralNSMF1825, fit.measures=T)
```

Metric Invariance

```
metricASP <- cfa(ASPmodel, data=ASPMFOv26,estimator = "WLSMV", group="Sex",
group.equal="loadings")
summary(metricASP, fit.measures=T)
```

Scalar Invariance

```
scalarASP <- cfa(ASPmodel, data=ASPMFOv26,estimator = "WLSMV", group="Sex",
group.equal=c("loadings", "Intercepts"))
summary (scalarASP, fit.measures=T)
```

Item Response Theory

Mokken Scale Analysis

Library - Mokken

```
AISP(ASPF18D, search = "normal", lowerbound = 0.4, alpha = 0.05)
MS1 <- ASPF18D[, c("ASP1", "ASP2", "ASP3", "ASP4", "ASP5", "ASP6", "ASP7",
"ASP8", "ASP9")]
coefH(MS1)
check.monotonicity (MS1)
summary(check.monotonicity(MS1))
iio.results <- check.iio(MS1)
summary(check.iio(MS1, item.selection=TRUE))
check.reliability(MS1d, MS = TRUE, alpha = TRUE, lambda.2 = TRUE, LCRC = FALSE,
nclass = nclass.default)
apply(MS1, 2, mean)
```

Generalised Partial Credit Model

Library – mirt

```
M18F1Mirt <- Males18_25NS[,c("P7", "P9", "P10", "P12", "P13", "P14", "P15", "P17",
"P18", "P22", "P27", "M3", "M10", "M23", "M24", "M26", "M28", "M29")]
results.M18F1NS<-mirt(data=M18F1Mirt, model=1, itemtype="gpcm",
SE=TRUE,verbose=FALSE)
coef.gpcm <- coef(results.M18F1NS, IRTpars=TRUE, simplify=TRUE)
items.gpcm <- as.data.frame(coef.gpcm$items)
print(items.gpcm)
plot(M18F1NS, type="trace")
itemplot (M18F1NS, 1,  type="trace")
```

Exploratory Graph Analysis

Library - EGAnet

ega.MP <- EGA(data=MP1[,5:68], model="glasso", plot.EGA = TRUE)

summary (ega.MP)

glasso estimates the Gaussian graphical model using graphical LASSO with extended Bayesian information criterion to select the optimal regularisation parameter.

Item Redundancy

M18Mach <- (Males18_25NS[1:53])

redundancy.analysis(

M18Mach,

n = NULL,

method = c("wTO"),

type = c("adapt"),

key = NULL,

reduce = TRUE,

reduce.method = c("remove"),

lavaan.args = list(),

adhoc = TRUE,

plot.redundancy = FALSE,

plot.args = list())

Robust Correlations for Non-Normal Data with Outliers

To check for normality

Library MVN

S3M18Norm <- S3_Males_18_25[c("CM", "Ent", "SP", "ER",

"PrimaryPsych","SecondaryPsych", "NPITotal", "MachTotal", "ASP", "Extraversion",

"Agreeableness", "Conscientious", "NegEmo", "OpenMind", "SelfEsteem")]

> mvn(S3M18Norm, mvnTest = "mardia")

ConvF18 <- S3_Females_18_25[c("CM", "Ent", "SP", "ER", "PrimaryPsych",

"SecondaryPsych", "NPITotal", "MachTota", "ASP", "Extraversion", "Agreeableness",

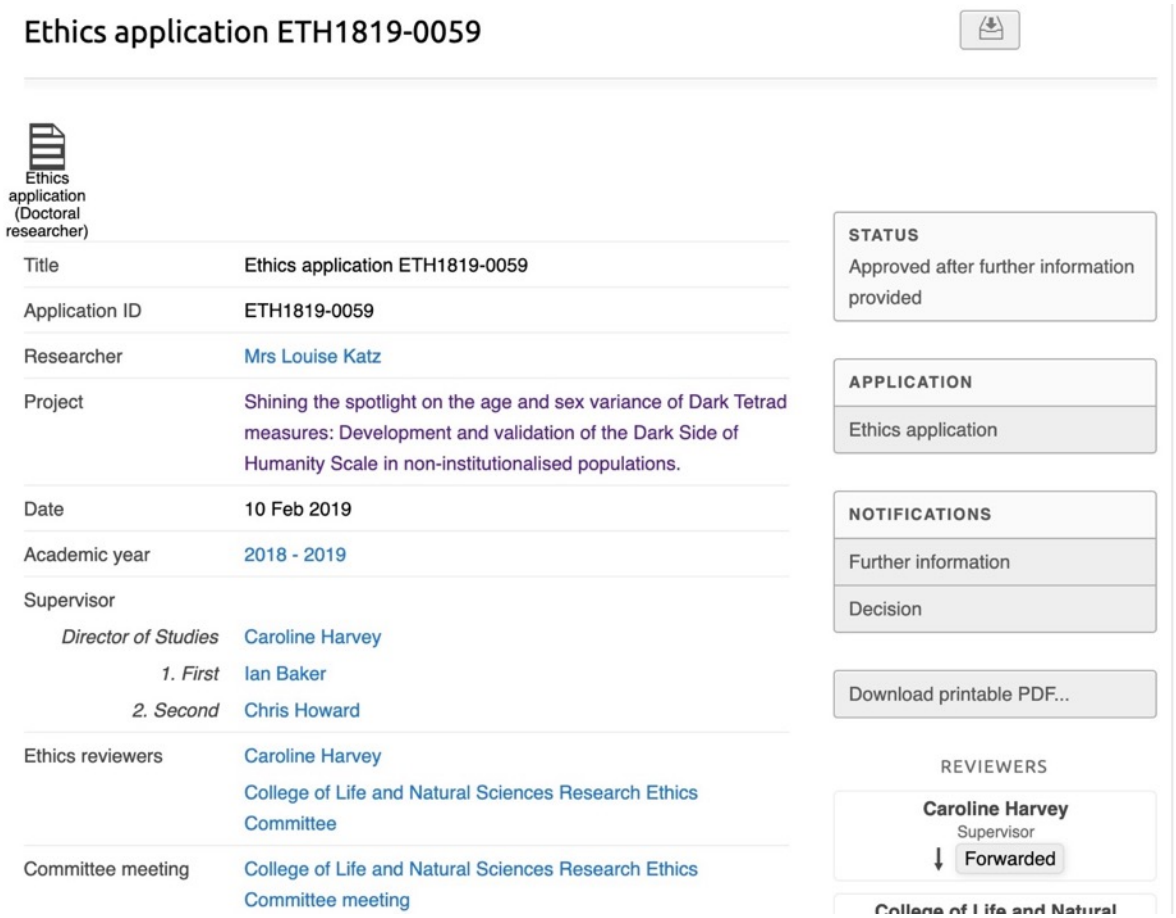
"Conscientious", "NegEmo", "OpenMind", "SelfEsteem")]

pball(ConvF18)

Appendix 3

The Dark Tetrad Measures: An Interrogation of the Factor Structures, Construct Validity and Invariance Across Sex and Age.

3.1 Ethical Approval Application



Ethics application ETH1819-0059

Ethics application (Doctoral researcher)

Title	Ethics application ETH1819-0059
Application ID	ETH1819-0059
Researcher	Mrs Louise Katz
Project	Shining the spotlight on the age and sex variance of Dark Tetrad measures: Development and validation of the Dark Side of Humanity Scale in non-institutionalised populations.
Date	10 Feb 2019
Academic year	2018 - 2019
Supervisor	
Director of Studies	Caroline Harvey
1. First	Ian Baker
2. Second	Chris Howard
Ethics reviewers	Caroline Harvey College of Life and Natural Sciences Research Ethics Committee
Committee meeting	College of Life and Natural Sciences Research Ethics Committee meeting

STATUS
Approved after further information provided

APPLICATION
Ethics application

NOTIFICATIONS
Further information
Decision

Download printable PDF...

REVIEWERS
Caroline Harvey
Supervisor
Forwarded

College of Life and Natural

3.1.1 Ethical Approval Confirmation

Kedleston Road, Derby
DE22 1GB, UK

T: +44 (0)1332 591060
E: researchoffice@derby.ac.uk
Sponsor License No: QGN14R294

Dear Louise,

Ethics Ref No: ETH1819-0059

Thank you for submitting this revised application to the Human Sciences Research Ethics Committee.

I have now reviewed the revised documents you sent following the feedback you received on your initial application, and I am satisfied that all of the issues raised with the application have been dealt with. The application can now therefore be approved.

The following documents have now been re-reviewed:

Ethics application formResponses to reviewer comments

If any changes to the study described in the application or supporting documentation is necessary, you must notify the committee and may be required to make a resubmission of the application.

Please note ethical approval for the study in application ETH1819-0059

is valid for a period of 5 years i.e. May 2024.

Good luck with the study.

Yours sincerely

Dr Ed Stupple

3.2 Measures Used in Chapter Four

3.2.1 Machiavellianism: The Mach IV (Christie & Geis, 1970)

Instructions

Please indicate to what extent you agree or disagree with each statement.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Disagree Somewhat	Neither agree nor disagree	Agree Somewhat	Agree	Strongly Agree

1. The best way to handle people is to tell them what they want to hear.
2. When you ask someone to do something for you, it is best to give the real reasons for wanting it rather than giving reasons which might carry more weight.
3. Anyone who completely trusts anyone else is asking for trouble.
4. It is hard to get ahead without cutting corners here and there.
5. Honesty is the best policy in all cases.
6. It is safest to assume that all people have a vicious streak and it will come out when they are given a chance.
7. Never tell anyone the real reason you did something unless it is useful to do so.
8. One should take action only when sure it is morally right.
9. It is wise to flatter important people.
10. All in all, it is better to be humble and honest than important and dishonest.
11. Barnum was very wrong when he said there's a sucker born every minute.
12. People suffering from incurable diseases should have the choice of being put painlessly to death.
13. It is possible to be good in all respects.
14. Most people are basically good and kind.
15. There is no excuse for lying to someone else.
16. Most men forget more easily the death of their father than the loss of their property.
17. Most people who get ahead in the world lead clean, moral lives.
18. Generally speaking, men won't work hard unless they're forced to do so.
19. The biggest difference between most criminals and other people is that criminals are stupid enough to get caught.
20. Most men are brave.

3.2.2 Psychopathy: The Levenson Self-Report Psychopathy Scale (LSRP; Levenson, Kiehl & Fitzpatrick, 1995)

Instructions.

Please indicate to what degree you agree or disagree with each statement.

1	2	3	4
Disagree Strongly	Disagree Somewhat	Agree Somewhat	Agree Strongly

Primary Psychopathy

1. Success is based on survival of the fittest; I am not concerned about the losers.
2. For me what's right is whatever I can get away with
3. In today's world, I feel justified in doing anything I can get away with to succeed.
4. My main purpose in life is getting as many goodies as I can.
5. Making a lot of money is my most important goal.
6. I let others worry about higher values, my main concern is with the bottom line.
7. People who are stupid enough to get ripped off usually deserve it.
8. Looking out for myself is my top priority.
9. I tell other people what they want to hear so that they will do what I want them to do.
10. I would be upset if my success came at someone else's expense.
11. I often admire a really clever scam.
12. I make a point of trying not to hurt others in pursuit of my goals.
13. I enjoy manipulating other people's feelings.
14. I feel bad if my words or actions cause someone else to feel emotional pain.
15. Even if I were trying very hard to sell something, I wouldn't like about it.
16. Cheating is not justified because it is unfair to others.

Secondary Psychopathy

1. I find myself in the same kinds of trouble, time after time.
2. I am often bored.
3. I find that I am able to pursue one goal for a long time.
4. I don't plan anything very far in advance.
5. I quickly lose interest in tasks I start.
6. Most of my problems are due to the fact that other people just don't understand me.
7. Before I do anything, I carefully consider the possible consequences.
8. I have been in a lot of shouting matches with other people.
9. When I get frustrated, I often 'let off steam' by blowing my top.
10. Love is overrated.

3.2.3. Narcissism: Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979)

Instructions

Below are paired statements, for each, please choose one which you associate with the most.

1. A. I have a natural talent for influencing people.
   B. I am not good at influencing people.
2. A. Modesty doesn't become me.
   B. I am essentially a modest person.
3. A. I would do almost anything on a dare.
   B. I tend to be a fairly cautious person.
4. A. When people compliment me I sometimes get embarrassed.
   B. I know that I am good because everybody keeps telling me so.
5. A. The thought of ruling the world frightens the hell out of me.
   B. If I ruled the world it would be a better place.
6. A. I can usually talk my way out of anything.
   B. I try to accept the consequences of my behaviours.
7. A. I prefer to blend in with the crowd.
   B. I like to be the centre of attention.
8. A. I will be a success.
   B. I am not too concerned about success.
9. A. I am no better or worse than most people.
   B. I think I am a special person.
10. A. I am not sure if I would make a good leader.
    B. I see myself as a good leader.
11. A. I am assertive.
    B. I wish I were more assertive.
12. A. I like to have authority over other people.
    B. I don't mind following orders.
13. A. I find it easy to manipulate people.
    B. I don't like it when I find myself manipulating people.
14. A. I insist upon getting the respect that is due me.
    B. I usually get the respect that I deserve.
15. A. I don't particularly like to show off my body.
    B. I like to show off my body.

16. A. I can read people like a book
    B. People are sometimes hard to understand.
17. A. If I feel competent I am willing to take responsibility for making decisions.
    B. I like to take responsibility for making decisions.
18. A. I just want to be reasonably happy.
    B. I want to be something in the eyes of the world.
19. A. My body is nothing special.
    B. I like to look at my body.
20. A. I try not to be a show-off.
    B. I will usually show off if I get the chance.
21. A. I always know what I am doing.
    B. Sometimes I am not sure of what I'm doing.
22. A. I sometimes depend on people to get things done.
    B. I rarely depend on anyone else to get things done.
23. A. Sometimes I tell good stories.
    B. Everybody likes to hear my stories.
24. A. I expect a great deal from other people.
    B. I like to do things for other people.
25. A. I will never be satisfied until I get all that I deserve.
    B. I take my satisfactions as they come.
26. A. Compliments embarrass me.
    B. I like to be complimented.
27. A. I have a strong will to power.
    B. Power for its own sake doesn't interest me.
28. A. I don't care about new fads and fashions
    B. I like to start new fads and fashions.
29. A. I like to look at myself in the mirror.
    B. I am not particularly interested in looking at myself in the mirror.
30. A. I really like to be the center of attention.
    B. It makes me uncomfortable to be the center of attention.
31. A. I can live my life in any way I want to.
    B. People can't always live their lives in terms of what they want.
32. A. Being an authority doesn't mean that much to me.
    B. People always seem to recognise my authority.
33. A. I would prefer to be a leader.
    B. It makes little difference to me whether I am a leader or not.

34. A. I am going to be a great person.
    B. I hope I am going to be successful.
35. A. People sometimes believe what I tell them.
    B. I can make anybody believe anything I want them to.
36. A. I am a born leader.
    B. Leadership is a quality that takes a long time to develop.
37. A. I wish somebody would someday write my biography.
    B. I don't like people to pry into my life for any reason.
38. A. I get upset when people don't notice how I look when I go out in public.
    B. I don't mind blending into the crowd when I go out in public.
39. A. I am more capable than other people.
    B. There is a lot that I can learn from other people.
40. A. I am much like everybody else.
    B. I am an extraordinary person.

3.2.4 Dark Triad: The Dirty Dozen (Jonason & Webster, 2010)

Instructions

Please indicate how much you agree or disagree with each statement.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

Machiavellian

1. I tend to manipulate others to get my way.
2. I have used deceit or lied to get my way.
3. I have used flattery to get my way.
4. I tend to exploit others towards my own end.

Psychopathy

5. I tend to lack remorse.
6. I tend to be unconcerned with the morality of my actions.
7. I tend to be callous or insensitive.
8. I tend to be cynical.

Narcissism

9. I tend to want others to admire me.
10. I tend to want others to pay attention to me.
11. I tend to seek prestige or status.
12. I tend to expect special favours from others.

3.2.5 Dark Triad: The Short Dark Triad (Jones & Paulhus, 2014)

Please indicate to what extent you agree or disagree with each statement.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

Machiavellianism

1. It's not wise to tell your secrets.
2. I like to use clever manipulation to get my way.
3. Whatever it takes, you must get the important people on your side.
4. Avoid direct conflict with others because they may be useful in the future.
5. It's wise to keep track of information that you can use against people later.
6. You should wait for the right time to get back at people.
7. There are things you should hide from other people to preserve your reputation.
8. Make sure your plans benefit yourself, not others.
9. Most people can be manipulated.

Narcissism

1. People see me as a natural leader.
2. I hate being the centre of attention.
3. Many group activities tend to be dull without me.
4. I know that I am special because everyone keeps telling me so.
5. I like to get acquainted with important people.
6. I feel embarrassed if someone compliments me.
7. I have been compared to famous people.
8. I am an average person.
9. I insist on getting the respect I deserve.

Psychopathy

1. I like to get revenge on authorities.
2. I avoid dangerous situations.
3. Payback needs to be quick and nasty.
4. People often say I'm out of control
5. It's true that I can be mean to others.
6. People who mess with me always regret it.
7. I have never gotten into trouble with the law.
8. I enjoy having sex with people I hardly know.
9. I'll say anything to get what I want.

3.2.6 Everyday Sadism: Varieties of Sadistic Tendencies (VAST; Paulhus & Jones, 2015)

Please rate your agreement or disagreement

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

1. In video games, I like the realistic blood spurts.
2. I sometimes replay my favourite scenes from gory slasher films.
3. I enjoy watching cage fighting (or MMA), where there is no escape.
4. I sometimes look away in horror movies.
5. In car-racing, it's the accidents that I enjoy most.
6. There's way too much violence in sports.
7. I love the YouTube clips of people fighting.
8. I enjoy physically hurting people.
9. I would never purposely humiliate someone.
10. I was purposely cruel to someone in high school.
11. I enjoy hurting my partner during sex (or pretending to).
12. I can dominate others using fear.
13. I enjoy making people suffer.
14. I enjoy mocking losers to their face.
15. I never said mean things to my parents.
16. I enjoy tormenting animals – especially the nasty ones.

3.2.7 Everyday Sadism: Short Sadistic Impulse Scale (SSIS; O'Meara, Davies, & Hammond, 2011)

Instructions

Please indicate which statement you associate with

0	1
Unlike Me	Like Me

1. I enjoy seeing people hurt.
2. I would enjoy hurting someone physically, sexually or emotionally.
3. Hurting people would be exciting.
4. I have hurt people for my own enjoyment.
5. People would enjoy hurting others if they gave it a go.
6. I have fantasies which involve hurting other people.
7. I have hurt people because I could.
8. I wouldn't intentionally hurt anyone.
9. I have humiliated others to keep them in line.
10. Sometimes I get so angry I want to hurt people.

3.2.8 Everyday Sadism: The Assessment of Sadistic Personality (ASP; Plouffe, Saklofske & Smith, 2017)

Instructions

Please indicate to what extent you agree with the following statements.

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree

1. Being mean to others can be exciting.
2. When I mock someone, it is funny to see them get upset.
3. I have made fun of people so that they know I am in control.
4. Watching people get into fights excites me.
5. I would hurt somebody if it meant that I would be in control.
6. I would not purposely hurt anybody, even if I didn't like them.
7. I think about hurting people who irritate me.
8. I get pleasure from mocking people in front of their friends.
9. I never get tired of pushing people around.

3.3. Examples of R Statistical Output

As there was a large amount of statistical output for each study, for brevity, examples are provided. A full appendix with all the output is available on request.

3.3.1 Univariate and Multivariate Normality

LSRP
Females 18-25

```
> mardia <- mvn(data= LSRPEmergingAdultsF, mvnTest = "mardia", univariateTest =
 "SW")
> mardia$multivariateNormality
             Test        Statistic              p value Result
1 Mardia Skewness 3918.32542414561 3.79786823627841e-14     NO
2 Mardia Kurtosis 4.64869163660037 3.34047174588292e-06     NO
3             MVN             <NA>                  <NA>     NO
```

ASP
Males 18-25

```
> mardia <- mvn(data= ASPEmergingAdultMale, mvnTest = "mardia", univariateTest =
 "SW")
> mardia$multivariateNormality
             Test        Statistic              p value Result
1 Mardia Skewness 478.134118265719 3.04827370345941e-32     NO
2 Mardia Kurtosis 6.43029285153084 1.27358346091455e-10     NO
3             MVN             <NA>                  <NA>     NO
```

Dirty Dozen
Females 26 and Over

```
> mardia <- mvn(data= DD26andoverF, mvnTest = "mardia", univariateTest = "SW")
> mardia$multivariateNormality
             Test        Statistic               p value Result
1 Mardia Skewness 1464.88093974705 9.80176557496253e-132     NO
2 Mardia Kurtosis 27.2131797776979                     0     NO
3             MVN             <NA>                  <NA>     NO
```

Mach IV
Males 26 and Over

```
> mardia <- mvn(data= MachIVover26M, mvnTest = "mardia", univariateTest = "SW")
> mardia$multivariateNormality
             Test        Statistic              p value Result
1 Mardia Skewness 2094.34915857025 1.08907284380442e-19     NO
2 Mardia Kurtosis  9.1777275028726                     0     NO
3             MVN             <NA>                  <NA>     NO
```

3.3.2 Confirmatory Factor Analysis

LSRP Females 26 and over

```
> summary(fits$LSRPov26F, standardized=TRUE, fit.measures=TRUE)
lavaan 0.6-5 ended normally after 37 iterations

  Estimator                                         DWLS
  Optimization method                           NLMINB
  Number of free parameters                        103

  Number of observations                           151

Model Test User Model:
                                              Standard      Robust
  Test Statistic                               544.164     501.641
  Degrees of freedom                               298         298
  P-value (Chi-square)                           0.000       0.000
  Scaling correction factor                                  1.602
  Shift parameter                                          162.043
    for the simple second-order correction

Model Test Baseline Model:

  Test statistic                              5274.978    2168.469
  Degrees of freedom                               325         325
  P-value                                        0.000       0.000
  Scaling correction factor                                  2.685

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.950       0.890
  Tucker-Lewis Index (TLI)                       0.946       0.880

Root Mean Square Error of Approximation:

  RMSEA                                          0.074       0.067
  90 Percent confidence interval - lower         0.064       0.057
  90 Percent confidence interval - upper         0.084       0.078
  P-value RMSEA <= 0.05                          0.000       0.004

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                        NA
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.107       0.107
```

Narcissistic Personality Inventory Females 26 and Over

Unidimensional model

```
Warning message:
In lav_model_estimate(lavmodel = lavmodel, lavpartable = lavpartable,  :
  lavaan WARNING: the optimizer warns that a solution has NOT been found!
```

Narcissistic Personality Inventory
Seven-Factor Model
Males 18-25

lavaan 0.6-5 ended normally after 173 iterations

Estimator	DWLS
Optimization method	NLMINB
Number of free parameters	100
Number of observations	156

Model Test User Model:

	Standard	Robust
Test Statistic	710.337	767.749
Degrees of freedom	680	680
P-value (Chi-square)	0.204	0.011
Scaling correction factor		2.085
Shift parameter		427.090
for the simple second-order correction		

Model Test Baseline Model:

Test statistic	3669.756	1714.882
Degrees of freedom	741	741
P-value	0.000	0.000
Scaling correction factor		3.007

User Model versus Baseline Model:

Comparative Fit Index (CFI)	0.990	0.910
Tucker-Lewis Index (TLI)	0.989	0.902

Root Mean Square Error of Approximation:

RMSEA	0.017	0.029
90 Percent confidence interval - lower	0.000	0.015
90 Percent confidence interval - upper	0.030	0.039
P-value RMSEA <= 0.05	1.000	1.000
Robust RMSEA		NA
90 Percent confidence interval - lower		NA
90 Percent confidence interval - upper		NA

Standardized Root Mean Square Residual:

SRMR	0.124	0.124

3.3.3 Measurement Invariance

ASP Males and Females Over 26

Configural Invariance

```
> summary (configuralASP26, fit.measures=TRUE)
lavaan 0.6-5 ended normally after 200 iterations

  Estimator                                         DWLS
  Optimization method                           NLMINB
  Number of free parameters                         54

  Number of observations per group:
    1                                              151
    2                                              150

Model Test User Model:
                                              Standard      Robust
  Test Statistic                                24.325      66.585
  Degrees of freedom                            54          54
  P-value (Chi-square)                          1.000       0.117
  Scaling correction factor                                 0.487
  Shift parameter for each group:
      1                                                     8.352
      2                                                     8.297
    for the simple second-order correction
  Test statistic for each group:
    1                                           11.351      31.654
    2                                           12.974      34.931

Model Test Baseline Model:

  Test statistic                               828.278     370.242
  Degrees of freedom                            72          72
  P-value                                       0.000       0.000
  Scaling correction factor                                 2.536

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                   1.000       0.958
  Tucker-Lewis Index (TLI)                      1.052       0.944

  Robust Comparative Fit Index (CFI)                           NA
  Robust Tucker-Lewis Index (TLI)                              NA

Root Mean Square Error of Approximation:

  RMSEA                                         0.000       0.039
  90 Percent confidence interval - lower        0.000       0.000
  90 Percent confidence interval - upper        0.000       0.068
  P-value RMSEA <= 0.05                         1.000       0.696

  Robust RMSEA                                                 NA
  90 Percent confidence interval - lower                     0.000
  90 Percent confidence interval - upper                       NA

Standardized Root Mean Square Residual:

  SRMR                                          0.050       0.050

Parameter Estimates:

  Information                                 Expected
  Information saturated (h1) model        Unstructured
  Standard errors                           Robust.sem
```

Metric Invariance

```
> summary (metricASP, fit.measures=TRUE)
lavaan 0.6-5 ended normally after 98 iterations

  Estimator                                         DWLS
  Optimization method                             NLMINB
  Number of free parameters                           54
  Number of equality constraints                       8
  Row rank of the constraints matrix                   8

  Number of observations per group:
    1                                                151
    2                                                150

Model Test User Model:
                                              Standard      Robust
  Test Statistic                                34.034      62.324
  Degrees of freedom                                62          62
  P-value (Chi-square)                           0.999       0.465
  Scaling correction factor                                  0.919
  Shift parameter for each group:
      1                                                     12.690
      2                                                     12.606
    for the simple second-order correction
  Test statistic for each group:
    1                                           17.970      32.240
    2                                           16.065      30.084

Model Test Baseline Model:

  Test statistic                               828.278     370.242
  Degrees of freedom                                72          72
  P-value                                        0.000       0.000
  Scaling correction factor                                  2.536

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    1.000       0.999
  Tucker-Lewis Index (TLI)                       1.043       0.999

  Robust Comparative Fit Index (CFI)                            NA
  Robust Tucker-Lewis Index (TLI)                               NA

Root Mean Square Error of Approximation:

  RMSEA                                          0.000       0.006
  90 Percent confidence interval - lower         0.000       0.000
  90 Percent confidence interval - upper         0.000       0.050
  P-value RMSEA <= 0.05                          1.000       0.951

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                     0.000
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.064       0.064

Parameter Estimates:

  Information                                 Expected
  Information saturated (h1) model        Unstructured
  Standard errors                           Robust.sem
```

3.3.4 Mokken Scale Analysis

The Assessment of Sadistic Personality Males 18-25

AISP

```
      0.4
ASP1    0
ASP2    1
ASP3    1
ASP4    1
ASP5    1
ASP6R   0
ASP7    1
ASP8    1
ASP9    1
```

Scale One

```
$method
[1] "MIIO"

$item.summary
     ItemH #ac #vi #vi/#ac maxvi sum sum/#ac tmax #tsig crit
ASP4  0.40    6    0       0     0    0       0    0     0    0
ASP7  0.46    6    0       0     0    0       0    0     0    0
ASP3  0.49    6    0       0     0    0       0    0     0    0
ASP2  0.45    7    0       0     0    0       0    0     0    0
ASP5  0.58    6    0       0     0    0       0    0     0    0
ASP8  0.53    7    0       0     0    0       0    0     0    0
ASP9  0.56    6    0       0     0    0       0    0     0    0

$backward.selection
     step 1
ASP4      0
ASP7      0
ASP3      0
ASP2      0
ASP5      0
ASP8      0
ASP9      0

$HT
[1] 0.1461171
```

```
$MS
[1] 0.8465838

$alpha
[1] 0.8431435

$lambda.2
[1] 0.846603
```

```
    ASP2     ASP3     ASP4     ASP5     ASP7     ASP8     ASP9
2.006410 2.006410 2.275641 1.820513 2.192308 1.685897 1.628205
```

```
> coefH(ASPM18MS1)
$Hij
     ASP2    se      ASP3    se      ASP4    se      ASP5    se      ASP7
ASP2                  0.558  (0.069)  0.241  (0.085)  0.540  (0.064)  0.352
ASP3  0.558  (0.069)                  0.329  (0.087)  0.541  (0.078)  0.474
ASP4  0.241  (0.085)  0.329  (0.087)                  0.557  (0.080)  0.480
ASP5  0.540  (0.064)  0.541  (0.078)  0.557  (0.080)                  0.534
ASP7  0.352  (0.096)  0.474  (0.087)  0.480  (0.080)  0.534  (0.081)
ASP8  0.595  (0.072)  0.488  (0.078)  0.398  (0.080)  0.608  (0.076)  0.423
ASP9  0.445  (0.062)  0.561  (0.067)  0.441  (0.070)  0.736  (0.053)  0.471
     se      ASP8    se      ASP9    se
ASP2 (0.096)  0.595  (0.072)  0.445  (0.062)
ASP3 (0.087)  0.488  (0.078)  0.561  (0.067)
ASP4 (0.080)  0.398  (0.080)  0.441  (0.070)
ASP5 (0.081)  0.608  (0.076)  0.736  (0.053)
ASP7          0.423  (0.084)  0.471  (0.079)
ASP8 (0.084)                  0.704  (0.054)
ASP9 (0.079)  0.704  (0.054)

$Hi
     Item H  se
ASP2   0.452 (0.046)
ASP3   0.487 (0.050)
ASP4   0.401 (0.057)
ASP5   0.581 (0.043)
ASP7   0.456 (0.058)
ASP8   0.532 (0.048)
ASP9   0.557 (0.042)

$H
 Scale H se
   0.492 (0.040)
```

Appendix 4

Reconstructing the Dark Tetrad: Preliminary Development of The Dark Side of Humanity Scale.

4.1 Indicators for the Item Pool

Machiavellian

Items from Study 1

I tend to manipulate others to get my way.

I have used deceit or lied to get my way.

I like to use clever manipulation to get my way

It's wise to keep track of information that you can use against people

It is hard to get ahead without cutting corners here and there.

Anyone who completely trusts anyone is asking for trouble.

Never tell anyone the real reason you did something unless it is useful to do so.

One should take action only when it is morally right.

Honesty is the best policy in all cases

Misanthropy

Five Factor Machiavellianism Inventory (FFMI; Collison et al., 2018).

Original - It is important to be wary of others motives.

Change – I am wary of others motives.

Item Pool of the Mach IV (Christie & Geis, 1970)

Original – Friends should be chosen with an eye toward what they might be able to do for you.

Change – I chose my friends by what they can do for me.

Original – Never tell anyone the real reason you did something unless it is useful to do so.

Change – I never tell anyone the real reason I did something, unless it is useful to do so.

Original – Anyone who completely trusts anyone else is asking for trouble.

Change – If I completely trust anyone else, I'm asking for trouble.

Original – Never trust anyone who has a grudge against you.

Change – I don't trust anyone who holds a grudge against me.

Mach III (Christie & Geis, 1970)

I have a great deal of faith in human nature (R)

I will go out of my way to help other people (R)

Machiavellian Personality Scale (MPS; Dahling et al., 2009).

Original – I dislike committing to groups because I don't trust others.

Change – I don't like committing to groups because I don't trust other people.

If I show any weakness, other people will take advantage of it.

Original – The only good reason to talk to others is to get information that I can use to my benefit.

Change - The only good reason I talk to others is to get information that I can use to my benefit.

Immorality

Item Pool of the Mach IV (Christie & Geis, 1970)

Original – One should only take action when sure it is morally right.

Change – I only take action when it is morally right.

Original – Honesty is the best policy in all cases.

Change – For me, honesty is the best policy in all cases.

Original – If one is morally right, compromise is out of the question.

Change – If I feel I am morally right then compromise is out of the question.

Original – The most important thing in life is winning.

Change – For me, the most important thing in life is winning.

Original – There is no point in keeping a promise if it is to your advantage to break it.

Change – I will break a promise if it works to my advantage.

Original – It's hard to get ahead without cutting corners here and there.

Change – It's hard for me to get ahead without cutting corners here and there.

Mach III (Christie & Geis, 1970)

I'm not crafty or sly (R)

Original - Being honest all the time won't lead to success

Change - If I'm honest all the time it won't lead to the success of my objectives.

Machiavellian Personality Scale (MPS; Dahling et al., 2009).

I am willing to be unethical if I believe it will help me succeed.

I believe that lying is necessary to maintain a competitive advantage over others.

I am willing to sabotage the efforts of other people if they threaten my own goals.

Five Factor Machiavellianism Inventory (FFMI; Collison et al., 2018).

Original – Sometimes you have to lie to get things done.

Change – Sometimes I have to lie to get things done.

Ambitious/calculating

Mach III (Christie & Geis, 1970)

I am not an ambitious person (R)

I feel inferior to others (R)

When I'm under a great deal of stress, sometimes I feel like I'm going to pieces (R)

I like to carefully consider the consequences before I make a decision

It's easy for me to outsmart my peers.

I am confident interacting with others

I am very sure of myself.

I am a very persuasive person.

People would describe me as emotionally stable.

Machiavellian Personality Scale (MPS; Dahling et al., 2009).

I enjoy being able to control the situation

I enjoy having control over other people.

Original - I like to give the orders in interpersonal situations.

Change – I like to give orders to others.

Political Skill Inventory (Ferris et al., 2005)

I am particularly good at sensing the motivations and hidden agendas of others.

Five Factor Machiavellianism Inventory (FFMI; Collison et al., 2018)

I do not have a problem with speaking my mind.

Strategic Planning

Five Factor Machiavellianism Inventory (FFMI; Collison et al., 2018)

I never seem to be able to get organised (R)

I like to map out projects before I begin.

Mach III (Christie & Geis, 1970)

'Act first, think later', describes me well.

I am not easily flustered.

I prefer to be spontaneous rather than planning everything out (R)

I don't make many spur of the moment decisions,

I like to map out projects before I begin

Machiavellian Personality Scale (MPS; Dahling et al., 2009).

I like to share my plans and ideas with other people (R)

Locus of Control of Behaviour (Craig et al., 1984)

I can anticipate difficulties and take action to avoid them.

When I make plans, I am almost certain that I can make them work

Political Skill Inventory (Ferris et al., 2005)

I size up situations before deciding how to present an idea to others.

Psychopathy

Items from Study 1

I tend to lack remorse.

I tend to be unconcerned with the morality of my actions.

I'll say anything to get what I want.

I tell other people what they want to hear so that they will do what I want them to.

Success is based on survival of the fittest, I am not concerned about the losers.

My main purpose in life is getting as many goodies as I can.

For me, what's right is whatever I can get away with.

In today's world, I feel justified in doing anything I can get away with.

Callous Indifference

Development and validation of an Expanded version of the three-factor LSRP (Christian & Sellbom 2016).

I'm not a very emotional person

I tend to cry in sad movies (R)

I feel bad when I do something wrong (R)

My friends consider me a warm person (R)

Psychopathic Personality Traits Scale (Boduszek et al., 2016).

Before criticising somebody, I try to imagine and understand how it would make them feel (R)

What other people feel doesn't concern me

Seeing people cry doesn't really upset me

I find it difficult to understand what other people feel.

The Abbreviated Psychopathy Measure (APM; Semel, 2018)

I can be good at pretending to care about people, but must of the time I really don't care

I could look people straight in the eye and it means nothing to me to lie or to cheat them.

People might describe me as mean and cruel.

Triarchic Psychopathy Brief Measure (Patrick et al., 2009)

I don't care much if what I do hurts others

Self-Report Psychopathy Scale II (Williams & Paulhus, 2007)

It's sometimes fun to see how far I can push someone before they catch on.

Egocentricity

Psychopathic Personality Traits Scale (Boduszek et al., 2016).

I tend to focus on my own thoughts and ideas rather than on what others might be thinking

Original - It's natural for human behaviour to be motivated by self-interest

Change – It's natural for my behaviour to be motivated by self-interest.

Triarchic Psychopathy Brief Measure (Patrick et al., 2009)

I have a hard time waiting patiently for things I want

The Abbreviated Psychopathy Measure (APM; Semel, 2018)

Playing by the rules sounds nice, but getting what I want is more important

Manipulation

The concept and measurement of social psychopathy (Smith, 1985)

Original - There is always a way to get someone to trust you.

Change – I always find a way to get someone to trust me.

Triarchic Psychopathy Brief Measure (Patrick et al., 2009)

I can convince people to do what I want

Psychopathic Personality Traits Scale (Boduszek et al., 2016).

I know how to pay someone compliments to get something out of them

I know how to simulate emotions like pain and hurt to make others feel sorry for me.

I know how to make another person feel guilty

I sometimes provoke people on purpose to see their reaction

Fearless

Triarchic Psychopathy Brief Measure (Patrick et al., 2009)

I jump into things without thinking

I am well equipped to deal with stress

I can get over things that would traumatise others

I find things are more fun if a little danger is involved

I'm afraid of far fewer things than most people

I stay away from physical danger as much as I can (R)

It worries me to go into an unfamiliar situation without knowing all the details (R)

I get scared easily (R)

I function well in new situations, even when unprepared

Development and validation of an Expanded version of the three-factor LSRP (Christian & Sellbom 2016).

I know rules are there, but I don't tend to follow them

The Abbreviated Psychopathy Measure (APM; Semel, 2018)

I am willing to take great risks in life without fear of loss or failure.

Everyday Sadism

In relation to everyday sadism, as well as the items derived from study one, and those from other validated scales e.g., Schadenfreude scale (Crysel & Webster, 2018), nine items have been written which tap into sadism in the form of online trolling. For individuals who embody everyday sadism, evidence suggests that this behaviour is an occurrence of everyday sadism, such that trolling is driven by sadistic pleasure and is a significant positive predictor of trolling behaviour, where individuals experience feelings of pleasure from hurting others (Buckels, Trapnell et al., 2019; Sest & March, 2017), thus providing justification for inclusion of scale items tapping into online trolling.

Items from Study 1

I have fantasies which involve hurting other people.
Hurting people would be exciting.
I would enjoy hurting someone physically, sexually or emotionally.
I enjoy seeing people hurt.
I have humiliated others to keep them in line.
I think about hurting people who irritate me.
I have made fun of people so that they know I am in control.
When I mock someone, it is funny to see them get upset.
I never get tired of pushing people around.
I would hurt somebody if it meant that I would be in control.
I get pleasure from mocking people in front of their friends.
I can dominate others using fear.
I enjoy mocking losers to their face.
I enjoy making people suffer.
I enjoy physically hurting people.
I enjoy tormenting animals – especially the nasty ones.

Vicarious Sadism

Schadenfreude scale (Crysel & Webster, 2018).

I laugh when someone just misses the bus

It's funny when a person walks into a closed sliding door.

I feel good when I see people walking in the rain while I'm driving my car.

Original - I like to see someone successful get fired.

Change – It give me pleasure to see someone successful get fired.

I take pleasure in seeing someone trip over.

I enjoy watching videos where performers may get hurt.

I enjoy watching people fight (either on TV or in real life)

Original – I like to watch people in pain

Change – I enjoy watching people in pain.

I enjoy watching the gory murder scenes in films over and over again.

Sadistic Fantasies

Social Dominance Orientation Scale (Pratto et al., 1994)

Original – We should use more graphic forms of executions (such as a firing squad).

Change - I would enjoy witnessing graphic forms of executions, such as a firing squad.

Original – The crime problems have gotten so bad that maybe we should bring back public hangings.

Change - The crime problem has become so bad that if we brought back public hangings, I would take great pleasure from watching them.

Online Sadism

Items written for the item pool

Posting offensive comments on social media forums and seeing people's reactions is a source of entertainment for me.

I get a lot of pleasure from being abusive to people I don't know on forums such as twitter.

Trolling behaviour is both cruel and unnecessary ®

I take pleasure from seeing people's reactions when I post offensive images on social media forums.

When famous women post pictures of themselves on social media forums, it gives me so much pleasure to tell them that they are fat and ugly.

I enjoy watching other people get abused on social media forums.
I post offensive comments on social media forums just so I can take pleasure from the hurt I cause.
Being anonymous on social media forums lets me enjoy abusing people with different religious beliefs to my own.
Social media forums provide me with an anonymous space to enjoy being cruel to others.
Social media forums give me the opportunity to enjoy insulting people for their sexual orientations.
Trolling behaviour makes me angry ®
I take pleasure in being the cause of someone else's pain.

Narcissism

Items from Study One

I like to show off my body.

I tend to seek prestige or status.

I tend to expect special favours from others.

Grandiose Narcissism

Entitlement

The Five Factor Narcissism Inventory Short Form (FFNI; Sherman et al., 2015).

I deserve to receive special treatment

I don't think the rules apply to me as much as they apply to others

I only associate with people of my calibre.

I do not waste my time hanging out with people who are beneath me.

Grandiose Narcissism Scale (GNS: Foster et al., 2015).

I expect people to bend the rules for me

I deserve to get what I want.

I deserve more out of life than other people.

Hypersensitive Narcissism Scale (HNS: Cheek et al., 2013).

I resent others who have what I lack

Exploitativeness

Pathological Narcissism Inventory (Pincus et al., 2009)

I can usually talk my way out of anything.

I find it easy to manipulate people.

I can make anyone believe anything I want them to.

I find it easy to manipulate people.

The Five Factor Narcissism Inventory Short Form (FFNI; Sherman et al., 2015).

It is easy to get people to do what I want

Grandiose Narcissism Scale (GNS: Foster et al., 2015).

If I have to take advantage of somebody to get what I want, so be it

I've been known to use people to get what I want

Grandiose Fantasies

Pathological Narcissism Inventory (Pincus et al., 2009)

I often fantasise about being rewarded for my efforts.

I often fantasise about being recognised for my accomplishments.

I often fantasise about being admired and respected.

I often fantasise about having a huge impact on the world around me

I want to amount to something in the eyes of the world

Hypersensitive Narcissism Scale (HNS:Cheek et al., 2013).

I sometimes have fantasies about being violent without knowing why

Entitlement Rage

Pathological Narcissism Inventory (Pincus et al., 2009)

It irritates me when people don't notice how good I am.

I get mad when people don't notice all that I do for them

I can get pretty angry when others disagree with me

The Five Factor Narcissism Inventory Short Form (FFNI; Sherman et al., 2015).

It really makes me angry when I don't get what I deserve.

I hate being criticised so much that I can't control my temper when it happens

I have at times gone into a rage when not treated right

Grandiose and Vulnerable Narcissism (GVN; Rohmann et al., 2015).

Original - I can get seethingly angry if I am criticised unjustly.

Change - I fly into a fury if I am criticised unjustly.

I fly into a rage if somebody expects me to do tasks that are really beneath my skill level.

I get into a temper if I don't get the recognition that I deserve.

If something doesn't work out, I could just explode.

I can get really nasty if I don't get what I want.

Shame

Hypersensitive Narcissism Scale (HNS:Cheek et al., 2013).

My feelings are easily hurt by ridicule or by the slighting remarks of others

I tend to feel humiliated when criticised.

Defeat or disappointment usually shame or anger me but I try not to show it

The Five Factor Narcissism Inventory Short Form (FFNI; Sherman et al., 2015).

I feel ashamed when people judge me.

I feel foolish when I make a mistake in front of others.

I feel awful when I get put down in front of others

When I realise I have failed at something I feel humiliated

I'm pretty indifferent to the criticism of others ®

Pathological Narcissism Inventory (Pincus et al., 2009)

I often hide my needs for fear that others will see me as needy and dependent.

Contingent self-esteem

Grandiose and Vulnerable Narcissism (GVN; Rohmann et al., 2015).

I'm a person who needs a lot of approval from others

Pathological Narcissism Inventory (Pincus et al., 2009)

My self-esteem fluctuates a lot

When people don't notice me, I start to feel bad about myself

It's hard for me to feel good about myself when I'm alone.

I am preoccupied with thoughts and concerns that most people are not interested in me

When others don't notice me, I start to feel worthless

It's hard to feel good about myself unless I know other people admire me

I sometimes need important others in my life to reassure me of my self-worth.

The Five Factor Narcissism Inventory Short Form (FFNI; Sherman et al., 2015).

I don't really care what others think of me ®

I often feel as if I need compliments from others in order to be sure of myself

Devaluing

Pathological Narcissism Inventory (Pincus et al., 2009)

When others disappoint me I often get angry at myself

I sometimes feel ashamed about my expectations of others when they disappoint me

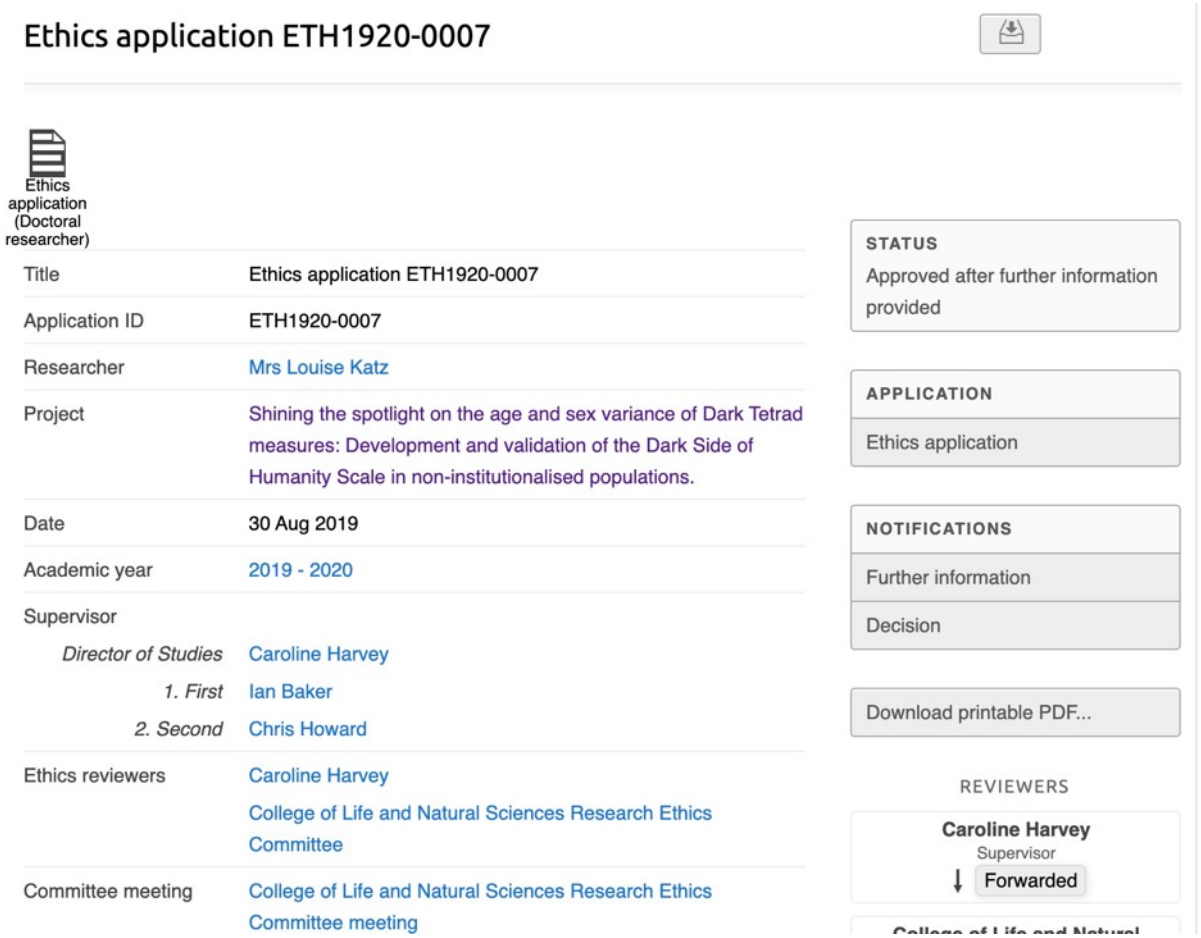
Sometimes I avoid people because I'm concerned they won't acknowledge what I do for them

Sometimes I avoid people because I'm afraid they won't do what I want them to do.

Sometimes it's easier to be alone than to face not getting everything I want from other people

Sometimes I avoid people because I'm concerned that they'll disappoint me

4.2 Ethical Approval Application



Ethics application ETH1920-0007

Ethics application (Doctoral researcher)

Title	Ethics application ETH1920-0007
Application ID	ETH1920-0007
Researcher	Mrs Louise Katz
Project	Shining the spotlight on the age and sex variance of Dark Tetrad measures: Development and validation of the Dark Side of Humanity Scale in non-institutionalised populations.
Date	30 Aug 2019
Academic year	2019 - 2020
Supervisor	
Director of Studies	Caroline Harvey
1. First	Ian Baker
2. Second	Chris Howard
Ethics reviewers	Caroline Harvey College of Life and Natural Sciences Research Ethics Committee
Committee meeting	College of Life and Natural Sciences Research Ethics Committee meeting

STATUS
Approved after further information provided

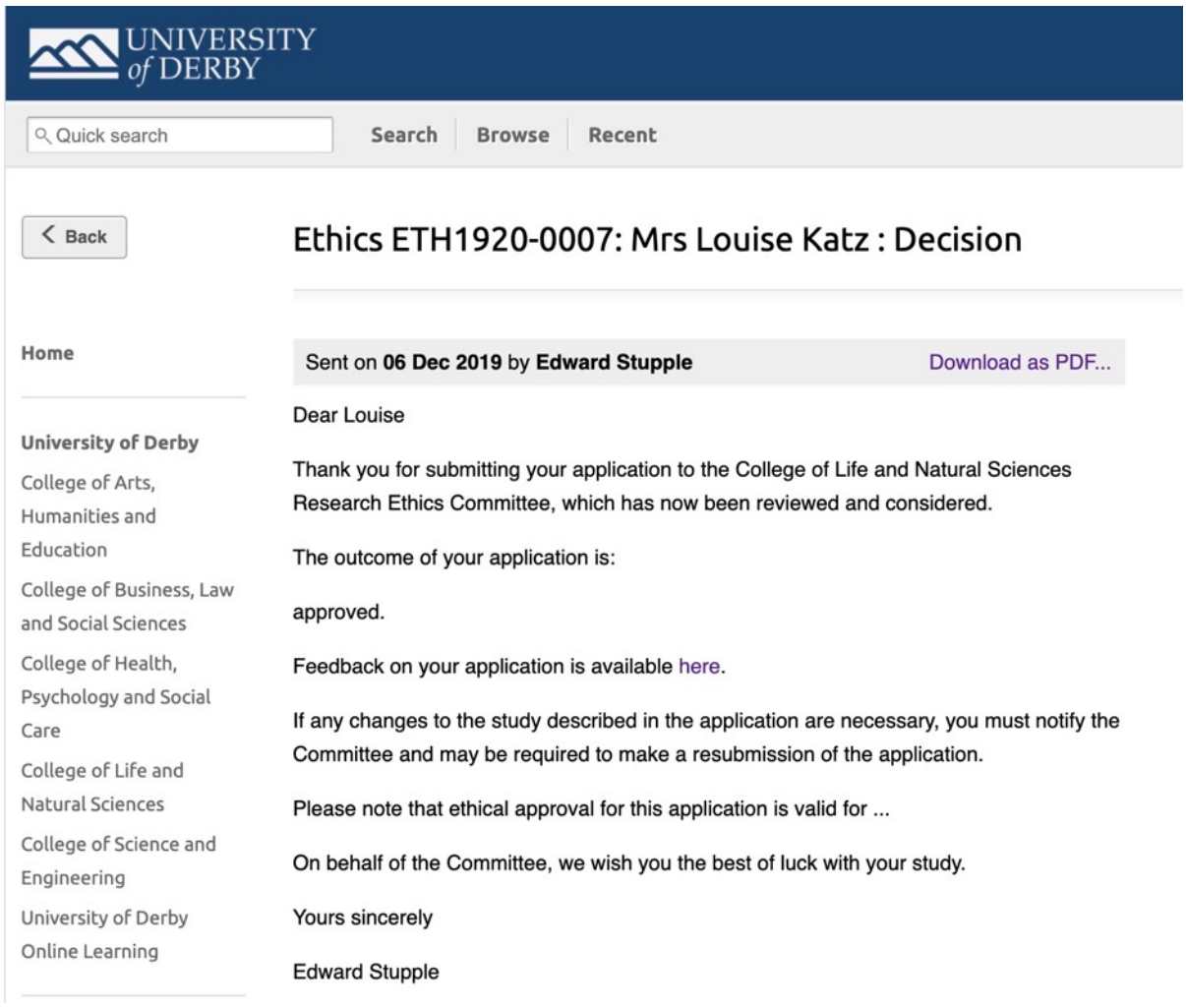
APPLICATION
Ethics application

NOTIFICATIONS
Further information
Decision

Download printable PDF...

REVIEWERS
Caroline Harvey
Supervisor
Forwarded

4.2.1 Ethical Approval Confirmation



UNIVERSITY of DERBY

Quick search | Search | Browse | Recent

Back

Ethics ETH1920-0007: Mrs Louise Katz : Decision

Home

University of Derby
College of Arts, Humanities and Education
College of Business, Law and Social Sciences
College of Health, Psychology and Social Care
College of Life and Natural Sciences
College of Science and Engineering
University of Derby Online Learning

Sent on **06 Dec 2019** by **Edward Stupple** Download as PDF...

Dear Louise

Thank you for submitting your application to the College of Life and Natural Sciences Research Ethics Committee, which has now been reviewed and considered.

The outcome of your application is:

approved.

Feedback on your application is available here.

If any changes to the study described in the application are necessary, you must notify the Committee and may be required to make a resubmission of the application.

Please note that ethical approval for this application is valid for ...

On behalf of the Committee, we wish you the best of luck with your study.

Yours sincerely

Edward Stupple

Appendix 4.3

Examples of R Statistical Output.

4.3.1 Network Redundancy Analysis

Everyday Sadism

Males 26 and Over

```
$reduced$data
# A tibble: 158 x 26
      S1    S2    S4    S5    S7    S8   S10   S12   S13   S15   S17   S18   S21
   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
 1     1     1     1     1     1     1     1     1     1     1     1     1     1
 2     1     2     1     1     1     1     1     1     1     1     1     1     1
 3     1     2     1     1     4     1     1     1     1     1     1     1     1
 4     2     2     2     2     2     2     2     3     2     1     1     2     2
 5     1     1     1     1     1     1     1     1     1     1     1     1     1
 6     2     3     1     1     1     1     1     1     1     1     1     1     1
 7     1     1     1     1     1     1     1     1     1     1     1     1     1
 8     1     4     1     1     1     1     1     1     1     1     1     1     1
 9     4     4     1     1     1     1     1     1     1     1     1     1     1
10     2     2     2     4     2     2     2     4     3     2     1     2     1
# … with 148 more rows, and 13 more variables: S22 <dbl>, S23 <dbl>, S24 <dbl>,
#   S25 <dbl>, S26 <dbl>, S28 <dbl>, S30 <dbl>, S32 <dbl>, S33 <dbl>, S34 <dbl>,
#   S36 <dbl>, S37 <dbl>, S39 <dbl>

$reduced$merged
     [,1]  [,2]  [,3] [,4]  [,5]  [,6]  [,7] [,8]  [,9]  [,10] [,11] [,12] [,13]
[1,] "S27" "S31" "S8" "S12" "S15" "S23" "S3" "S34" "S9"  "S13" "S20" "S29" "S37"
[2,] "S24" "S32" "S6" "S11" "S16" "S19" "S4" "S35" "S10" "S14" "S25" "S32" "S38"
```

Machiavellianism

Females 26 and Over

```
$reduced
$reduced$data
# A tibble: 177 x 30
      M1    M2    M3    M4    M5    M7   M10   M11   M12   M14   M15   M18   M20
   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
 1     5     1     1     5     1     4     2     6     2     3     2     3     5
 2     4     3     4     3     3     4     1     4     2     2     2     4     4
 3     4     2     1     3     3     4     2     3     2     2     1     3     4
 4     4     1     1     5     3     3     1     2     2     2     4     3     4
 5     5     2     2     2     2     2     1     4     2     2     2     3     4
 6     4     2     1     3     1     4     1     3     1     1     1     3     4
 7     5     2     4     5     4     4     3     5     2     3     4     4     3
 8     3     1     1     3     2     4     2     4     2     2     1     4     2
 9     4     2     3     5     2     4     3     6     2     4     2     4     4
10     2     1     1     6     1     6     1     5     1     5     1     6     3
# … with 167 more rows, and 17 more variables: M21 <dbl>, M22 <dbl>, M23 <dbl>,
#   M24 <dbl>, M27 <dbl>, M30 <dbl>, M31 <dbl>, M32 <dbl>, M33 <dbl>, M35 <dbl>,
#   M39 <dbl>, M40 <dbl>, M41 <dbl>, M43 <dbl>, M45 <dbl>, M48 <dbl>, M50 <dbl>

$reduced$merged
     [,1]  [,2]  [,3] [,4]  [,5]  [,6]  [,7]  [,8]  [,9]  [,10] [,11] [,12]
[1,] "M25" "M35" "M8" "M2"  "M29" "M34" "M40" "M42" "M43" "M50" "M1"  "M6"
[2,] "M26" "M46" "M7" "M17" "M30" "M39" "M44" "M49" "M37" "M52" "M9"  "M7"
[3,] "M28" "M47" ""   ""    "M16" "M13" "M36" "M41" "M38" ""    ""    ""
[4,] "M24" "M53" ""   ""    ""    ""    ""    ""    ""    ""    ""    ""
     [,13] [,14]
[1,] "M19" "M48"
[2,] "M20" "M51"
[3,] ""    ""
[4,] ""    ""


$adhoc
$redundant
$redundant$M41
[1] "M43"
```

Narcissism

Males 18-25

```
attr(,"class")
[1] "node.redundant"

$reduced
$reduced$data
# A tibble: 175 x 33
      N1    N2    N3    N4    N5    N6    N7    N8    N9   N10   N11   N19   N22
   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
 1     3     4     4     3     3     3     4     4     2     1     5     5     2
 2     6     6     3     3     4     4     3     3     4     5     4     6     5
 3     2     2     3     3     3     2     3     2     3     2     3     2     3
 4     3     3     3     2     1     2     1     3     2     1     3     4     2
 5     1     1     1     1     1     1     1     1     1     1     6     5     1
 6     1     2     2     2     2     1     2     3     2     2     3     3     4
 7     2     3     3     3     3     5     4     4     3     3     5     3     5
 8     2     2     3     3     3     3     3     4     3     3     4     2     3
 9     4     3     3     4     3     3     2     3     2     3     4     4     3
10     3     1     1     1     3     1     4     3     1     1     1     4     1
# … with 165 more rows, and 20 more variables: N23 <dbl>, N24 <dbl>, N25 <dbl>,
#   N27 <dbl>, N28 <dbl>, N29 <dbl>, N31 <dbl>, N32 <dbl>, N33 <dbl>, N34 <dbl>,
#   N37 <dbl>, N39 <dbl>, N40 <dbl>, N41 <dbl>, N43 <dbl>, N44 <dbl>, N47 <dbl>,
#   N54 <dbl>, N57 <dbl>, N58 <dbl>

$reduced$merged
     [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]  [,9]  [,10] [,11] [,12]
[1,] "N14" "N18" "N49" "N24" "N35" "N50" "N55" "N30" "N40" "N41" "N43" "N53"
[2,] "N15" "N19" "N52" "N26" "N36" "N51" "N56" "N31" "N45" "N42" "N46" "N54"
[3,] "N11" "N20" "N44" ""    "N37" "N47" "N58" ""    ""    ""    ""    ""
[4,] "N12" "N17" "N48" ""    "N38" ""    ""    ""    ""    ""    ""    ""
[5,] "N13" "N21" ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
[6,] "N16" ""    ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
```

Psychopathy

Females 18-25

```
$reduced
$reduced$data
# A tibble: 157 x 22
      P2    P4    P7    P9   P10   P12   P13   P14   P15   P17   P20   P21   P22
   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
 1     3     4     4     5     5     3     5     3     3     5     4     5     4
 2     2     3     2     6     3     4     4     4     1     2     5     4     3
 3     5     5     4     6     3     4     4     4     4     3     3     5     4
 4     6     1     1     5     1     2     4     1     2     1     2     6     2
 5     1     4     1     4     4     3     4     2     2     4     5     6     2
 6     4     1     2     2     2     2     4     4     3     3     4     3     3
 7     3     1     1     1     1     1     1     1     1     1     1     1     1
 8     2     2     1     1     1     1     1     1     1     1     1     3     1
 9     1     3     1     1     1     2     1     1     1     1     1     1     1
10     3     4     1     5     6     6     6     5     5     2     6     1     1
# … with 147 more rows, and 9 more variables: P23 <dbl>, P27 <dbl>, P30 <dbl>,
#   P33 <dbl>, P35 <dbl>, P36 <dbl>, P37 <dbl>, P39 <dbl>, P40 <dbl>


$reduced$merged
     [,1]  [,2]  [,3]  [,4]  [,5]  [,6] [,7]  [,8] [,9]  [,10] [,11] [,12] [,13]
[1,] "P24" "P10" "P15" "P17" "P32" "P5" "P8"  "P6" "P13" "P19" "P27" "P30" "P34"
[2,] "P25" "P11" "P16" "P18" "P40" "P7" "P39" "P7" "P29" "P20" "P28" "P41" "P35"
[3,] "P26" ""    ""    ""    "P31" "P1" "P3"  ""   ""    ""    ""    ""    ""
[4,] "P23" ""    ""    ""    ""    ""   ""    ""   ""    ""    ""    ""    ""
     [,14]
[1,] "P37"
[2,] "P38"
[3,] ""
[4,] ""
```

4.3.2 Generalised Partial Credit Model

Initial analysis with all homogenous items.

Machiavellianism

Males 18-25 (Text output from R)

	a	b1	b2	b3	b4	b5
M1	0.05235654	-26.62392783	-12.51246504	-22.8626924	7.637269	31.1879246
M3	0.67018261	-1.01830658	-0.21862891	-0.5909878	1.799310	2.2755265
M4	0.27912071	-7.58766014	-1.69815161	1.0815025	1.280684	4.4312224
M10	0.68725613	0.03392839	0.89206166	0.7870773	1.735319	5.1243153
M11	0.17541555	-3.59864886	-2.07469739	-2.9401276	1.279536	2.5691877
M12	0.12435585	-3.48371211	1.47167745	4.6410467	6.816061	3.9683650
M14	0.47585176	0.46107716	0.49741715	-0.1408739	2.481201	4.3884458
M19	0.55472695	-1.57383857	0.16011735	1.2277621	2.906738	3.8982825
M20	0.14799992	-5.04488008	-9.90445285	-3.0127125	2.717925	10.9461166
M21	0.14918626	-2.42604726	-4.31544246	-2.4377804	4.038921	14.4302449
M23	1.27144258	-1.67821651	-0.06924927	-0.2026344	1.615006	1.9857441
M24	0.73088251	-1.88759203	-1.28589533	0.4207110	1.504965	2.3799958
M26	1.16185831	-0.59404826	0.13733275	0.6190678	1.713431	2.2035438
M28	1.44821379	-0.54772503	-0.23534531	0.3463472	1.416872	1.6767172
M29	1.78879777	-0.42757340	0.09712369	0.6524580	1.894422	2.3380116
M39	-0.12646726	12.64343636	1.71642127	5.3914579	1.968828	-6.5064114
M40	0.67066300	-1.90404836	-1.31406821	-0.2523836	1.838739	0.8738634

Everyday Sadism
Females 18-25

	a	b1	b2	b3	b4	b5
S4	1.2572043	1.0908501	0.9862076	1.178124	2.791320	2.2524221
S8	3.3692582	0.6793316	1.2921117	1.314171	1.703052	1.8777160
S10	3.6306835	0.8133364	1.1923537	1.492317	1.556591	2.3349301
S12	1.1336198	0.8072814	1.2578123	0.645284	2.677474	1.1628284
S13	2.4413124	0.7525332	1.3266812	1.174528	2.001032	1.4905616
S15	1.8080692	0.8822487	1.3085014	1.471502	2.484789	1.3686798
S18	0.9948632	1.6232764	1.6474778	1.026564	2.819169	2.3520855
S24	1.8111794	1.4827023	1.7488609	1.148819	1.660934	NA
S32	3.0040119	0.9182960	1.5009608	2.091057	1.316884	NA
S34	1.3613894	1.3353738	1.4060835	1.791961	2.349765	0.5596147
S36	7.5867274	1.0925425	1.4369371	1.571591	2.196124	NA

4.3.3 Final Generalised Partial Credit Model

Machiavellian

Males 18-25

```
> coef.gpcm <- coef(results.M18machmirtFinal1, IRTpars=TRUE, simplify=TRUE)
> items.gpcm <- as.data.frame(coef.gpcm$items)
> print(items.gpcm)
            a          b1          b2         b3       b4       b5
M3  0.6633032 -1.01481651 -0.22291902 -0.6042577 1.812728 2.294412
M10 0.6868813  0.03103741  0.88784077  0.7897528 1.745788 5.140576
M23 1.1625789 -1.71333054 -0.04955068 -0.2513948 1.692381 2.029187
M24 0.7173931 -1.88925564 -1.30395322  0.4210802 1.522135 2.407643
M26 1.2377638 -0.60379058  0.12210916  0.6153224 1.690753 2.193758
M28 1.3790240 -0.53621575 -0.25241081  0.3392062 1.451913 1.696052
M29 1.8566391 -0.43904447  0.08779288  0.6536623 1.891345 2.332530
```

Females 18-25

```
> coef.gpcm <- coef(results.F18machmirtFinal1, IRTpars=TRUE, simplify=TRUE)
> items.gpcm <- as.data.frame(coef.gpcm$items)
> print(items.gpcm)
            a         b1          b2          b3       b4       b5
M3  0.7635350 -0.9846804  0.6639566 0.07350727 2.983956 1.834217
M10 1.1793061  0.1024378  1.4827507 1.10403129 2.501162 2.459062
M23 1.2299866 -0.9206093  0.5183726 0.41869164 1.995508       NA
M24 0.8270002 -1.4585525 -0.1123714 0.55098666 1.962606 2.864868
M26 1.1426920 -0.5943419  0.6565469 0.93286173 2.664961 2.625739
M28 1.9698310 -0.5018235  0.5856581 1.32170543 1.892676       NA
M29 1.5298590  0.4984075  0.8540576 1.40204349 2.068651       NA
```

Males 26 and Over

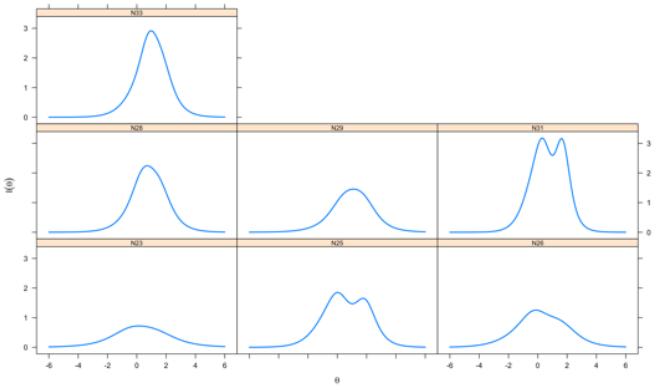
```
> coef.gpcm <- coef(results.M26machmirtFinal1, IRTpars=TRUE, simplify=TRUE)
> items.gpcm <- as.data.frame(coef.gpcm$items)
> print(items.gpcm)
            a          b1         b2          b3       b4       b5
M3  0.5483581 -0.68120107  0.6779030  0.09126751 3.249687 1.240913
M10 0.8254504  0.02939898  0.9549459  0.96728922 1.727916 4.216438
M23 0.9951211 -1.25569492  0.4918039 -0.52496937 2.106261 2.417067
M24 0.6624283 -2.01504538 -0.5009886  0.14525598 2.929952 2.078218
M26 1.6205588 -0.67617910  0.3975903  0.59701883 1.811088 2.583014
M28 1.1472173 -0.51222673  0.4052045  0.78247678 2.731578 1.802325
M29 2.0736657 -0.14779494  0.8392848  1.06311695 2.098616 1.635418
```

Females 26 and Over

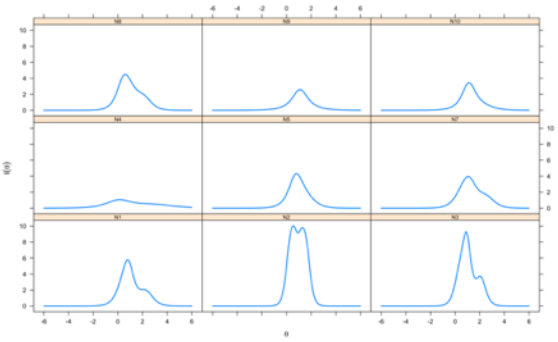
```
> coef.gpcm <- coef(results.F26machmirtFinal1, IRTpars=TRUE, simplify=TRUE)
> items.gpcm <- as.data.frame(coef.gpcm$items)
> print(items.gpcm)
            a         b1        b2        b3       b4       b5
M3  0.4663820  0.2679295 1.6745134 0.5779726 3.876529       NA
M10 0.6447838  0.6500217 2.3985420 2.6234313 1.543917 3.682133
M23 1.0094734 -0.1290756 1.3015432 0.9257826 3.640602       NA
M24 0.7782698 -0.7916877 0.5135583 0.4156558 2.513040 3.341665
M26 1.8307472 -0.4201697 0.6858831 1.0792064 2.222015       NA
M28 2.9809324 -0.0418781 0.6532347 1.0850167 2.309705       NA
M29 1.4022079  0.5304415 1.7631955 1.3515603 3.223923 2.021161
```

4.3.4 Item information Curves

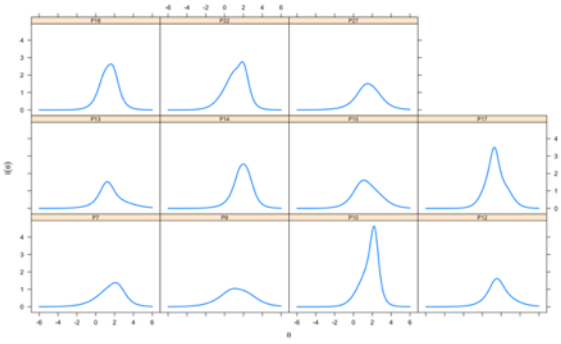
Vulnerable Narcissism
Males 18-25



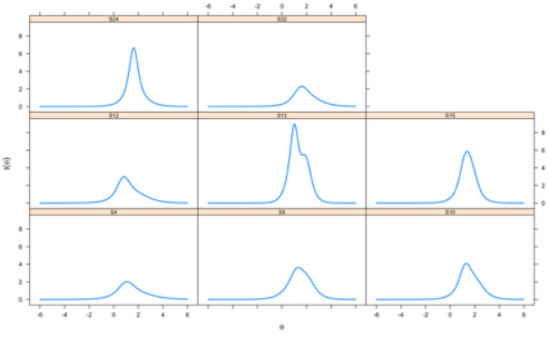
Grandiose Narcissism
Females 18-25



Psychopathy
Females 26 and over



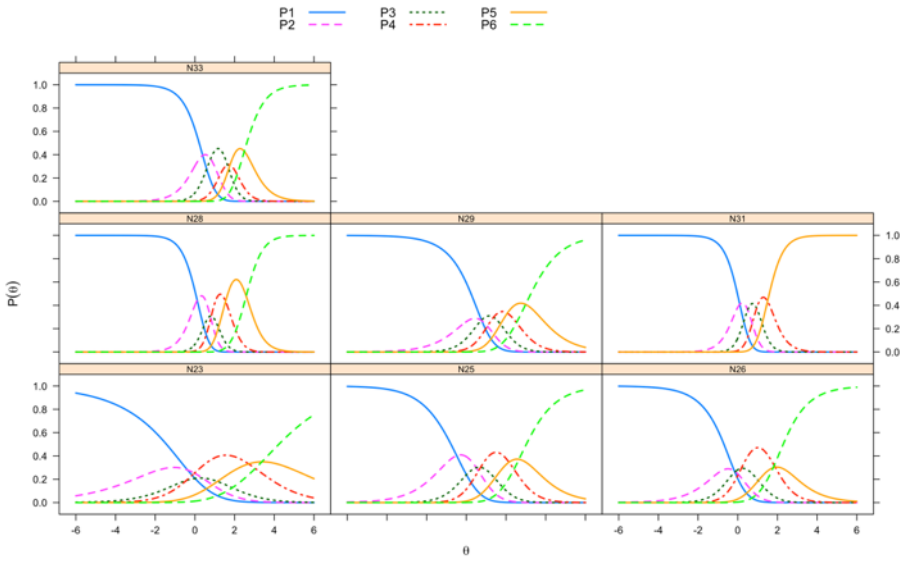
Everyday Sadism
Males 26 and Over



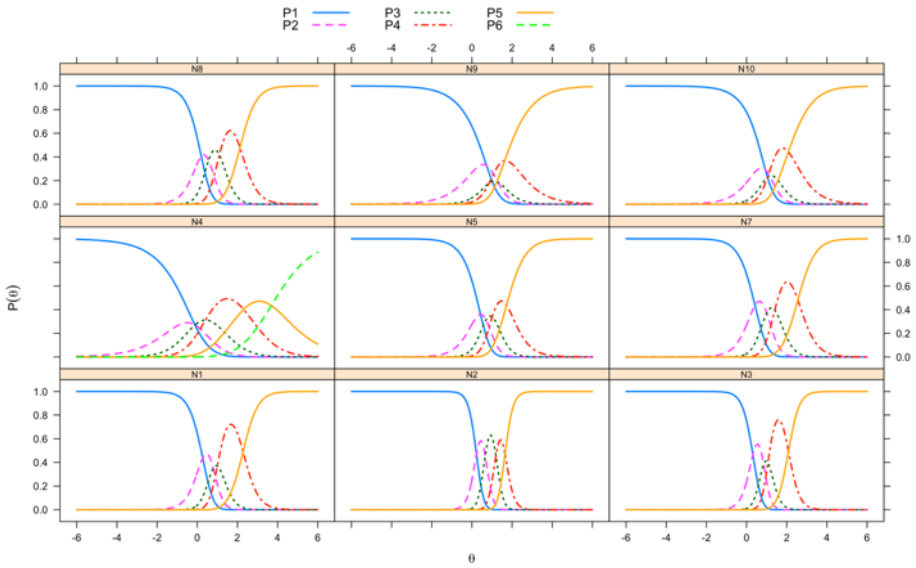
4.3.5 Item Response Curves

Vulnerable Narcissism

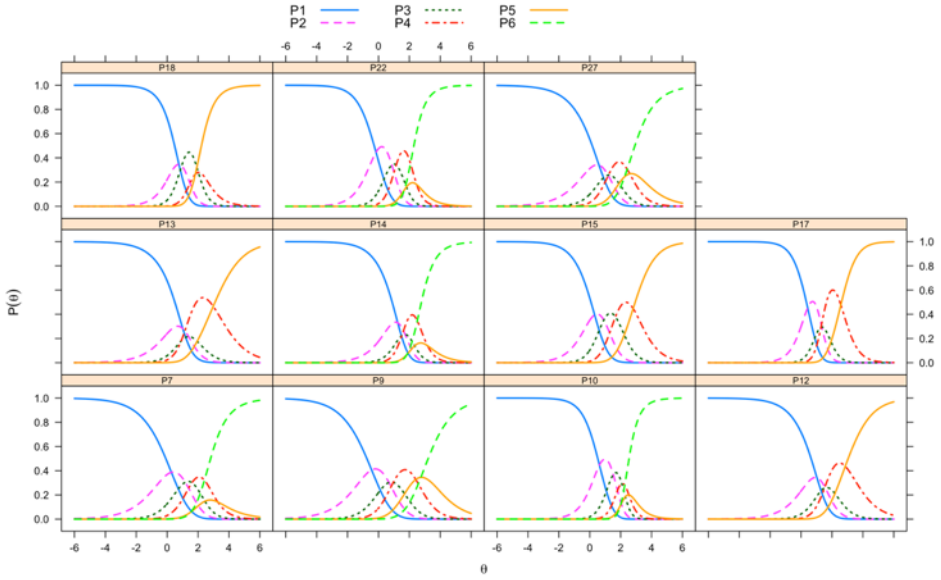
Males 26 and Over



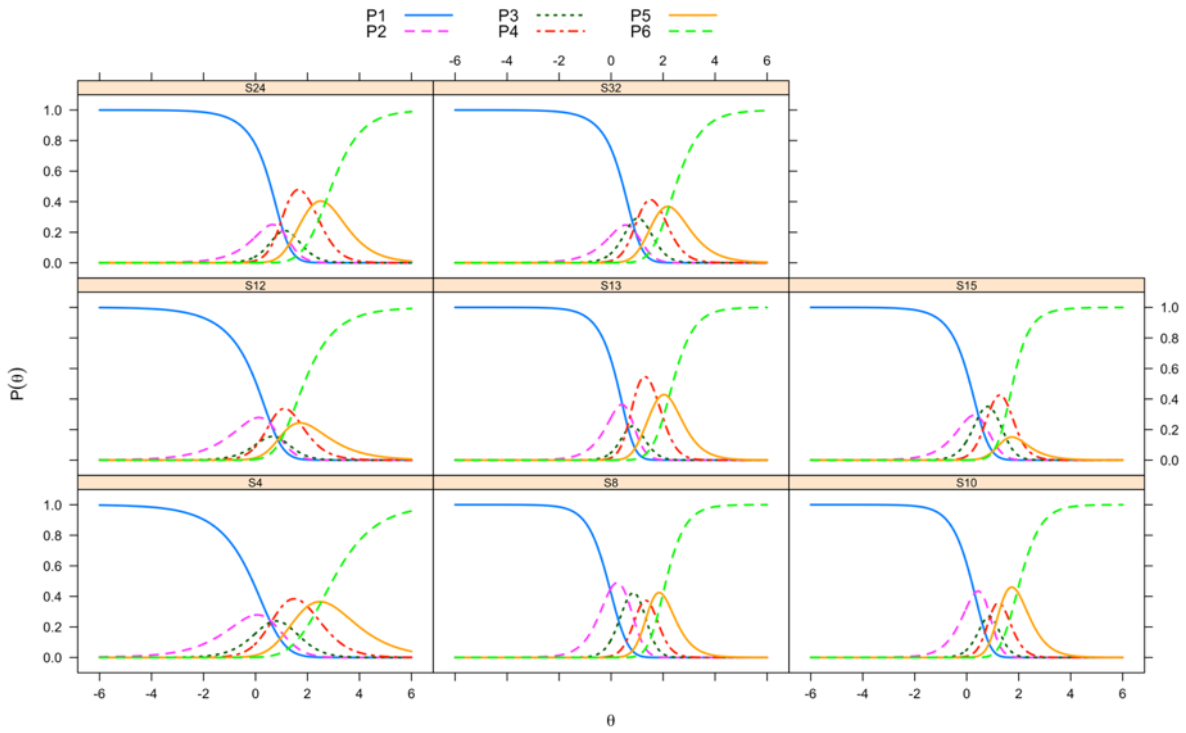
Grandiose Narcissism
Females 18-25



Psychopathy
Females 26 and Over



Everyday Sadism
Males 18-25



Appendix 4.4
Network Invariance Tests

Males and Females 18-25

```
NETWORK INVARIANCE TEST
     Test statistic M:  0.1512496
p-value 0.568

GLOBAL STRENGTH INVARIANCE TEST
     Global strength per group:  1.550353 1.703167
Test statistic S:  0.1528139
p-value 0.017
```

Females 18-25 and Males 26 and Over

```
NETWORK INVARIANCE TEST
     Test statistic M:  0.1683038
p-value 0.508

GLOBAL STRENGTH INVARIANCE TEST
     Global strength per group:  1.550353 1.587132
Test statistic S:  0.036779
p-value 0.621
```

Males and Females 26 and Over

```
NETWORK INVARIANCE TEST
     Test statistic M:  0.1698937
p-value 0.421

GLOBAL STRENGTH INVARIANCE TEST
     Global strength per group:  1.558441 1.587132
Test statistic S:  0.02869103
p-value 0.7
```

Males 18-25 and Females 26 and Over

```
NETWORK INVARIANCE TEST
     Test statistic M:  0.2404925
p-value 0.066

GLOBAL STRENGTH INVARIANCE TEST
     Global strength per group:  1.703167 1.558441
Test statistic S:  0.1447259
p-value 0.022
```

Appendix 4.5

Confirmatory Factor Analysis of EGA Models

Males 18-25

```
> summary (cfaM18.net, fit.measures=T)
Summary: Confirmatory Factor Analysis:
Length  Class   Mode
     1 lavaan     S4

 FIt Measures:
chisq.scaled           df       pvalue   cfi.scaled rmsea.scaled
     906.601      813.000        1.000        0.973        0.026
```

Females 18-25

```
> summary (cfaF18.net)
Summary: Confirmatory Factor Analysis:
Length  Class   Mode
     1 lavaan     S4

 FIt Measures:
chisq.scaled           df       pvalue   cfi.scaled rmsea.scaled
      923.55       809.00         1.00         0.95         0.03
```

Males 26 and Over

```
> summary (cfaM26.net)
Summary: Confirmatory Factor Analysis:
Length  Class   Mode
     1 lavaan     S4

 FIt Measures:
chisq.scaled           df       pvalue   cfi.scaled rmsea.scaled
     966.963      813.000        1.000        0.938        0.035
```

Females 26 and Over

```
> summary (cfaF26.net)
Summary: Confirmatory Factor Analysis:
Length  Class   Mode
     1 lavaan     S4

 FIt Measures:
chisq.scaled           df       pvalue   cfi.scaled rmsea.scaled
     932.029      809.000        1.000        0.926        0.029
```

Appendix 4.6

Confirmatory Factor Analysis Dark Side of Humanity Scale

Males 18-25

```
> summary (M18comba, fit.measures=TRUE)
lavaan 0.6-5 ended normally after 74 iterations

  Estimator                                         DWLS
  Optimization method                           NLMINB
  Number of free parameters                         90

  Number of observations                           175

Model Test User Model:
                                              Standard      Robust
  Test Statistic                               443.317     894.715
  Degrees of freedom                               813         813
  P-value (Chi-square)                           1.000       0.024
  Scaling correction factor                                  1.685
  Shift parameter                                          631.645
    for the simple second-order correction

Model Test Baseline Model:

  Test statistic                             23295.979    3850.363
  Degrees of freedom                               861         861
  P-value                                        0.000       0.000
  Scaling correction factor                                  7.505

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    1.000       0.973
  Tucker-Lewis Index (TLI)                       1.017       0.971

  Robust Comparative Fit Index (CFI)                            NA
  Robust Tucker-Lewis Index (TLI)                               NA

Root Mean Square Error of Approximation:

  RMSEA                                          0.000       0.024
  90 Percent confidence interval - lower         0.000       0.010
  90 Percent confidence interval - upper         0.000       0.033
  P-value RMSEA <= 0.05                          1.000       1.000

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                        NA
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.058       0.058

Parameter Estimates:

  Information                                 Expected
  Information saturated (h1) model         Unstructured
  Standard errors                            Robust.sem
```

Appendix 4.7

Measurement Invariance Dark Side of Humanity Scale

Males and Females 18-25

Configural Invariance

```
> summary (configuralNSMF1825, fit.measures=TRUE)
lavaan 0.6-5 ended normally after 83 iterations

  Estimator                                      DWLS
  Optimization method                           NLMINB
  Number of free parameters                        264

  Number of observations per group:
    1                                              157
    0                                              175

Model Test User Model:
                                              Standard      Robust
  Test Statistic                              1003.264    1817.400
  Degrees of freedom                          1626        1626
  P-value (Chi-square)                        1.000       0.001
  Scaling correction factor                               1.798
  Shift parameter for each group:
      1                                                   595.545
      0                                                   663.824
    for the simple second-order correction
  Test statistic for each group:
    1                                         559.948     906.997
    0                                         443.317     910.404

Model Test Baseline Model:

  Test statistic                            37498.269    6428.884
  Degrees of freedom                         1722        1722
  P-value                                    0.000       0.000
  Scaling correction factor                              7.601

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                1.000       0.959
  Tucker-Lewis Index (TLI)                   1.018       0.957

  Robust Comparative Fit Index (CFI)                        NA
  Robust Tucker-Lewis Index (TLI)                           NA

Root Mean Square Error of Approximation:

  RMSEA                                      0.000       0.027
  90 Percent confidence interval - lower     0.000       0.018
  90 Percent confidence interval - upper     0.000       0.033
  P-value RMSEA <= 0.05                      1.000       1.000

  Robust RMSEA                                              NA
  90 Percent confidence interval - lower                    NA
  90 Percent confidence interval - upper                    NA

Standardized Root Mean Square Residual:

  SRMR                                       0.063       0.063

Parameter Estimates:

  Information                                 Expected
  Information saturated (h1) model          Unstructured
  Standard errors                            Robust.sem
```

Metric Invariance

```
> summary (metricNSMF1825, fit.measures=TRUE)
lavaan 0.6-5 ended normally after 87 iterations

  Estimator                                         DWLS
  Optimization method                              NLMINB
  Number of free parameters                           264
  Number of equality constraints                       38
  Row rank of the constraints matrix                   38

  Number of observations per group:
    1                                                 157
    0                                                 175

Model Test User Model:
                                                  Standard      Robust
  Test Statistic                                  1281.875    1826.407
  Degrees of freedom                                 1664        1664
  P-value (Chi-square)                              1.000       0.003
  Scaling correction factor                                     2.369
  Shift parameter for each group:
      1                                                       607.861
      0                                                       677.553
    for the simple second-order correction
  Test statistic for each group:
    1                                              735.755     918.374
    0                                              546.120     908.033

Model Test Baseline Model:

  Test statistic                                37498.269    6428.884
  Degrees of freedom                                 1722        1722
  P-value                                           0.000       0.000
  Scaling correction factor                                     7.601

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                       1.000       0.965
  Tucker-Lewis Index (TLI)                          1.011       0.964

  Robust Comparative Fit Index (CFI)                              NA
  Robust Tucker-Lewis Index (TLI)                                 NA

Root Mean Square Error of Approximation:

  RMSEA                                             0.000       0.024
  90 Percent confidence interval - lower            0.000       0.015
  90 Percent confidence interval - upper            0.000       0.031
  P-value RMSEA <= 0.05                             1.000       1.000

  Robust RMSEA                                                    NA
  90 Percent confidence interval - lower                          NA
  90 Percent confidence interval - upper                          NA

Standardized Root Mean Square Residual:

  SRMR                                              0.070       0.070

Parameter Estimates:

  Information                                    Expected
  Information saturated (h1) model           Unstructured
  Standard errors                              Robust.sem
```

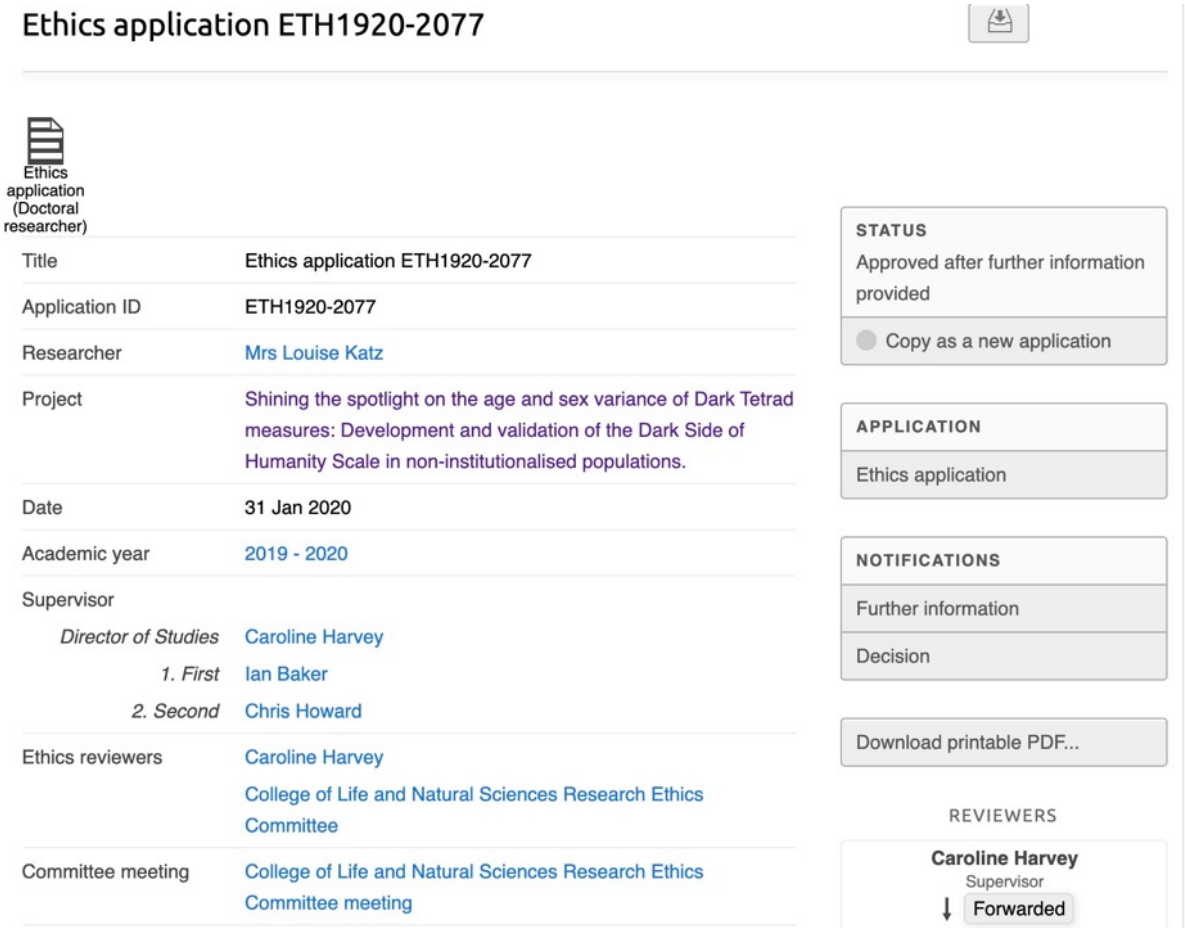
Scalar Invariance

```
> summary (scalerNSMF1825, fit.measures=TRUE)
lavaan 0.6-5 ended normally after 93 iterations

  Estimator                                         DWLS
  Optimization method                           NLMINB
  Number of free parameters                        268
  Number of equality constraints                    80
  Row rank of the constraints matrix                80

  Number of observations per group:
    1                                              157
    0                                              175

Model Test User Model:
                                              Standard      Robust
  Test Statistic                              1313.341    1869.301
  Degrees of freedom                             1702        1702
  P-value (Chi-square)                          1.000       0.003
  Scaling correction factor                                 2.372
  Shift parameter for each group:
      1                                                     622.160
      0                                                     693.490
    for the simple second-order correction
  Test statistic for each group:
    1                                           751.985     939.166
    0                                           561.357     930.135

Model Test Baseline Model:

  Test statistic                            37498.269    6428.884
  Degrees of freedom                             1722        1722
  P-value                                       0.000       0.000
  Scaling correction factor                                 7.601

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                   1.000       0.964
  Tucker-Lewis Index (TLI)                      1.011       0.964

  Robust Comparative Fit Index (CFI)                           NA
  Robust Tucker-Lewis Index (TLI)                              NA

Root Mean Square Error of Approximation:

  RMSEA                                         0.000       0.024
  90 Percent confidence interval - lower        0.000       0.015
  90 Percent confidence interval - upper        0.000       0.031
  P-value RMSEA <= 0.05                         1.000       1.000

  Robust RMSEA                                                 NA
  90 Percent confidence interval - lower                       NA
  90 Percent confidence interval - upper                       NA

Standardized Root Mean Square Residual:

  SRMR                                          0.071       0.071

Parameter Estimates:

  Information                                 Expected
  Information saturated (h1) model        Unstructured
  Standard errors                           Robust.sem
```

Appendix 5

The Construct Validity and Temporal Reliability of the Dark Side of Humanity Scale

5.1 Ethical Approval Application



Ethics application ETH1920-2077

Ethics application (Doctoral researcher)

Title	Ethics application ETH1920-2077
Application ID	ETH1920-2077
Researcher	Mrs Louise Katz
Project	Shining the spotlight on the age and sex variance of Dark Tetrad measures: Development and validation of the Dark Side of Humanity Scale in non-institutionalised populations.
Date	31 Jan 2020
Academic year	2019 - 2020
Supervisor	
Director of Studies	Caroline Harvey
1. First	Ian Baker
2. Second	Chris Howard
Ethics reviewers	Caroline Harvey College of Life and Natural Sciences Research Ethics Committee
Committee meeting	College of Life and Natural Sciences Research Ethics Committee meeting

STATUS
Approved after further information provided
Copy as a new application

APPLICATION
Ethics application

NOTIFICATIONS
Further information
Decision

Download printable PDF...

REVIEWERS
Caroline Harvey
Supervisor
Forwarded

5.1.1. Ethical Approval Confirmation

Kedleston Road, Derby
DE22 1GB, UK

T: +44 (0)1332 591060
E: researchoffice@derby.ac.uk
Sponsor License No: QGN14R294

Dear Louise

ETH1920-2077

Thank you for submitting your application to the College of Life and Natural Sciences Research Ethics Committee, which has now been reviewed and considered.

The outcome of your application is:

approved.

If any changes to the study described in the application are necessary, you must notify the Committee and may be required to make a resubmission of the application.

On behalf of the Committee, we wish you the best of luck with your study.

Yours sincerely

Charlotte Dakin

5.2 Measures Used in Chapter Six

The Mach IV can be found in Appendix 3.2.1, the Levenson Self-Report Scale in Appendix 3.2.2 and the Assessment of Sadistic Personality in 3.2.8.

5.2.1 The Dark Side of Humanity Scale

Instructions

Below you will find statements for which there are no right or wrong answers. Please consider each one and indicate how well that statement describes you.

1	2	3	4	5	6
Not at all like me	Mainly like me	A little unlike me	A little like me	Mainly like me	Very much like me

Successful Psychopathy

19. It's wise to keep track of information that I can use against people.
20. What other people feel doesn't concern me.
21. The only good reason I talk to others is to get information that I can use to my benefit.
22. I can be good at pretending to care about people but must of the time I really don't care.
23. It's sometimes fun to see how far I can push someone before they catch on.
24. Success is based on survival of the fittest, I am not concerned about the losers.
25. I could look people straight in the eye and it means nothing to me to lie or to cheat them.
26. I can simulate emotions like pain and hurt to make others feel sorry for me.
27. I am willing to be unethical if I believe it will help my plans succeed.
28. For me, what's right is whatever I can get away with.
29. I believe that lying is necessary to maintain a competitive advantage over others.
30. If I'm honest all the time it won't lead to the success of my objectives.
31. I am willing to sabotage the efforts of other people if they threaten my own goals.
32. I don't care much if what I do hurts others.
33. In today's world, I feel justified in doing anything I can get away with.
34. Playing by the rules sounds nice but getting what I want is more important.
35. I will break a promise if it works to my advantage.
36. People might describe me as mean and cruel.

Grandiose Entitlement

10. I deserve to receive special treatment.
11. I expect people to bend the rules for me.
12. I tend to expect special favours from others.
13. I deserve to get what I want.
14. I deserve more out of life than other people.
15. I don't think the rules apply to me as much as they apply to others.
16. I expect to be treated better than average.
17. I only associate with people of my calibre.
18. I do not waste my time hanging out with people who are beneath me.

Sadistic Cruelty

9. It gives me pleasure to see someone successful get fired.
10. I enjoy seeing people hurt.
11. I have fantasies which involve hurting other people.
12. Hurting people would be exciting.
13. I post offensive comments on social media forums just so I can take pleasure from the hurt I cause.
14. I would hurt somebody if it meant that I would be in control.
15. I get pleasure from mocking people in front of their friends.
16. I enjoy watching people in pain.

Entitlement Rage

Factor Four

8. It irritates me when people don't notice how good I am.
9. I get into a temper if I don't get the recognition that I deserve.
10. I can get pretty angry when others disagree with me.
11. I hate being criticised so much that I can't control my temper when it happens.
12. It really makes me angry when I don't get what I deserve.
13. I fly into a rage if somebody expects me to do tasks that are really beneath my skill level.
14. I can get really nasty if I don't get what I want.

5.2.2 Narcissism

Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979)

Instructions

Please indicate to what extent you agree or disagree with each statement.

1	2	3	4	5
Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

1. I have a natural talent for influencing people.
2. Modesty doesn't become me.
3. I would do almost anything on a dare.
4. I know that I am good because everybody keeps telling me so.
5. If I ruled the world it would be a better place.
6. I can usually talk my way out of anything.
7. I like to be the center of attention.
8. I will be a success.
9. I think I am a special person.
10. I see myself as a good leader.
11. I am assertive.
12. I like to have authority over other people.
13. I find it easy to manipulate people.
14. I insist on getting the respect that is due me.
15. I like to show off my body.
16. I can read people like a book.
17. I like to take responsibility for making decisions.
18. I want to amount to something in the eyes of the world.
19. I like to look at my body.
20. I will usually show off if I get the chance.
21. I always know what I'm doing.
22. I rarely depend on anyone else to get things done.
23. Everybody likes to hear my stories.
24. I expect a great deal from other people.
25. I will never be satisfied until I get all that I deserve.
26. I like to be complimented.
27. I have a strong will to power.

28. I like to start new fads and fashions.
29. I like to look at myself in the mirror.
30. I really like to be the center of attention.
31. I can live my life in any way I want to.
32. People always seem to recognise my authority.
33. I would prefer to be a leader.
34. I am going to be a great person.
35. I can make anybody believe anything I want them to.
36. I am a born leader.
37. I wish somebody would someday write my biography.
38. I get upset when people don’t notice how I look when I go out in public.
39. I am more capable than other people.
40. I am an extraordinary person.

5.2.3 Big Five Personality Traits

The Big Five Inventory-2 (BFI-2)

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please indicate the extent to which you agree or disagree with the following statements.

1	2	3	4	5
Disagree strongly	Disagree a little	Neutral: No opinion	Agree a little	Agree strongly

I am someone who

1. Is outgoing, sociable.
2. Is compassionate, has a soft heart.
3. Tends to be disorganised.
4. Is relaxed, handles stress well.
5. Has few artistic interests.
6. Has an assertive personality.
7. Is respectful, treats others with respect.
8. Tends to be lazy.
9. Stays optimistic after experiencing a setback.
10. Is curious about many different things.
11. Rarely feels excited or eager.
12. Tends to find fault with others.
13. Is dependable, steady.
14. Is moody, has up and down mood swings.
15. Is inventive, finds clever ways to do things.
16. Tends to be quiet.
17. Feels little sympathy for others.
18. Is systematic, likes to keep things in order.
19. Can be tense.
20. Is fascinated by art, music or literature.
21. Is dominant, acts as a leader.
22. Starts arguments with others.
23. Has difficulty getting started on tasks.
24. Feels secure, comfortable with self.
25. Avoids intellectual, philosophical discussions.
26. Is less active than other people.
27. Has a forgiving nature.

28. Can be somewhat careless.
29. Is emotionally stable, not easily upset.
30. Has little creativity.
31. Is sometimes shy, introverted.
32. Is helpful and unselfish with others,
33. Keeps things neat and tidy.
34. Worries a lot.
35. Values art and beauty.
36. Finds it hard to influence people.
37. Is sometimes rude to others.
38. Is efficient, gets things done,
39. Often feels sad.
40. Is complex, a deep thinker.
41. Is full of energy.
42. Is suspicious of other intentions.
43. Is reliable, can always be counted on.
44. Keeps their emotions under control,
45. Has difficulty imagining things.
46. Is talkative.
47. Can be cold and uncaring.
48. Leaves a mess, doesn't clean up.
49. Rarely feels anxious or afraid.
50. Thinks poetry and plays are boring.
51. Prefers to have others take charge.
52. Is polite, courteous to others.
53. Is persistent, works until the task is finished.
54. Tends to feel depressed, blue.
55. Has little interest in abstract ideas.
56. Shows a lot of enthusiasm.
57. Assumes the best about people.
58. Sometimes behaves irresponsibly.
59. Is temperamental, gets emotional easily.
60. Is original, comes up with new ideas.

5.2.4 Self-Esteem

Rosenberg Self-Esteem Scale (Rosenberg, 1965)

Please record the appropriate answer for each item, depending on whether you Strongly agree, agree, disagree, or strongly disagree with it.

1	2	3	4
Strongly Agree	Agree	Disagree	Strongly Disagree

1. On the whole I am satisfied with myself.
2. At times, I think I am no good at all.
3. I feel that I have a number of good qualities.
4. I am able to do things as well as most other people.
5. I feel I do not have much to be proud of.
6. I certainly feel useless at times.
7. I feel that I'm a person of worth.
8. I wish I could have more respect for myself.
9. All in all, I am inclined to think that I am a failure.
10. I take a positive attitude toward myself.

Appendix 5.3

Confirmatory Factor Analysis of the Dark Side of Humanity Scale.

Females 26 and Over

```
> summary (F26S3, fit.measures=T)
lavaan 0.6-7 ended normally after 70 iterations

  Estimator                                      DWLS
  Optimization method                           NLMINB
  Number of free parameters                         90

  Number of observations                           183

Model Test User Model:
                                              Standard      Robust
  Test Statistic                               465.960     905.720
  Degrees of freedom                               813         813
  P-value (Chi-square)                           1.000       0.013
  Scaling correction factor                                  1.681
  Shift parameter                                          628.604
       simple second-order correction

Model Test Baseline Model:

  Test statistic                             13474.677    2296.239
  Degrees of freedom                               861         861
  P-value                                        0.000       0.000
  Scaling correction factor                                  8.789

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    1.000       0.935
  Tucker-Lewis Index (TLI)                       1.029       0.932

  Robust Comparative Fit Index (CFI)                            NA
  Robust Tucker-Lewis Index (TLI)                               NA

Root Mean Square Error of Approximation:

  RMSEA                                          0.000       0.025
  90 Percent confidence interval - lower         0.000       0.013
  90 Percent confidence interval - upper         0.000       0.034
  P-value RMSEA <= 0.05                          1.000       1.000

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                        NA
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.074       0.074
```

Males 26 and Over

```
> summary (M26S3, fit.measures=T)
lavaan 0.6-7 ended normally after 70 iterations

  Estimator                                         DWLS
  Optimization method                           NLMINB
  Number of free parameters                         90

  Number of observations                           178

Model Test User Model:
                                              Standard      Robust
  Test Statistic                               482.864     927.473
  Degrees of freedom                               813         813
  P-value (Chi-square)                           1.000       0.003
  Scaling correction factor                                  1.524
  Shift parameter                                          610.562
       simple second-order correction

Model Test Baseline Model:

  Test statistic                             14374.001    2269.812
  Degrees of freedom                               861         861
  P-value                                        0.000       0.000
  Scaling correction factor                                  9.592

Jser Model versus Baseline Model:

  Comparative Fit Index (CFI)                    1.000       0.919
  Tucker-Lewis Index (TLI)                       1.026       0.914

  Robust Comparative Fit Index (CFI)                            NA
  Robust Tucker-Lewis Index (TLI)                               NA

Root Mean Square Error of Approximation:

  RMSEA                                          0.000       0.028
  90 Percent confidence interval - lower         0.000       0.017
  90 Percent confidence interval - upper         0.000       0.037
  P-value RMSEA <= 0.05                          1.000       1.000

  Robust RMSEA                                                  NA
  90 Percent confidence interval - lower                        NA
  90 Percent confidence interval - upper                        NA

Standardized Root Mean Square Residual:

  SRMR                                           0.067       0.067
```

Appendix 5.4

Robust Correlations
Convergent and Discriminant Validity

Females 18-25

```
pball(x = NPIConvF18)

Robust correlation matrix:
                    CM     Ent      SP      ER PrimaryPsych SecondaryPsych
CM              1.0000  0.5893  0.6469  0.5641       0.7106         0.5391
Ent             0.5893  1.0000  0.5295  0.7344       0.4619         0.4052
SP              0.6469  0.5295  1.0000  0.5230       0.5684         0.3872
ER              0.5641  0.7344  0.5230  1.0000       0.4177         0.4819
PrimaryPsych    0.7106  0.4619  0.5684  0.4177       1.0000         0.4589
SecondaryPsych  0.5391  0.4052  0.3872  0.4819       0.4589         1.0000
NPITotal        0.3373  0.4841  0.2106  0.4138       0.3004         0.1531
MachTotal       0.6450  0.4363  0.5410  0.4600       0.6201         0.3987
ASP             0.5805  0.4509  0.6144  0.4221       0.5518         0.4132
Extraversion   -0.1069  0.0389 -0.1907 -0.0355      -0.0759        -0.1432
Agreeableness  -0.5434 -0.3366 -0.4753 -0.4045      -0.4843        -0.5047
Conscientious  -0.3291 -0.2450 -0.2637 -0.2636      -0.2912        -0.5556
NegEmo          0.1574  0.0411  0.2741  0.3062       0.0414         0.3189
OpenMind       -0.1552  0.0209 -0.0813  0.0773      -0.2819        -0.1801
SelfEsteem      0.2511  0.0536  0.2139  0.1856       0.1510         0.3327
NPISuperiority  0.1912  0.3851  0.1169  0.3938       0.2022         0.0882
NPIExploit      0.3607  0.4525  0.2171  0.4220       0.3121         0.1757
NPIEntitle      0.2636  0.4080  0.1864  0.4216       0.2056         0.1848
                NPITotal MachTotal     ASP Extraversion Agreeableness
CM                0.3373    0.6450  0.5805      -0.1069       -0.5434
Ent               0.4841    0.4363  0.4509       0.0389       -0.3366
SP                0.2106    0.5410  0.6144      -0.1907       -0.4753
ER                0.4138    0.4600  0.4221      -0.0355       -0.4045
PrimaryPsych      0.3004    0.6201  0.5518      -0.0759       -0.4843
SecondaryPsych    0.1531    0.3987  0.4132      -0.1432       -0.5047
NPITotal          1.0000    0.2333  0.2863       0.5692       -0.1443
MachTotal         0.2333    1.0000  0.5592      -0.1492       -0.6422
ASP               0.2863    0.5592  1.0000       0.0061       -0.5227
Extraversion      0.5692   -0.1492  0.0061       1.0000        0.1480
Agreeableness    -0.1443   -0.6422 -0.5227       0.1480        1.0000
Conscientious     0.0124   -0.3142 -0.3208       0.2496        0.4269
NegEmo           -0.2414    0.2392  0.2172      -0.3964       -0.2881
OpenMind          0.2178   -0.1041 -0.1647       0.1950        0.2329
SelfEsteem       -0.3863    0.2330  0.2092      -0.5079       -0.3202
NPISuperiority    0.7494    0.1181  0.1127       0.3780       -0.0155
NPIExploit        0.7778    0.2322  0.3272       0.4367       -0.2307
NPIEntitle        0.7546    0.2589  0.1830       0.3454       -0.1381
                Conscientious  NegEmo OpenMind SelfEsteem NPISuperiority
CM                    -0.3291  0.1574  -0.1552     0.2511         0.1912
Ent                   -0.2450  0.0411   0.0209     0.0536         0.3851
SP                    -0.2637  0.2741  -0.0813     0.2139         0.1169
ER                    -0.2636  0.3062   0.0773     0.1856         0.3938
PrimaryPsych          -0.2912  0.0414  -0.2819     0.1510         0.2022
SecondaryPsych        -0.5556  0.3189  -0.1801     0.3327         0.0882
NPITotal               0.0124 -0.2414   0.2178    -0.3863         0.7494
MachTotal             -0.3142  0.2392  -0.1041     0.2330         0.1181
ASP                   -0.3208  0.2172  -0.1647     0.2092         0.1127
Extraversion           0.2496 -0.3964   0.1950    -0.5079         0.3780
Agreeableness          0.4269 -0.2881   0.2329    -0.3202        -0.0155
Conscientious          1.0000 -0.2927   0.1453    -0.3637        -0.0049
NegEmo                -0.2927  1.0000  -0.0204     0.6092        -0.1471
OpenMind               0.1453 -0.0204   1.0000    -0.1696         0.2267
SelfEsteem            -0.3637  0.6092  -0.1696     1.0000        -0.3711
NPISuperiority        -0.0049 -0.1471   0.2267    -0.3711         1.0000
NPIExploit            -0.0912 -0.1271   0.1991    -0.1451         0.4666
NPIEntitle             0.0723 -0.0903   0.2049    -0.2247         0.5587
```

p-values:

	CM	Ent	SP	ER	PrimaryPsych	SecondaryPsych
CM	NA	0.00000	0.00000	0.00000	0.00000	0.00000
Ent	0.00000	NA	0.00000	0.00000	0.00000	0.00000
SP	0.00000	0.00000	NA	0.00000	0.00000	0.00000
ER	0.00000	0.00000	0.00000	NA	0.00000	0.00000
PrimaryPsych	0.00000	0.00000	0.00000	0.00000	NA	0.00000
SecondaryPsych	0.00000	0.00000	0.00000	0.00000	0.00000	NA
NPITotal	0.00001	0.00000	0.00662	0.00000	0.00009	0.04960
MachTotal	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
ASP	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Extraversion	0.17171	0.61945	0.01417	0.65088	0.33228	0.06651
Agreeableness	0.00000	0.00001	0.00000	0.00000	0.00000	0.00000
Conscientious	0.00002	0.00151	0.00062	0.00062	0.00015	0.00000
NegEmo	0.04354	0.60000	0.00037	0.00006	0.59766	0.00003
OpenMind	0.04651	0.78940	0.29940	0.32384	0.00024	0.02062
SelfEsteem	0.00114	0.49427	0.00579	0.01702	0.05282	0.00001
NPISuperiority	0.01391	0.00000	0.13477	0.00000	0.00919	0.25967
NPIExploit	0.00000	0.00000	0.00509	0.00000	0.00004	0.02402
NPIEntitle	0.00062	0.00000	0.01651	0.00000	0.00808	0.01747

	NPITotal	MachTotal	ASP	Extraversion	Agreeableness
CM	0.00001	0.00000	0.00000	0.17171	0.00000
Ent	0.00000	0.00000	0.00000	0.61945	0.00001
SP	0.00662	0.00000	0.00000	0.01417	0.00000
ER	0.00000	0.00000	0.00000	0.65088	0.00000
PrimaryPsych	0.00009	0.00000	0.00000	0.33228	0.00000
SecondaryPsych	0.04960	0.00000	0.00000	0.06651	0.00000
NPITotal	NA	0.00256	0.00019	0.00000	0.06441
MachTotal	0.00256	NA	0.00000	0.05585	0.00000
ASP	0.00019	0.00000	NA	0.93791	0.00000
Extraversion	0.00000	0.05585	0.93791	NA	0.05776
Agreeableness	0.06441	0.00000	0.00000	0.05776	NA
Conscientious	0.87414	0.00004	0.00003	0.00123	0.00000
NegEmo	0.00178	0.00197	0.00507	0.00000	0.00017
OpenMind	0.00494	0.18324	0.03448	0.01209	0.00261
SelfEsteem	0.00000	0.00260	0.00700	0.00000	0.00003
NPISuperiority	0.00000	0.13100	0.14968	0.00000	0.84330
NPIExploit	0.00000	0.00269	0.00002	0.00000	0.00287
NPIEntitle	0.00000	0.00078	0.01864	0.00001	0.07683

	Conscientious	NegEmo	OpenMind	SelfEsteem	NPISuperiority
CM	0.00002	0.04354	0.04651	0.00114	0.01391
Ent	0.00151	0.60000	0.78940	0.49427	0.00000
SP	0.00062	0.00037	0.29940	0.00579	0.13477
ER	0.00062	0.00006	0.32384	0.01702	0.00000
PrimaryPsych	0.00015	0.59766	0.00024	0.05282	0.00919
SecondaryPsych	0.00000	0.00003	0.02062	0.00001	0.25967
NPITotal	0.87414	0.00178	0.00494	0.00000	0.00000
MachTotal	0.00004	0.00197	0.18324	0.00260	0.13100
ASP	0.00003	0.00507	0.03448	0.00700	0.14968
Extraversion	0.00123	0.00000	0.01209	0.00000	0.00000
Agreeableness	0.00000	0.00017	0.00261	0.00003	0.84330
Conscientious	NA	0.00014	0.06262	0.00000	0.95050
NegEmo	0.00014	NA	0.79444	0.00000	0.05936
OpenMind	0.06262	0.79444	NA	0.02941	0.00341
SelfEsteem	0.00000	0.00000	0.02941	NA	0.00000
NPISuperiority	0.95050	0.05936	0.00341	0.00000	NA
NPIExploit	0.24376	0.10381	0.01034	0.06293	0.00000
NPIEntitle	0.35616	0.24873	0.00828	0.00371	0.00000

Males 26 and over

```
Call:
pball(x = NPIConv)

Robust correlation matrix:
                    CM     Ent      SP      ER PrimaryPsych SecondaryPsych
CM              1.0000  0.6304  0.5979  0.5685       0.7570         0.4509
Ent             0.6304  1.0000  0.4643  0.6559       0.6819         0.3834
SP              0.5979  0.4643  1.0000  0.4498       0.5030         0.3907
ER              0.5685  0.6559  0.4498  1.0000       0.5039         0.4752
PrimaryPsych    0.7570  0.6819  0.5030  0.5039       1.0000         0.4590
SecondaryPsych  0.4509  0.3834  0.3907  0.4752       0.4590         1.0000
NPITotal        0.3035  0.5478  0.2483  0.4136       0.4460         0.1414
MachTotal       0.5904  0.4870  0.4408  0.4848       0.6855         0.4726
ASP             0.5915  0.5030  0.6445  0.4248       0.6050         0.3651
Extraversion   -0.0092  0.1168 -0.0499  0.0025       0.1041        -0.1649
Agreeableness  -0.6595 -0.4958 -0.4514 -0.4513      -0.5857        -0.4774
Conscientious  -0.2507 -0.0720 -0.2997 -0.1540      -0.1736        -0.5443
NegEmo          0.0006 -0.0306  0.1402  0.2040      -0.0793         0.3987
OpenMind       -0.1795 -0.0886 -0.2247 -0.0303      -0.2159        -0.2432
SelfEsteem      0.0623 -0.0924  0.1533  0.0254      -0.0139         0.4418
NPISuperiority  0.1393  0.4384  0.1694  0.3394       0.3196         0.1479
NPIExploit      0.4339  0.5019  0.3889  0.4748       0.4633         0.3361
NPIEntitle      0.3024  0.5503  0.2236  0.5086       0.4054         0.2736
                NPITotal MachTotal     ASP Extraversion Agreeableness
CM                0.3035    0.5904  0.5915      -0.0092       -0.6595
Ent               0.5478    0.4870  0.5030       0.1168       -0.4958
SP                0.2483    0.4408  0.6445      -0.0499       -0.4514
ER                0.4136    0.4848  0.4248       0.0025       -0.4513
PrimaryPsych      0.4460    0.6855  0.6050       0.1041       -0.5857
SecondaryPsych    0.1414    0.4726  0.3651      -0.1649       -0.4774
NPITotal          1.0000    0.2969  0.3954       0.5740       -0.2368
MachTotal         0.2969    1.0000  0.4391      -0.0660       -0.6044
ASP               0.3954    0.4391  1.0000       0.0922       -0.6065
Extraversion      0.5740   -0.0660  0.0922       1.0000        0.1411
Agreeableness    -0.2368   -0.6044 -0.6065       0.1411        1.0000
Conscientious     0.2542   -0.2539 -0.2948       0.4480        0.3216
NegEmo           -0.1632    0.1183  0.1372      -0.3483       -0.2823
OpenMind          0.0763   -0.1612 -0.3459       0.2109        0.3349
SelfEsteem       -0.3607    0.1496  0.1957      -0.4933       -0.2911
NPISuperiority    0.7525    0.2189  0.2333       0.3289       -0.0649
NPIExploit        0.7843    0.4366  0.4576       0.3865       -0.4170
NPIEntitle        0.7985    0.3619  0.3518       0.3867       -0.3019
                Conscientious  NegEmo OpenMind SelfEsteem NPISuperiority
CM                    -0.2507  0.0006  -0.1795     0.0623         0.1393
Ent                   -0.0720 -0.0306  -0.0886    -0.0924         0.4384
SP                    -0.2997  0.1402  -0.2247     0.1533         0.1694
ER                    -0.1540  0.2040  -0.0303     0.0254         0.3394
PrimaryPsych          -0.1736 -0.0793  -0.2159    -0.0139         0.3196
SecondaryPsych        -0.5443  0.3987  -0.2432     0.4418         0.1479
NPITotal               0.2542 -0.1632   0.0763    -0.3607         0.7525
MachTotal             -0.2539  0.1183  -0.1612     0.1496         0.2189
ASP                   -0.2948  0.1372  -0.3459     0.1957         0.2333
Extraversion           0.4480 -0.3483   0.2109    -0.4933         0.3289
Agreeableness          0.3216 -0.2823   0.3349    -0.2911        -0.0649
Conscientious          1.0000 -0.3923   0.2910    -0.6289         0.1327
NegEmo                -0.3923  1.0000  -0.1122     0.6070        -0.1034
OpenMind               0.2910 -0.1122   1.0000    -0.3114         0.1636
SelfEsteem            -0.6289  0.6070  -0.3114     1.0000        -0.3046
NPISuperiority         0.1327 -0.1034   0.1636    -0.3046         1.0000
NPIExploit            -0.0167 -0.0193   0.0179    -0.1083         0.4900
NPIEntitle             0.1424  0.0489   0.0798    -0.1630         0.5833
```

```
p-values:
                    CM      Ent       SP       ER PrimaryPsych SecondaryPsych
CM                  NA 0.00000 0.00000 0.00000       0.00000         0.00000
Ent            0.00000      NA 0.00000 0.00000       0.00000         0.00000
SP             0.00000 0.00000      NA 0.00000       0.00000         0.00000
ER             0.00000 0.00000 0.00000       NA       0.00000         0.00000
PrimaryPsych   0.00000 0.00000 0.00000 0.00000            NA         0.00000
SecondaryPsych 0.00000 0.00000 0.00000 0.00000       0.00000              NA
NPITotal       0.00004 0.00000 0.00083 0.00000       0.00000         0.05975
MachTotal      0.00000 0.00000 0.00000 0.00000       0.00000         0.00000
ASP            0.00000 0.00000 0.00000 0.00000       0.00000         0.00000
Extraversion   0.90342 0.12035 0.50861 0.97376       0.16666         0.02785
Agreeableness  0.00000 0.00000 0.00000 0.00000       0.00000         0.00000
Conscientious  0.00074 0.33941 0.00005 0.04011       0.02049         0.00000
NegEmo         0.99378 0.68521 0.06194 0.00630       0.29284         0.00000
OpenMind       0.01653 0.23942 0.00257 0.68842       0.00380         0.00107
SelfEsteem     0.40865 0.21983 0.04102 0.73614       0.85388         0.00000
NPISuperiority 0.06373 0.00000 0.02382 0.00000       0.00001         0.04885
NPIExploit     0.00000 0.00000 0.00000 0.00000       0.00000         0.00000
NPIEntitle     0.00004 0.00000 0.00270 0.00000       0.00000         0.00022
               NPITotal MachTotal     ASP Extraversion Agreeableness
CM              0.00004   0.00000 0.00000      0.90342       0.00000
Ent             0.00000   0.00000 0.00000      0.12035       0.00000
SP              0.00083   0.00000 0.00000      0.50861       0.00000
ER              0.00000   0.00000 0.00000      0.97376       0.00000
PrimaryPsych    0.00000   0.00000 0.00000      0.16666       0.00000
SecondaryPsych  0.05975   0.00000 0.00000      0.02785       0.00000
NPITotal             NA   0.00006 0.00000      0.00000       0.00146
MachTotal       0.00006        NA 0.00000      0.38132       0.00000
ASP             0.00000   0.00000      NA      0.22077       0.00000
Extraversion    0.00000   0.38132 0.22077           NA       0.06023
Agreeableness   0.00146   0.00000 0.00000      0.06023            NA
Conscientious   0.00062   0.00063 0.00006      0.00000       0.00001
NegEmo          0.02948   0.11588 0.06781      0.00000       0.00013
OpenMind        0.31171   0.03159 0.00000      0.00472       0.00000
SelfEsteem      0.00000   0.04620 0.00885      0.00000       0.00008
NPISuperiority  0.00000   0.00333 0.00172      0.00001       0.38953
NPIExploit      0.00000   0.00000 0.00000      0.00000       0.00000
NPIEntitle      0.00000   0.00000 0.00000      0.00000       0.00004
               Conscientious  NegEmo OpenMind SelfEsteem NPISuperiority
CM                   0.00074 0.99378  0.01653    0.40865        0.06373
Ent                  0.33941 0.68521  0.23942    0.21983        0.00000
SP                   0.00005 0.06194  0.00257    0.04102        0.02382
ER                   0.04011 0.00630  0.68842    0.73614        0.00000
PrimaryPsych         0.02049 0.29284  0.00380    0.85388        0.00001
SecondaryPsych       0.00000 0.00000  0.00107    0.00000        0.04885
NPITotal             0.00062 0.02948  0.31171    0.00000        0.00000
MachTotal            0.00063 0.11588  0.03159    0.04620        0.00333
ASP                  0.00006 0.06781  0.00000    0.00885        0.00172
Extraversion         0.00000 0.00000  0.00472    0.00000        0.00001
Agreeableness        0.00001 0.00013  0.00000    0.00008        0.38953
Conscientious             NA 0.00000  0.00008    0.00000        0.07751
NegEmo               0.00000      NA  0.13607    0.00000        0.16951
OpenMind             0.00008 0.13607       NA    0.00002        0.02911
SelfEsteem           0.00000 0.00000  0.00002         NA        0.00004
NPISuperiority       0.07751 0.16951  0.02911    0.00004             NA
NPIExploit           0.82522 0.79811  0.81278    0.15029        0.00000
NPIEntitle           0.05797 0.51680  0.28951    0.02969        0.00000
```

5.5 Robust Test Re-Test Correlations

Males 18-25

```
pball(x = T1T2M18)

Robust correlation matrix:
        CMT1  ENTT1   SPT1   ERT1   CMT2  ENTT2   SPT2   ERT2
CMT1  1.0000 0.5740 0.5820 0.4572 0.6378 0.3382 0.4020 0.3634
ENTT1 0.5740 1.0000 0.5190 0.6726 0.3513 0.6150 0.3595 0.5148
SPT1  0.5820 0.5190 1.0000 0.4462 0.3815 0.2889 0.5446 0.3114
ERT1  0.4572 0.6726 0.4462 1.0000 0.2950 0.4114 0.3499 0.5261
CMT2  0.6378 0.3513 0.3815 0.2950 1.0000 0.6212 0.7104 0.5596
ENTT2 0.3382 0.6150 0.2889 0.4114 0.6212 1.0000 0.5880 0.6736
SPT2  0.4020 0.3595 0.5446 0.3499 0.7104 0.5880 1.0000 0.5348
ERT2  0.3634 0.5148 0.3114 0.5261 0.5596 0.6736 0.5348 1.0000

p-values:
         CMT1   ENTT1    SPT1    ERT1    CMT2   ENTT2    SPT2    ERT2
CMT1       NA 0.00000 0.00000 0.00000 0.00000 0.00062 0.00004 0.00022
ENTT1 0.00000      NA 0.00000 0.00000 0.00036 0.00000 0.00026 0.00000
SPT1  0.00000 0.00000      NA 0.00000 0.00010 0.00373 0.00000 0.00170
ERT1  0.00000 0.00000 0.00000      NA 0.00303 0.00002 0.00038 0.00000
CMT2  0.00000 0.00036 0.00010 0.00303      NA 0.00000 0.00000 0.00000
ENTT2 0.00062 0.00000 0.00373 0.00002 0.00000      NA 0.00000 0.00000
SPT2  0.00004 0.00026 0.00000 0.00038 0.00000 0.00000      NA 0.00000
ERT2  0.00022 0.00000 0.00170 0.00000 0.00000 0.00000 0.00000      NA
```

Females 26 and Over

```
pball(x = T1T2F26)

Robust correlation matrix:
        CMT1  ENTT1   SPT1   ERT1   CMT2  ENTT2   SPT2   ERT2
CMT1  1.0000 0.6399 0.6654 0.6589 0.6687 0.4042 0.4536 0.5099
ENTT1 0.6399 1.0000 0.5460 0.7430 0.3801 0.5838 0.2284 0.5332
SPT1  0.6654 0.5460 1.0000 0.6005 0.5024 0.2658 0.4405 0.2963
ERT1  0.6589 0.7430 0.6005 1.0000 0.4090 0.4767 0.2987 0.6625
CMT2  0.6687 0.3801 0.5024 0.4090 1.0000 0.5761 0.6906 0.5665
ENTT2 0.4042 0.5838 0.2658 0.4767 0.5761 1.0000 0.4749 0.7046
SPT2  0.4536 0.2284 0.4405 0.2987 0.6906 0.4749 1.0000 0.4344
ERT2  0.5099 0.5332 0.2963 0.6625 0.5665 0.7046 0.4344 1.0000

p-values:
       CMT1   ENTT1    SPT1    ERT1  CMT2   ENTT2    SPT2    ERT2
CMT1     NA 0.00000 0.00000 0.00000 0e+00 0.00002 0.00000 0.00000
ENTT1 0e+00      NA 0.00000 0.00000 5e-05 0.00000 0.01797 0.00000
SPT1  0e+00 0.00000      NA 0.00000 0e+00 0.00565 0.00000 0.00194
ERT1  0e+00 0.00000 0.00000      NA 1e-05 0.00000 0.00178 0.00000
CMT2  0e+00 0.00005 0.00000 0.00001    NA 0.00000 0.00000 0.00000
ENTT2 2e-05 0.00000 0.00565 0.00000 0e+00      NA 0.00000 0.00000
SPT2  0e+00 0.01797 0.00000 0.00178 0e+00 0.00000      NA 0.00000
ERT2  0e+00 0.00000 0.00194 0.00000 0e+00 0.00000 0.00000      NA
```

5.6. Robust T-Test

Females 26 and Over
Entitlement Rage

```
yuen(formula = ERT1 ~ ERT2, data = T1T2FemalesOver26)

Test statistic: 3.7363 (df = 12.05), p-value = 0.00282

Trimmed mean difference:  -1.67857
95 percent confidence interval:
-2.657     -0.7002

Explanatory measure of effect size: 0.64
```

References In Support of the Appendices

Boduszek, D., Debowska, A., Dhingra, K., & DeLisi, M. (2016). Introduction and validation of Psychopathic Personality Traits Scale (PPTS) in a large prison sample. *Journal of Criminal Justice*, *46*, 9-17.

Buckels, E. E., Trapnell, P. D., Andjelovic, T., & Paulhus, D. L. (2019). Internet trolling and everyday sadism: Parallel effects on pain perception and moral judgment. *Journal of Personality*, *87*(2), 328-340.

Cheek, J. M., Hendin, H. M., & Wink, P. M. (2013). An expanded version of the hypersensitive narcissism scale: The maladaptive covert narcissism scale. In *meeting of the Association for Research in Personality, Charlotte, NC*.

Christian, E., & Sellbom, M. (2016). Development and validation of an expanded version of the three-factor Levenson Self-Report Psychopathy Scale. *Journal of Personality Assessment*, *98*(2), 155-168.

Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. NY, USA: Academic Press.

Collison, K. L., Vize, C. E., Miller, J. D., & Lynam, D. R. (2018). Development and preliminary validation of a five-factor model measure of Machiavellianism. *Psychological Assessment*, *30*(10), 1401.

Crysel, L. C., & Webster, G. D. (2018). Schadenfreude and the spread of political misfortune. *PloS one*, *13*(9).

Dahling, J. J., Whitaker, B. G., & Levy, P. E. (2009). The development and validation of a new Machiavellianism scale. *Journal of Management*, *35*(2), 219-257.

Ferris, G. R., Treadway, D. C., Kolodinsky, R. W., Hochwarter, W. A., Kacmar, C. J., Douglas, C., & Frink, D. D. (2005). Development and validation of the political skill inventory. *Journal of Management*, *31*(1), 126-152.

Foster, J. D., McCain, J. L., Hibberts, M. F., Brunell, A. B., & Johnson, R. B. (2015). The Grandiose Narcissism Scale: A global and facet-level measure of grandiose narcissism. *Personality and Individual Differences*, *73*, 12-16.

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology*, *21*(3), 913-938.

Pincus, A. L., Ansell, E. B., Pimentel, C. A., Cain, N. M., Wright, A. G., & Levy, K. N. (2009). Initial construction and validation of the Pathological Narcissism Inventory. *Psychological Assessment*, *21*(3), 365.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741.

Rohmann, E., Neumann, E., Herner, M. J., & Bierhoff, H. (2015). Grandiose and vulnerable narcissism: Self-construal, attachment, and love in romantic relationships. *European Psychologist, 17*, 279–290.

Semel R. A. (2018). The Abbreviated Psychopathy Measure (APM): Two pilot studies of a brief measure of psychopathy traits. *International Journal of Psychological Studies, 10.*

Sest, N., & March, E. (2017). Constructing the cyber-troll: Psychopathy, sadism, and empathy. *Personality and Individual Differences*, *119*, 69-72.

Sherman, E. D., Miller, J. D., Few, L. R., Campbell, W. K., Widiger, T. A., Crego, C., & Lynam, D. R. (2015). Development of a Short Form of the Five-Factor Narcissism Inventory: The FFNI-SF. *Psychological Assessment*, *27*(3), 1110.

Smith, R. J. (1985). The concept and measurement of social psychopathy. *Journal of Research in Personality, 19,* 219-231.

Williams, K. M., Paulhus, D. L., & Hare, R. D. (2007). Capturing the four-factor structure of psychopathy in college students via self-report. *Journal of Personality Assessment, 88*(2), 205-219.